

LinearRegressionGenderScopus2minimalRobustLoop.R

MikeT

Wed Jul 17 07:23:04 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 <- 1300:1315 #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\\Brazil 1996-2018
328 fields\\cov3\\ScopusFind Brazil"; Country <- "Br"
#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\\France 1996-2018
329 fields\\cov3\\ScopusFind France"; Country <- "Fr"
#RootFileName <- "E:\\data\\Scopus\\All fields regression\\Germany 1996-2018
330 fields\\cov3\\ScopusFind Germany"; Country <- "De"
#RootFileName <- "E:\\data\\Scopus\\All fields regression\\India 1996-2018
329 fields\\cov3\\ScopusFind India"; Country <- "In"

```

```

RootFileName <- "E:\\data\\Scopus\\All fields regression\\Ireland 1996-2018
329 fields\\cov3\\ScopusFind Ireland"; Country <- "Ie"
#RootFileName <- "E:\\data\\Scopus\\All fields regression\\Mexico 1996-2018
326 fields\\cov3\\ScopusFind Mexico"; Country <- "Mx"
#RootFileName <- "E:\\data\\Scopus\\All fields regression\\Norway 1996-2018
327 fields\\cov3\\ScopusFind Norway"; Country <- "No"
#RootFileName <- "E:\\data\\Scopus\\All fields regression\\Poland 1996-2018
326 fields\\cov3\\ScopusFind Poland"; Country <- "PL"
#RootFileName <- "E:\\data\\Scopus\\All fields regression\\Russia 1996-2018
318 fields\\cov3\\ScopusFind Russian Federation"; Country <- "Ru"
#RootFileName <- "E:\\data\\Scopus\\All fields regression\\Spain 1996-2018
329 fields\\cov3\\ScopusFind Spain"; Country <- "Es"
#RootFileName <- "E:\\data\\Scopus\\All fields regression\\Sweden 1996-2018
328 fields\\cov3\\ScopusFind Sweden"; Country <- "Se"
#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 1996-2018 330
fields\\cov3\\ScopusFind United Kingdom"; Country <- "UK"
#RootFileName <- "E:\\data\\Scopus\\All fields regression\\USA 1996-2018 331
fields\\cov3\\United States "; Country <- "Usa"
column.names <- c("Articles",
"FirstF","FirstM","FirstP","LastF","LastM","LastP")
GenderTeamSize <- array(integer(7*FieldCount),dim = c(FieldCount, 7),
dimnames <- list(row.names, column.names))
column.names <- c("FFA1","FLA1","2","3","4","5+","FFA2",
"FLA2","FFA3","FLA4")
RegCoef <- array(integer(10*FieldCount),dim = c(FieldCount, 10), dimnames <-
list(row.names, column.names))
column.names <- c("FFA1p","FLA1p","2p","3p","4p","5+p","FFA2p",
"FLA2p","FFA3p","FLA4p")
RegP <- array(integer(10*FieldCount),dim = c(FieldCount, 10), dimnames <-
list(row.names, column.names))
RegStar <- array(character(10*FieldCount),dim = c(FieldCount, 10), dimnames
<- list(row.names, column.names))
column.names <- c("1Ctry","Gend 1st nth","Fem1 96","Fem1 18","Chg1","FemN
96","FemN 18", "ChgN")
BasicStats <- array(integer(8*FieldCount),dim = c(FieldCount, 8), dimnames <-
list(row.names, column.names))
column.names <- c("1CtyMFto14","Fem1All","Fem1 96","Fem1
14","Chg1","FemNAll","FemN96","FemN14", "ChgN", "1CtyMF96", "1CtyMF14")
BasicStats2 <- array(integer(11*FieldCount),dim = c(FieldCount, 11), dimnames
<- list(row.names, column.names))
column.names <- c("CitationSet","Female1st","FemaleLast")
GenderAnalysed <- array(integer(3*FieldCount),dim = c(FieldCount, 3),
dimnames <- list(row.names, column.names))

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

```

```

}
BasicStats[i,8]<-NA; BasicStats2[i,8]<-NA; RegCoef[i,8]<-NA ; RegP[i,8]<-NA
BasicStats2[i,9]<-NA; RegCoef[i,9]<-NA ; RegP[i,9]<-NA
BasicStats2[i,10]<-NA;RegCoef[i,10]<-NA ; RegP[i,10]<-NA
BasicStats2[i,11]<-NA
}

FieldCount = 0
#iCode <- 3302 #test
for (iCode in FieldCodeSet) {
  FieldCount <- FieldCount + 1 #
  SampleFile <- paste(RootFileName, iCode, "_cov.txt", sep="");
  BasicStats[FieldCount,1] <- 0; BasicStats2[FieldCount,1] <- 0;
  if (file.exists(SampleFile)) {
    tryCatch({
      print(""); print("")
      print("#####")
      print(paste("Analysis of AJSC",iCode))
      print("#####")
      AllScopusData <- read.table(file=SampleFile, head=TRUE, sep = "\t")
      names(AllScopusData)[3] <- "NLCS"
      names(AllScopusData)[4] <- "Year"
      names(AllScopusData)[5] <- "OneField"
      names(AllScopusData)[6] <- "Fields"
      #AllScopusData <- AllScopusData[AllScopusData$Year!=2004,] #Uncomment
this to remove year with high outliers if code below reveals a problem
      if (MaxCountriesAllowed>0) AllScopusData <-
AllScopusData[AllScopusData$UniqueCountries <= MaxCountriesAllowed,] #Comment
out to include non-us authors after 1st
      AllScopusDataOlder <- AllScopusData[AllScopusData$Year <=
LastYearForCitationAnalysis,] #2015 Ensure citation window of at least 3
years 2013 for 5 years
      AllScopusDataOlder$Year <- factor(AllScopusDataOlder$Year) #Treat each
year separately
      AllScopusDataOlderFirstGendered <-
AllScopusDataOlder[AllScopusDataOlder$FirstAuthorFemale > -1,] #Female 1,
Male 0, Unknown -1
      AllScopusDataOlderFirstGendered$FirstAuthorFemale <-
factor(AllScopusDataOlderFirstGendered$FirstAuthorFemale)
      AllScopusDataOlderFirstLastGendered <-
AllScopusDataOlderFirstGendered[AllScopusDataOlderFirstGendered$LastAuthorFem
ale > -1,] #Female 1, Male 0, Unknown -1
      AllScopusDataOlderFirstLastGendered$LastAuthorFemale <-
factor(AllScopusDataOlderFirstLastGendered$LastAuthorFemale) #Female 1, Male
0, Unknown -1
      AllYearsBothGendered <-AllScopusData[AllScopusData$FirstAuthorFemale >
-1,]
      AllYearsBothGendered <-
AllYearsBothGendered[AllYearsBothGendered$LastAuthorFemale > -1,]
      AllYearsBothGendered1996 <-

```



```

AllYearsBothGendered[AllYearsBothGendered$Year == 1996,]
  AllYearsBothGenderedLast <-
AllYearsBothGendered[AllYearsBothGendered$Year ==
LastYearForCitationAnalysis,]
  AllYearsBothGendered2018 <-
AllYearsBothGendered[AllYearsBothGendered$Year == 2018,]
  BasicStats[FieldCount,1] <- nrow(AllScopusData)
  BasicStats[FieldCount,2] <- nrow(AllYearsBothGendered) /
nrow(AllScopusData) * 100 #Percentage
  BasicStats[FieldCount,3] <- count_if(1,
AllYearsBothGendered1996$FirstAuthorFemale)/ nrow(AllYearsBothGendered1996) *
100
  BasicStats[FieldCount,4] <- count_if(1,
AllYearsBothGendered2018$FirstAuthorFemale)/ nrow(AllYearsBothGendered2018) *
100
  BasicStats[FieldCount,5] <- BasicStats[FieldCount,4] -
BasicStats[FieldCount,3]
  BasicStats[FieldCount,6] <- count_if(1,
AllYearsBothGendered1996$LastAuthorFemale)/ nrow(AllYearsBothGendered1996) *
100
  BasicStats[FieldCount,7] <- count_if(1,
AllYearsBothGendered2018$LastAuthorFemale)/ nrow(AllYearsBothGendered2018) *
100
  BasicStats[FieldCount,8] <- BasicStats[FieldCount,7] -
BasicStats[FieldCount,6]
  ##c("US gendered96to14","Fem1 All","Fem1 96","Fem1 14","Chg","FemN
All","FemN 96","FemN 14", "Chg")
  BasicStats2[FieldCount,1] <- nrow(AllScopusDataOlderFirstLastGendered)
  BasicStats2[FieldCount,2] <- count_if(1,
AllScopusDataOlderFirstLastGendered$FirstAuthorFemale) /
nrow(AllScopusDataOlderFirstLastGendered) * 100 #Percentage
  BasicStats2[FieldCount,3] <- count_if(1,
AllYearsBothGendered1996$FirstAuthorFemale)/ nrow(AllYearsBothGendered1996) *
100
  BasicStats2[FieldCount,4] <- count_if(1,
AllYearsBothGenderedLast$FirstAuthorFemale)/ nrow(AllYearsBothGenderedLast) *
100
  BasicStats2[FieldCount,5] <- BasicStats2[FieldCount,4] -
BasicStats2[FieldCount,3]
  BasicStats2[FieldCount,6] <- count_if(1,
AllScopusDataOlderFirstLastGendered$LastAuthorFemale) /
nrow(AllScopusDataOlderFirstLastGendered) * 100
  BasicStats2[FieldCount,7] <- count_if(1,
AllYearsBothGendered1996$LastAuthorFemale)/ nrow(AllYearsBothGendered1996) *
100
  BasicStats2[FieldCount,8] <- count_if(1,
AllYearsBothGenderedLast$LastAuthorFemale)/ nrow(AllYearsBothGenderedLast) *
100
  BasicStats2[FieldCount,9] <- BasicStats2[FieldCount,8] -
BasicStats2[FieldCount,7]

```

```

BasicStats2[FieldCount,10] <- nrow(AllYearsBothGendered1996)
BasicStats2[FieldCount,11] <- nrow(AllYearsBothGenderedLast)
#print("MNLCS for all years [All, first gendered, first & last
gendered], just to check nothing is odd")
#print(tapply(AllScopusDataOlder$NLCS, AllScopusDataOlder$Year, mean))
#print(tapply(AllScopusDataOlderFirstGendered$NLCS,
AllScopusDataOlderFirstGendered$Year, mean))
#print(tapply(AllScopusDataOlderFirstLastGendered$NLCS,
AllScopusDataOlderFirstLastGendered$Year, mean))

print("Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]")
print(table(AllScopusDataOlder$Year))
print(table(AllScopusDataOlderFirstGendered$Year))
print(table(AllScopusDataOlderFirstLastGendered$Year))
}, error = function(e) return("failed narrow field data entry and basic
processing"))
tryCatch({
#####
print("Heteroscedasticity checks, confirming that there are problems
with these")
#####
#Check for outliers caused by field normalisation - should be evident
in heteroskedacity caused by individual years
print(bartlett.test(NLCS~Year, data=AllScopusDataOlderFirstGendered))
#Homogeneity of Variances test. Big fail is OK because older years have
greater variability
YearLm <- lm(NLCS~Year, data=AllScopusDataOlderFirstGendered)
plot(YearLm, which = 1) #Residuals vs. fitted. Check that width of
residual bar is not huge for a few years
AllScopusDataOlderFirstGendered$YMresiduals <- resid(YearLm)
YearGenderLm <- lm(YMresiduals~FirstAuthorFemale,
data=AllScopusDataOlderFirstGendered)
print(bartlett.test(YMresiduals~FirstAuthorFemale,
data=AllScopusDataOlderFirstGendered)) #Homogeneity of Variances test - this
is the key test - should not fail by much, but might 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$LastAut
horFemale)
  FemaleTeamSizes2018 <- as.numeric(GenderAuthors2018$'1')
  MaleTeamSizes2018 <- as.numeric(GenderAuthors2018$'0')
  GenderAnalysed[FieldCount,3] <- length(FemaleTeamSizes2018)
  print(paste("Female last author team size 2018 geometric mean:",
geometric.mean(FemaleTeamSizes2018)))
  print(paste("Male last author team size 2018 geometric mean:",
geometric.mean(MaleTeamSizes2018)))
  wilc<- wilcox.test(FemaleTeamSizes2018, MaleTeamSizes2018, alternative
= "two.sided")
  print(wilc)
  GenderTeamSize[FieldCount,5] <- geometric.mean(FemaleTeamSizes2018)
  GenderTeamSize[FieldCount,6] <- geometric.mean(MaleTeamSizes2018)
  GenderTeamSize[FieldCount,7] <- wilc$p.value
}, error = function(e) return("Failed team size tests"))
tryCatch({
  #Use 5 for 5+ authors
  for (i in 1:nrow(AllScopusDataOlderFirstLastGendered)) {
    AllScopusDataOlderFirstLastGendered$UniqueAuthors[i] <-
min(MaxAuthorsToCountInRegression,AllScopusDataOlderFirstLastGendered$UniqueA
uthors[i])
  }

AllScopusDataOlderFirstLastGendered$UniqueAuthors=factor(AllScopusDataOlderFi
rstLastGendered$UniqueAuthors) #Don't assume any team size influence formula
#####
print("Regression 1: First author gender, last author gender, team
size, Year as factors")
#####
if (iCode %in% FieldNeedsExtraTime) {
  FirstLastAuthorTeamLmrob <-

```

```

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

```

```

lmrob(NLCS~FirstAuthorFemale+LastAuthorFemale+Year,
data=AllScopusDataOlderFirstLastGendered)
}
print(vif(FirstLastAuthorLmrob)) # check under 5
hist(resid(FirstLastAuthorLmrob),main='Residuals from first and last
author',xlab='Standardised Residuals',ylab='Frequency') #Normality not needed
but useful - check for outliers - if too many outside +/- 1.96 then possible
field normalisation problem
AllScopusDataOlderFirstLastGendered$residuals <-
resid(FirstLastAuthorLmrob)
Outliers <- AllScopusDataOlderFirstLastGendered[
abs(AllScopusDataOlderFirstLastGendered$residuals) > outlier_threshold, ]
Outliers <- Outliers[c(1,3,4,5,6,16)]
Outliers$ScopusId <- gsub('scopus_id:', '', Outliers$ScopusId)
print(paste("List of ", nrow(Outliers), "outliers with residuals above
", outlier_threshold))
print(Outliers)
sum <- summary(FirstLastAuthorLmrob)
print(sum) #Robust to heteroscedacity
RegCoef[FieldCount,7] <- sum$coefficients[2]; RegP[FieldCount,7] <-
sum$coefficients[2,4];
RegCoef[FieldCount,8] <- sum$coefficients[3]; RegP[FieldCount,8] <-
sum$coefficients[3,4];
}, error = function(e) return("Failed regression 4"))
tryCatch({
#####
print("Regression 3: First author gender, Year as factors")
#####
#First author regression
if (iCode %in% FieldNeedsExtraTime) {
FirstAuthorLmrob <- lmrob(NLCS~FirstAuthorFemale+Year,
data=AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf, k.max = 1000))
} else {
FirstAuthorLmrob <- lmrob(NLCS~FirstAuthorFemale+Year,
data=AllScopusDataOlderFirstLastGendered)
}
print(vif(FirstAuthorLmrob)) # check under 5
hist(resid(FirstAuthorLmrob),main='Residuals from first
author',xlab='Standardised Residuals',ylab='Frequency') #Normality not needed
but useful - check for outliers - if too many outside +/- 1.96 then possible
field normalisation problem
AllScopusDataOlderFirstLastGendered$residuals <-
resid(FirstAuthorLmrob)
Outliers <- AllScopusDataOlderFirstLastGendered[
abs(AllScopusDataOlderFirstLastGendered$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
##    2    4    4    5    3    5    3    3    5    7    1    4    9    9    6
## 2011 2012
##    3    4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    2    1    3    2    3    2    2    4    5    0    2    8    3    5
## 2011 2012
##    2    3
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    2    1    3    2    3    2    2    4    4    0    2    8    2    5
## 2011 2012
##    2    3
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.75186951540277"
## [1] "Male first author team size 2018 geometric mean: 3.37411222668032"

## Warning in wilcox.test.default(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 = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.6205885667505"
## [1] "Male last author team size 2018 geometric mean: 3.49625946174963"

## Warning in wilcox.test.default(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 = 0.8
## 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.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

## 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"

## Warning in outlierStats(ret, x, control): Detected possible local
breakdown of SM-estimate in coefficient 'Year1997'.
## Use lmrob argument 'setting="KS2014"' to avoid this problem.

##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 9.167  1          3.028
## Year              9.167 15          1.077

## [1] "Regression 4: Last author gender, Year as factors"

## Warning in outlierStats(ret, x, control): Detected possible local
breakdown of SM-estimate in coefficient 'Year1997'.
## Use lmrob argument 'setting="KS2014"' to avoid this problem.

## 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:  47"
## [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
##   14   19   12    2   14    8   10   13   13   19   17   23   21   28   26
## 2011 2012
##   46   68
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010

```

```

##      7   10    7    1    2    5    8    4    8   14   13   13   18   23   18
## 2011 2012
##   34   50
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4   10    7    0    2    4    4    3    7   13   13   11   17   21   17
## 2011 2012
##   29   45
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 5.13892231147464"
## [1] "Male first author team size 2018 geometric mean: 3.23167158913068"

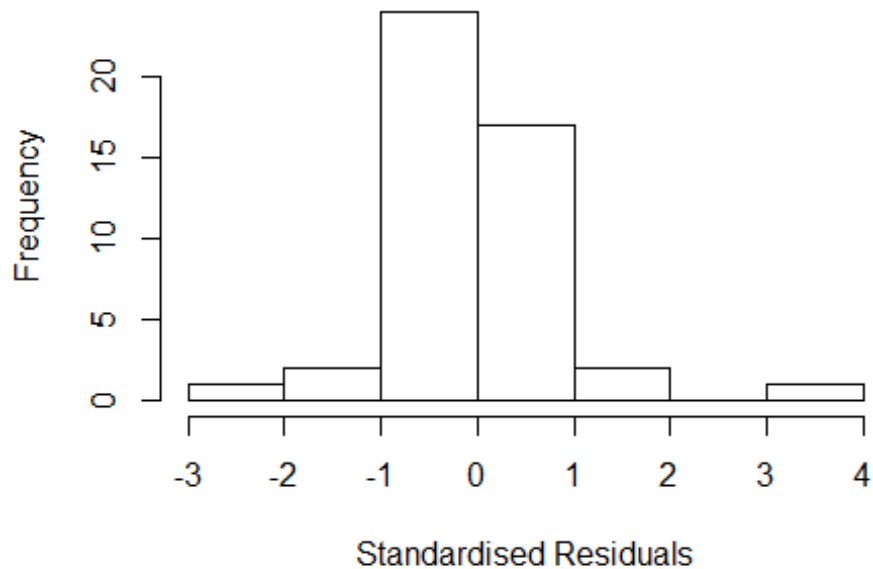
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## 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] "Female last author team size 2018 geometric mean: 5.98511147895538"
## [1] "Male last author team size 2018 geometric mean: 3.77390506953534"

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

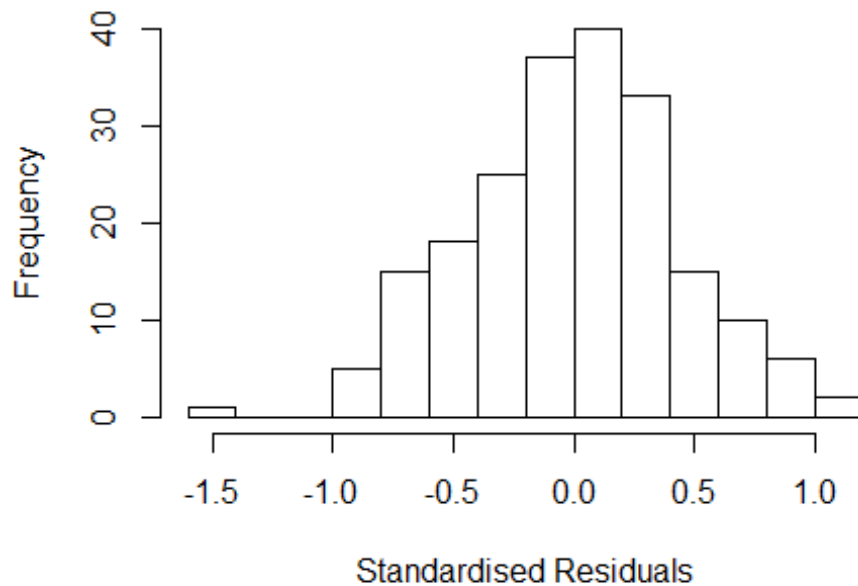
```

Residuals from first author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 72, 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 1.461 1 1.209
## LastAuthorFemale 1.296 1 1.138
## UniqueAuthors 4.296 4 1.200
## Year 6.248 15 1.063
```

Residuals from first and last author and team size



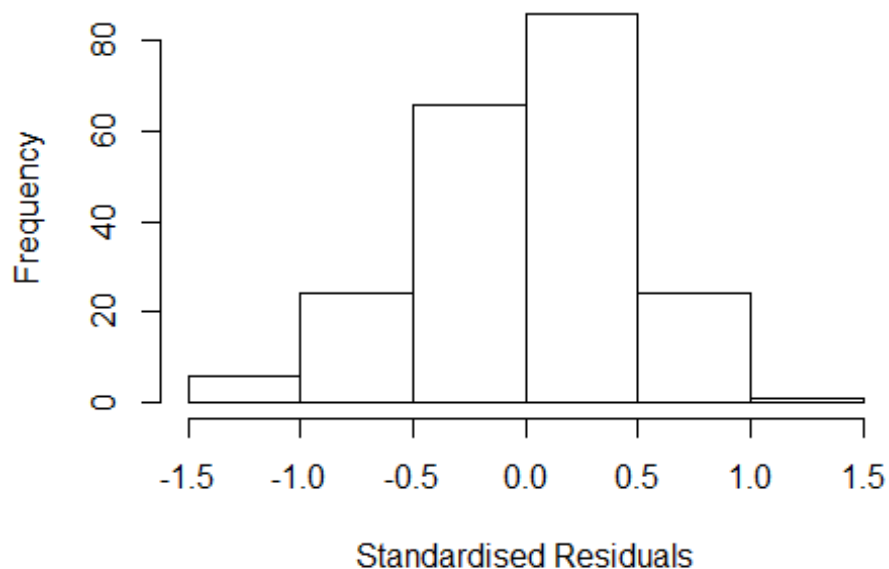
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4416 -0.2993 0.0168 0.2741 1.0228
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.87547 0.13087 6.69 2.6e-10 ***
## FirstAuthorFemale1 0.00742 0.06853 0.11 0.91393
## LastAuthorFemale1 -0.05793 0.08482 -0.68 0.49548
## UniqueAuthors2 0.39702 0.13889 2.86 0.00474 **
## UniqueAuthors3 0.68071 0.13372 5.09 8.7e-07 ***
## UniqueAuthors4 0.54603 0.14145 3.86 0.00016 ***
## UniqueAuthors5 0.72402 0.14240 5.08 9.0e-07 ***
## Year1997 -0.10561 0.21180 -0.50 0.61863
## Year1998 -0.18114 0.19488 -0.93 0.35384
## Year2000 0.01478 0.60926 0.02 0.98067
```

```

## Year2001          0.00497    0.26440    0.02  0.98502
## Year2002         -0.32793    0.35651   -0.92  0.35886
## Year2003         -0.57850    0.19638   -2.95  0.00364 **
## Year2004         -0.31129    0.18239   -1.71  0.08955 .
## Year2005         -0.57183    0.19890   -2.87  0.00451 **
## Year2006         -0.14482    0.20485   -0.71  0.48048
## Year2007         -0.16325    0.24208   -0.67  0.50092
## Year2008         -0.31987    0.17118   -1.87  0.06325 .
## Year2009         -0.32195    0.17601   -1.83  0.06899 .
## Year2010         -0.32520    0.17742   -1.83  0.06841 .
## Year2011         -0.12202    0.18620   -0.66  0.51309
## Year2012         -0.28200    0.17454   -1.62  0.10786
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.447
## Multiple R-squared:  0.251, Adjusted R-squared:  0.165
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 191 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.278  0.865  0.958  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      4.83e-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 1.389 1 1.179
## LastAuthorFemale 1.329 1 1.153
## Year 1.734 15 1.019

```


Residuals from first and last author



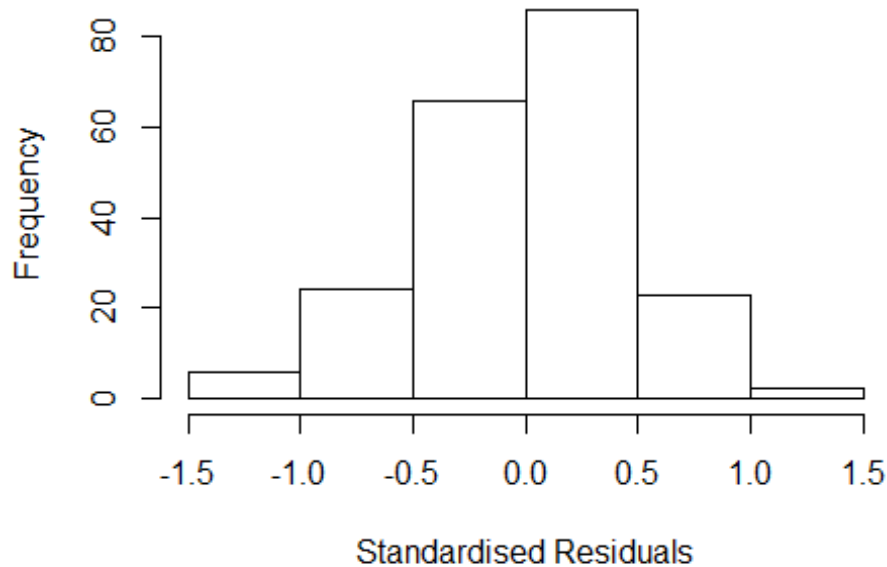
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3277 -0.3421 0.0351 0.3265 1.2164
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0299 0.2427 4.24 3.4e-05 ***
## FirstAuthorFemale1 0.0661 0.0742 0.89 0.37
## LastAuthorFemale1 -0.0231 0.0976 -0.24 0.81
## Year1997 -0.0769 0.2861 -0.27 0.79
## Year1998 0.1975 0.2764 0.71 0.48
## Year2000 -0.1864 0.6965 -0.27 0.79
## Year2001 0.3570 0.3861 0.92 0.36
## Year2002 -0.0316 0.3813 -0.08 0.93
## Year2003 -0.4213 0.3752 -1.12 0.26
## Year2004 0.0818 0.2688 0.30 0.76
## Year2005 -0.2579 0.2837 -0.91 0.36
## Year2006 0.1642 0.2668 0.62 0.54
```

```

## Year2007          0.1319      0.3268      0.40      0.69
## Year2008          0.0175      0.2632      0.07      0.95
## Year2009         -0.0610      0.2700     -0.23      0.82
## Year2010          0.0766      0.2697      0.28      0.78
## Year2011          0.2317      0.2708      0.86      0.39
## Year2012          0.1318      0.2598      0.51      0.61
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.477
## Multiple R-squared:  0.0884, Adjusted R-squared:  0.00635
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 189 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.420  0.876   0.953   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      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 1.374 1          1.172
## Year              1.374 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
## -1.3205 -0.3374 0.0433 0.3273 1.2217
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0304 0.2419 4.26 3.2e-05 ***
## FirstAuthorFemale1 0.0642 0.0757 0.85 0.40
## Year1997 -0.0828 0.2850 -0.29 0.77
## Year1998 0.1949 0.2755 0.71 0.48
## Year2000 -0.1975 0.7063 -0.28 0.78
## Year2001 0.3532 0.3897 0.91 0.37
## Year2002 -0.0338 0.3810 -0.09 0.93
## Year2003 -0.4218 0.3742 -1.13 0.26
## Year2004 0.0788 0.2671 0.30 0.77
## Year2005 -0.2592 0.2832 -0.92 0.36
## Year2006 0.1593 0.2656 0.60 0.55
## Year2007 0.1278 0.3232 0.40 0.69
```

```

## Year2008          0.0115      0.2614      0.04      0.96
## Year2009         -0.0627      0.2685     -0.23      0.82
## Year2010          0.0718      0.2668      0.27      0.79
## Year2011          0.2259      0.2699      0.84      0.40
## Year2012          0.1241      0.2554      0.49      0.63
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.484
## Multiple R-squared:  0.0876, Adjusted R-squared:  0.0107
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 189 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.438  0.881  0.955  0.908  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      4.83e-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 1.287  1          1.134
## Year            1.287 15          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
## -1.2965 -0.3279  0.0451  0.3213  1.1815
##

```

```

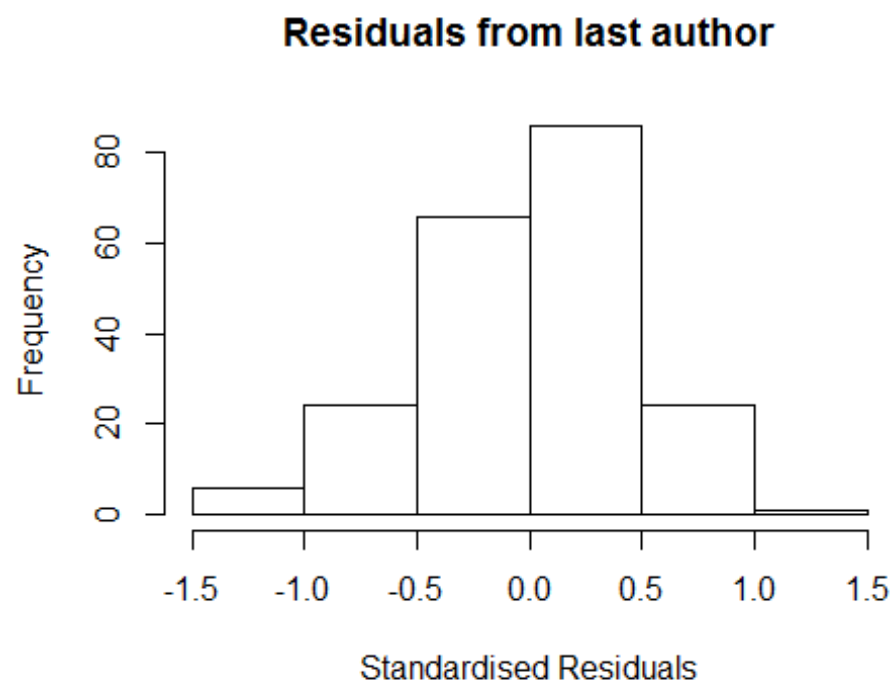
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.03024    0.24211   4.26 3.3e-05 ***
## LastAuthorFemale1 -0.01733    0.09830  -0.18  0.86
## Year1997        -0.05949    0.28643  -0.21  0.84
## Year1998         0.21804    0.27390   0.80  0.43
## Year2000        -0.15658    0.63307  -0.25  0.80
## Year2001         0.36110    0.37401   0.97  0.34
## Year2002         0.00999    0.37854   0.03  0.98
## Year2003        -0.40027    0.38892  -1.03  0.30
## Year2004         0.12049    0.26582   0.45  0.65
## Year2005        -0.22226    0.28132  -0.79  0.43
## Year2006         0.18492    0.26622   0.69  0.49
## Year2007         0.18077    0.31134   0.58  0.56
## Year2008         0.04425    0.25921   0.17  0.86
## Year2009        -0.04034    0.26776  -0.15  0.88
## Year2010         0.10609    0.26657   0.40  0.69
## Year2011         0.26630    0.26832   0.99  0.32
## Year2012         0.15833    0.25690   0.62  0.54
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.482
## Multiple R-squared:  0.0844, Adjusted R-squared:  0.00733
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 193 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.450  0.889  0.951  0.909  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      4.83e-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 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
##    7    6   13    7   16    9    4    8    5   10   10    8   12    4    2
## 2012
##    1
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    4    4    2    7    4    2    3    4    8    9    4   11    3    2
## 2012
##    1
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    4    4    2    6    4    2    2    3    7    7    4   11    2    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: 2.28942848510666"
## [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 refinements did
not
## converge (to refine.tol=1e-07) in 1000 (= k.max) steps

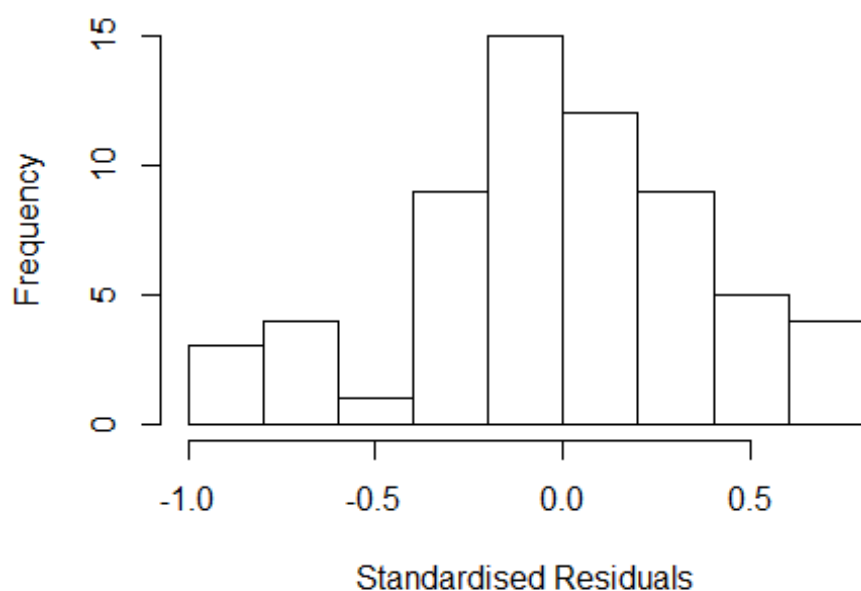
```

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

```
##               GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.411e+01  1      3.756e+00
## LastAuthorFemale  2.963e+13  1      5.443e+06
## Year              3.998e+14 15      3.067e+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
## -9.55e-01 -2.11e-01 -3.89e-16 2.88e-01 7.00e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6659 0.9227 0.72 0.474
## FirstAuthorFemale1 0.2659 0.1580 1.68 0.099 .
## LastAuthorFemale1 0.1804 0.1875 0.96 0.341
## Year1997 0.0391 0.9349 0.04 0.967
## Year1998 0.6684 0.9240 0.72 0.473
## Year1999 0.1136 0.9823 0.12 0.908
## Year2000 0.2741 0.9294 0.29 0.769
## Year2001 0.3216 0.9618 0.33 0.740
## Year2002 0.5657 0.9165 0.62 0.540
## Year2003 0.3136 0.9266 0.34 0.737
## Year2004 0.0652 0.9128 0.07 0.943
## Year2005 0.3171 0.9247 0.34 0.733
```

```

## Year2006          0.2889      0.9391      0.31      0.760
## Year2007          0.6493      0.9431      0.69      0.495
## Year2008          0.3592      0.9261      0.39      0.700
## Year2009          0.9126      0.9447      0.97      0.339
## Year2010          0.9989      0.9172      1.09      0.282
## Year2012          0.1562      0.9149      0.17      0.865
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.433
## Multiple R-squared:  0.285, Adjusted R-squared:  0.0089
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 52 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.606  0.841  0.957   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      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 -53.71  1      NaN
## Year              -53.71 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
## -1.03e+00 -2.06e-01  1.49e-16  2.64e-01  7.48e-01

```

```
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.7482    1.5493   0.48    0.63
## FirstAuthorFemale1 0.2817    0.1768   1.59    0.12
## Year1997         -0.0507    1.5548  -0.03    0.97
## Year1998          0.6250    1.5568   0.40    0.69
## Year1999          0.0313    1.5865   0.02    0.98
## Year2000          0.2659    1.5445   0.17    0.86
## Year2001          0.2433    1.5575   0.16    0.88
## Year2002          0.4755    1.5431   0.31    0.76
## Year2003          0.2313    1.5517   0.15    0.88
## Year2004          0.0279    1.5413   0.02    0.99
## Year2005          0.2542    1.5539   0.16    0.87
## Year2006          0.2866    1.5504   0.18    0.85
## Year2007          0.5602    1.5592   0.36    0.72
## Year2008          0.2955    1.5512   0.19    0.85
## Year2009          0.8303    1.5626   0.53    0.60
## Year2010          1.0812    1.5402   0.70    0.49
## Year2012          0.0582    1.5402   0.04    0.97
##
## Robust residual standard error: 0.425
## Multiple R-squared:  0.267, Adjusted R-squared:  0.0065
## Convergence in 41 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 54 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.533  0.865  0.955  0.911  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      1.61e-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"

## 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: 62"
## [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
##    25    11    26    36    31    30    34    22    27    36    39    29    43    40    44
## 2011 2012
##    39    38
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     4     1     4     7     6     2     9     6     7     5    15    10    13    12    21
## 2011 2012
##    13    13
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     3     1     4     6     5     2     7     5     6     5    15    10    13     9    17
## 2011 2012
##    12    10
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.52701905583787"
## [1] "Male first author team size 2018 geometric mean: 3.17659518715315"

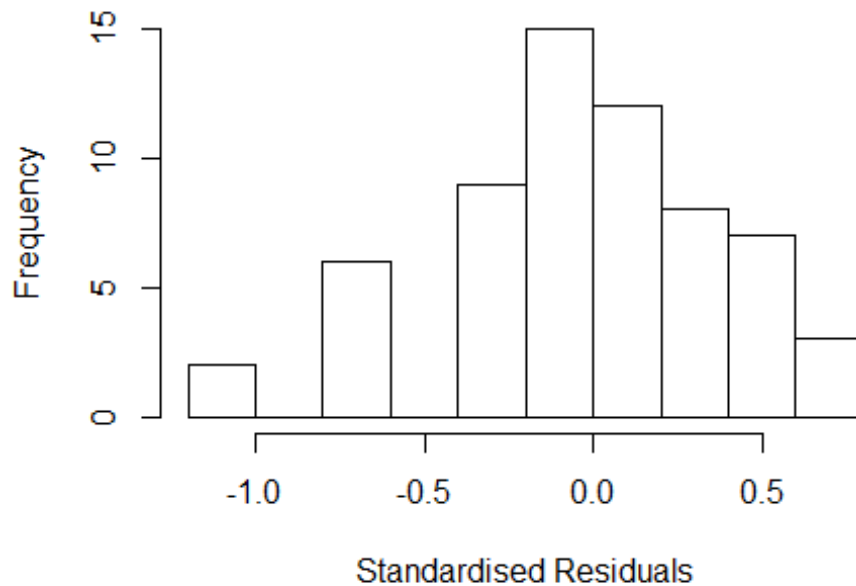
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 23, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.1924941018511"
## [1] "Male last author team size 2018 geometric mean: 3.15647100867328"

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

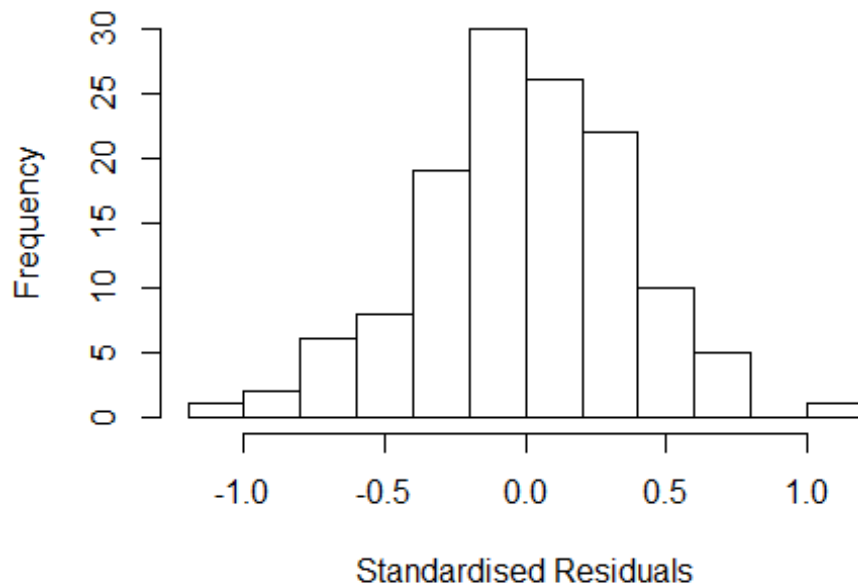
```

Residuals from first author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 24, p-value = 0.07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale  3.298  1      1.816
## LastAuthorFemale  2.899  1      1.703
## UniqueAuthors    99.736  4      1.778
## Year             239.481 16      1.187
```


Residuals from first and last author and team size



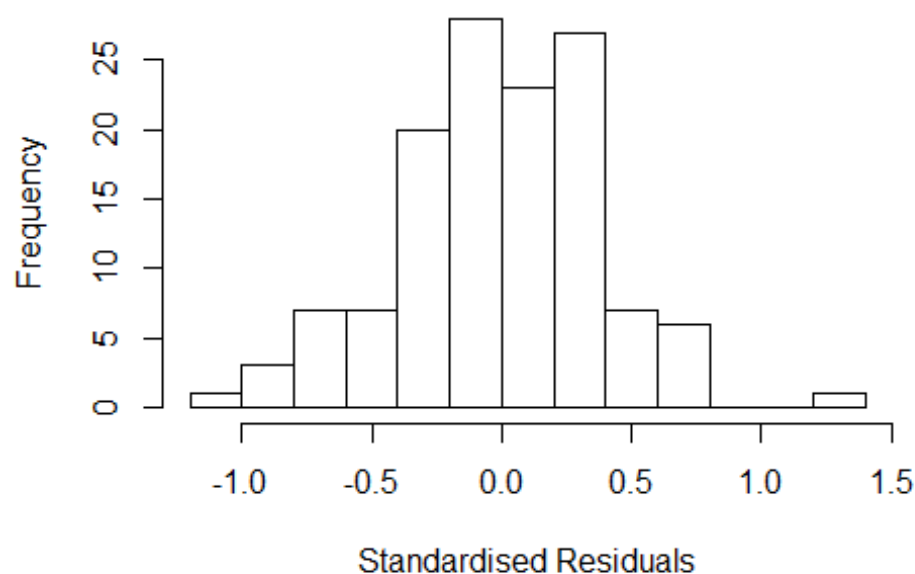
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q    Median        3Q        Max
## -1.153230 -0.231527 -0.000186  0.267514  1.061010
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.61422    0.08214   7.48 2.2e-11 ***
## FirstAuthorFemale1 0.00957    0.08049   0.12 0.90560
## LastAuthorFemale1 0.05017    0.09309   0.54 0.59108
## UniqueAuthors2    0.31649    0.14929   2.12 0.03632 *
## UniqueAuthors3    0.44397    0.13534   3.28 0.00140 **
## UniqueAuthors4    0.37166    0.15036   2.47 0.01502 *
## UniqueAuthors5    0.53136    0.13840   3.84 0.00021 ***
## Year1997          1.15178    0.08214  14.02 < 2e-16 ***
## Year1998           0.11852    0.29438   0.40 0.68803
## Year1999          -0.03589    0.09616  -0.37 0.70972
```

```

## Year2000      0.10735      0.16104      0.67  0.50646
## Year2001     -0.08899      0.20855     -0.43  0.67046
## Year2002      0.15186      0.17847      0.85  0.39672
## Year2003     -0.11866      0.19932     -0.60  0.55290
## Year2004     -0.06363      0.16265     -0.39  0.69641
## Year2005     -0.00554      0.16293     -0.03  0.97292
## Year2006      0.21514      0.14520      1.48  0.14134
## Year2007      0.14899      0.15877      0.94  0.35014
## Year2008      0.13009      0.18699      0.70  0.48810
## Year2009      0.40104      0.23290      1.72  0.08796 .
## Year2010      0.12126      0.13185      0.92  0.35980
## Year2011      0.02090      0.16390      0.13  0.89878
## Year2012      0.30441      0.19047      1.60  0.11294
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.369
## Multiple R-squared:  0.262, Adjusted R-squared:  0.11
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 114 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.308  0.889  0.951  0.907  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      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 2.549 1      1.596
## LastAuthorFemale 2.488 1      1.577
## Year      4.669 16      1.049

```

Residuals from first and last author



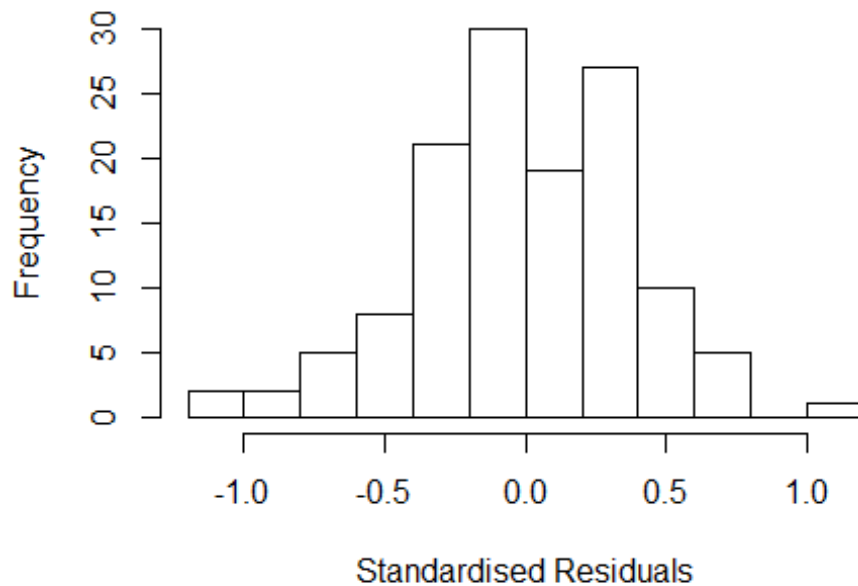
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.034388 -0.245235 -0.000192 0.271193 1.201276
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7434 0.1718 4.33 3.3e-05 ***
## FirstAuthorFemale1 0.0348 0.0822 0.42 0.673
## LastAuthorFemale1 0.0882 0.0972 0.91 0.366
## Year1997 1.0226 0.1718 5.95 3.1e-08 ***
## Year1998 0.0677 0.4507 0.15 0.881
## Year1999 0.1255 0.2119 0.59 0.555
## Year2000 0.2339 0.1949 1.20 0.233
## Year2001 0.1390 0.2444 0.57 0.571
## Year2002 0.3622 0.2097 1.73 0.087 .
## Year2003 0.1385 0.2764 0.50 0.617
## Year2004 0.2144 0.2379 0.90 0.369
## Year2005 0.2872 0.2203 1.30 0.195
```

```

## Year2006          0.4629      0.1948      2.38      0.019 *
## Year2007          0.4381      0.2000      2.19      0.031 *
## Year2008          0.3115      0.2309      1.35      0.180
## Year2009          0.5970      0.2366      2.52      0.013 *
## Year2010          0.3580      0.2041      1.75      0.082 .
## Year2011          0.2978      0.1980      1.50      0.135
## Year2012          0.5663      0.2344      2.42      0.017 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.406
## Multiple R-squared:  0.159, Adjusted R-squared:  0.0229
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 113 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.362  0.901  0.948  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      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 2.433 1          1.560
## Year              2.433 16          1.028

```

Residuals from first author



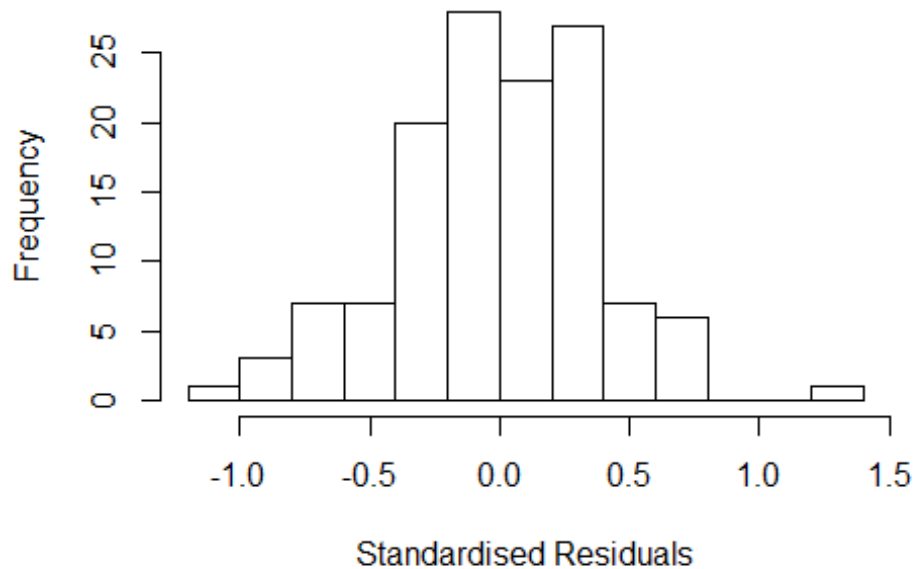
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.07777 -0.24741 -0.00984 0.27806 1.17130
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7683 0.2015 3.81 0.00022 ***
## FirstAuthorFemale1 0.0501 0.0786 0.64 0.52506
## Year1997 0.9977 0.2015 4.95 2.6e-06 ***
## Year1998 0.0429 0.4649 0.09 0.92659
## Year1999 0.1111 0.2390 0.47 0.64278
## Year2000 0.2216 0.2217 1.00 0.31989
## Year2001 0.1946 0.2586 0.75 0.45334
## Year2002 0.3382 0.2342 1.44 0.15156
## Year2003 0.1222 0.2876 0.42 0.67166
## Year2004 0.2224 0.2602 0.85 0.39443
## Year2005 0.2529 0.2428 1.04 0.29988
## Year2006 0.4583 0.2230 2.06 0.04218 *
```

```

## Year2007          0.4250      0.2257      1.88  0.06224 .
## Year2008          0.3016      0.2564      1.18  0.24197
## Year2009          0.6154      0.2581      2.38  0.01879 *
## Year2010          0.3448      0.2290      1.51  0.13485
## Year2011          0.2759      0.2227      1.24  0.21790
## Year2012          0.5714      0.2500      2.29  0.02419 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.404
## Multiple R-squared:  0.155, Adjusted R-squared:  0.0271
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 118 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.381  0.894  0.956  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      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 2.382 1      1.543
## Year            2.382 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
## -1.03550 -0.24885 0.00655 0.28163 1.18807
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7413 0.1694 4.38 2.7e-05 ***
## LastAuthorFemale1 0.0956 0.0937 1.02 0.310
## Year1997 1.0247 0.1694 6.05 2.0e-08 ***
## Year1998 0.0695 0.4473 0.16 0.877
## Year1999 0.1383 0.2102 0.66 0.512
## Year2000 0.2484 0.1923 1.29 0.199
## Year2001 0.1511 0.2328 0.65 0.518
## Year2002 0.3696 0.2051 1.80 0.074 .
## Year2003 0.1541 0.2655 0.58 0.563
## Year2004 0.2151 0.2338 0.92 0.360
## Year2005 0.3105 0.2104 1.48 0.143
## Year2006 0.4740 0.1945 2.44 0.016 *
```

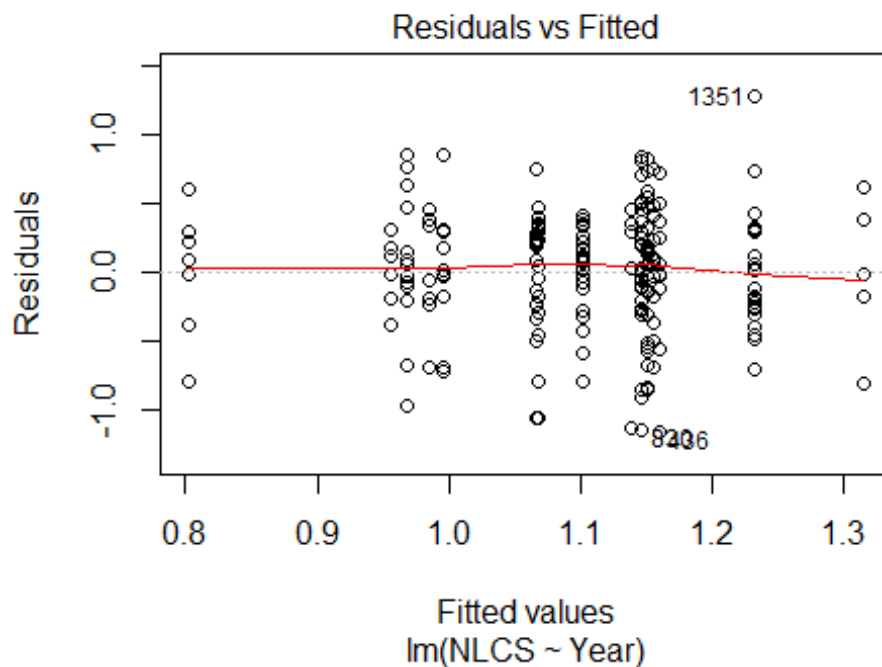
```

## Year2007          0.4538      0.1921      2.36      0.020 *
## Year2008          0.3192      0.2286      1.40      0.165
## Year2009          0.6002      0.2343      2.56      0.012 *
## Year2010          0.3693      0.2012      1.84      0.069 .
## Year2011          0.3075      0.1958      1.57      0.119
## Year2012          0.5816      0.2255      2.58      0.011 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.409
## Multiple R-squared:  0.158, Adjusted R-squared:  0.0302
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 114 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.379  0.906  0.955  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      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 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
##   29   39   45   49   59   49   57   47   42   64   61   60   78   88   84
## 2011 2012
##   79   90
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7    7    5    6    6    6   10   10   11   13   14    9   16   27   13
## 2011 2012

```



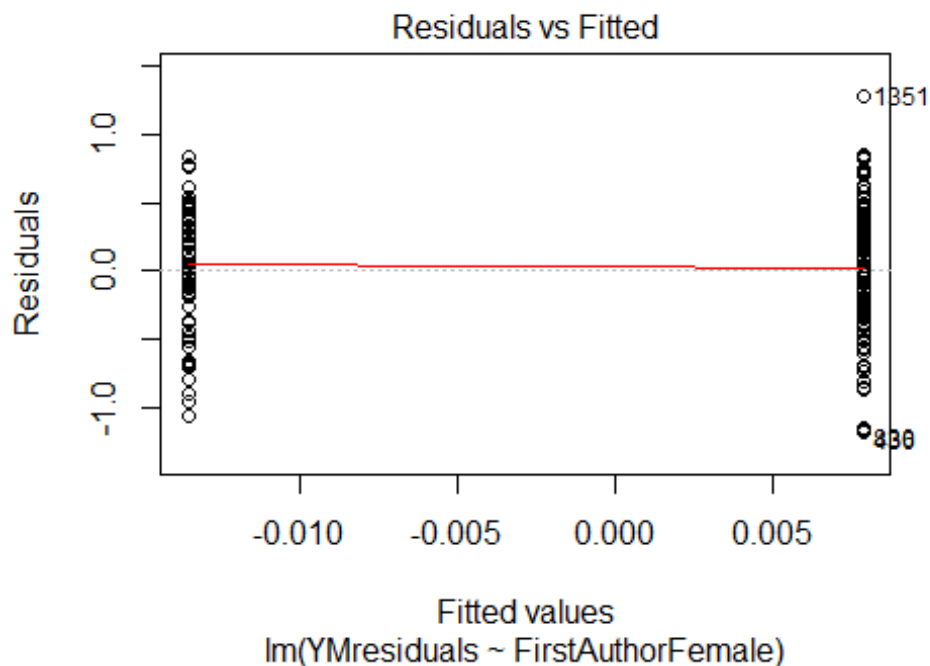
```
## 26 21
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 7 6 5 6 5 5 5 7 8 13 12 9 14 25 11
## 2011 2012
## 19 17
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 16, df = 16, p-value = 0.5
```



```
##
## 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: 4.3785313089672"
## [1] "Male first author team size 2018 geometric mean: 3.49342533860972"
## Warning in wilcox.test.default(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: 4.14349619657705"
## [1] "Male last author team size 2018 geometric mean: 3.70671710665565"

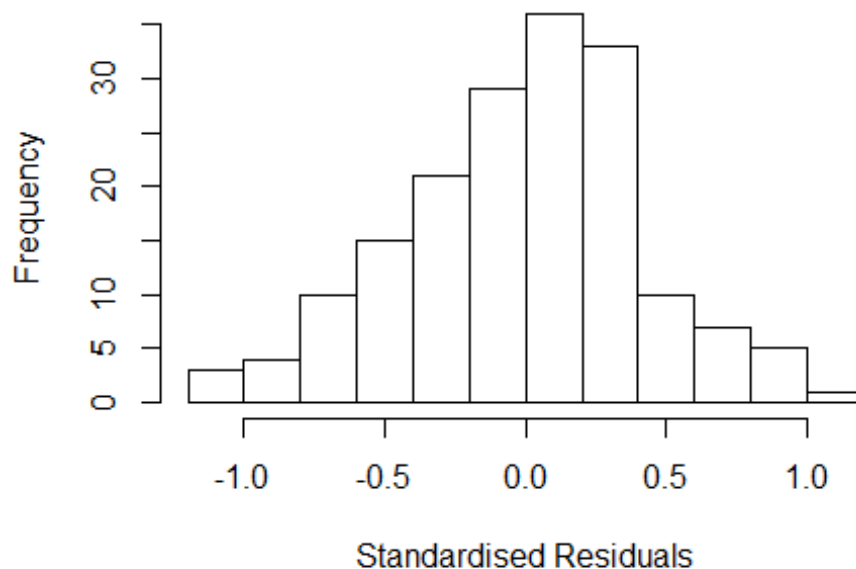
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 65, p-value = 0.7
## 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 | 1.830 | 1 | 1.353 |
| LastAuthorFemale | 1.887 | 1 | 1.374 |
| UniqueAuthors | 8.568 | 4 | 1.308 |
| Year | 17.155 | 16 | 1.093 |

Residuals from first and last author and team size



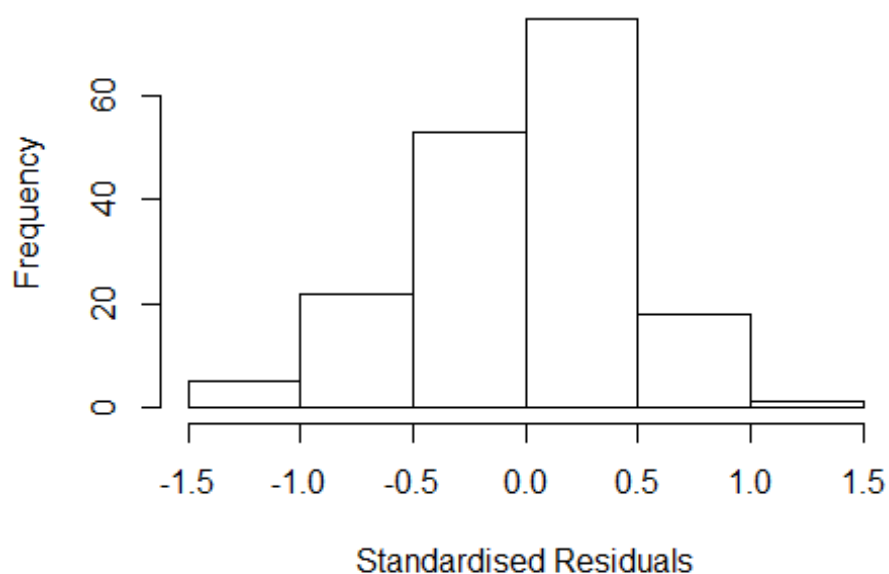
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0242 -0.2884 0.0287 0.2646 1.1894
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6672 0.3066 2.18 0.03107 *
## FirstAuthorFemale1 -0.0594 0.0798 -0.74 0.45808
## LastAuthorFemale1 -0.1515 0.0847 -1.79 0.07565 .
## UniqueAuthors2 0.4213 0.1994 2.11 0.03629 *
## UniqueAuthors3 0.5130 0.1941 2.64 0.00909 **
## UniqueAuthors4 0.4521 0.1918 2.36 0.01972 *
## UniqueAuthors5 0.6231 0.1831 3.40 0.00085 ***
## Year1997 -0.0607 0.2917 -0.21 0.83540
## Year1998 0.0510 0.2835 0.18 0.85747
## Year1999 0.0581 0.2477 0.23 0.81479
```

```

## Year2000          -0.1321      0.2532   -0.52  0.60263
## Year2001           0.1668      0.2799    0.60  0.55214
## Year2002           0.1134      0.2562    0.44  0.65855
## Year2003           0.1806      0.3278    0.55  0.58255
## Year2004           0.0592      0.2481    0.24  0.81156
## Year2005          -0.0874      0.2523   -0.35  0.72956
## Year2006          -0.0357      0.2986   -0.12  0.90503
## Year2007          -0.0133      0.2726   -0.05  0.96117
## Year2008           0.0923      0.2793    0.33  0.74160
## Year2009           0.1393      0.2619    0.53  0.59558
## Year2010           0.0530      0.2503    0.21  0.83249
## Year2011          -0.0738      0.2306   -0.32  0.74944
## Year2012           0.2401      0.2488    0.96  0.33616
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.426
## Multiple R-squared:  0.201, Adjusted R-squared:  0.0841
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 160 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.416  0.877  0.958  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      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 1.553 1      1.246
## LastAuthorFemale  2.279 1      1.510
## Year              3.436 16      1.039

```

Residuals from first and last author



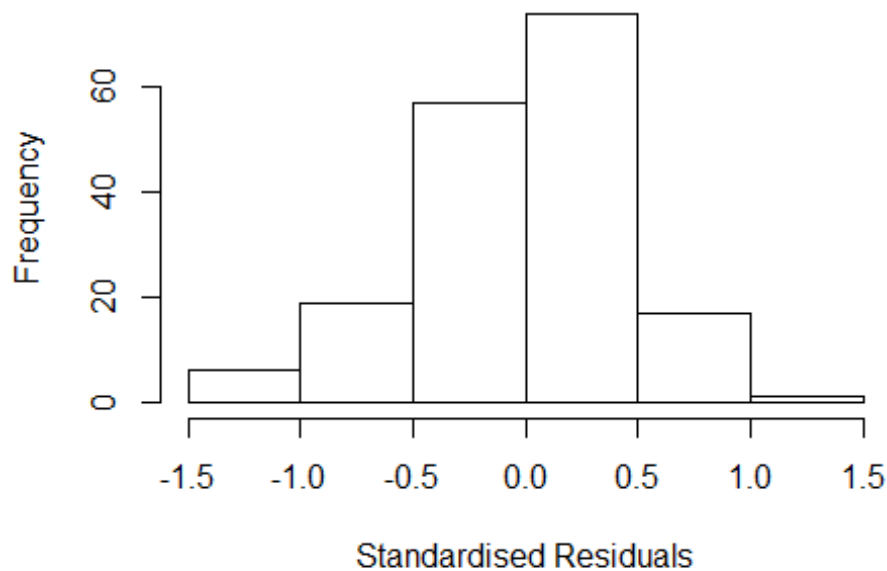
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3162 -0.2831  0.0467  0.2600  1.2345
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.07206    0.17338   6.18 5.3e-09 ***
## FirstAuthorFemale1 -0.00851    0.08008  -0.11   0.92
## LastAuthorFemale1 -0.14268    0.10868  -1.31   0.19
## Year1997        -0.17393    0.25530  -0.68   0.50
## Year1998         0.24417    0.26577   0.92   0.36
## Year1999         0.09359    0.25167   0.37   0.71
## Year2000        -0.07587    0.19235  -0.39   0.69
## Year2001         0.18607    0.25957   0.72   0.47
## Year2002         0.20798    0.23051   0.90   0.37
## Year2003         0.23494    0.30713   0.76   0.45
## Year2004         0.17989    0.20890   0.86   0.39
## Year2005         0.04404    0.20971   0.21   0.83
```

```

## Year2006      -0.08650    0.25826   -0.33    0.74
## Year2007      0.04326    0.22471    0.19    0.85
## Year2008      0.21067    0.25362    0.83    0.41
## Year2009      0.07320    0.20074    0.36    0.72
## Year2010      0.15219    0.21049    0.72    0.47
## Year2011      0.04639    0.17783    0.26    0.79
## Year2012      0.20446    0.19201    1.06    0.29
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.43
## Multiple R-squared:  0.0731, Adjusted R-squared:  -0.0345
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 163 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.329  0.857   0.958   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      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 1.588 1      1.260
## Year              1.588 16      1.015

```

Residuals from first author



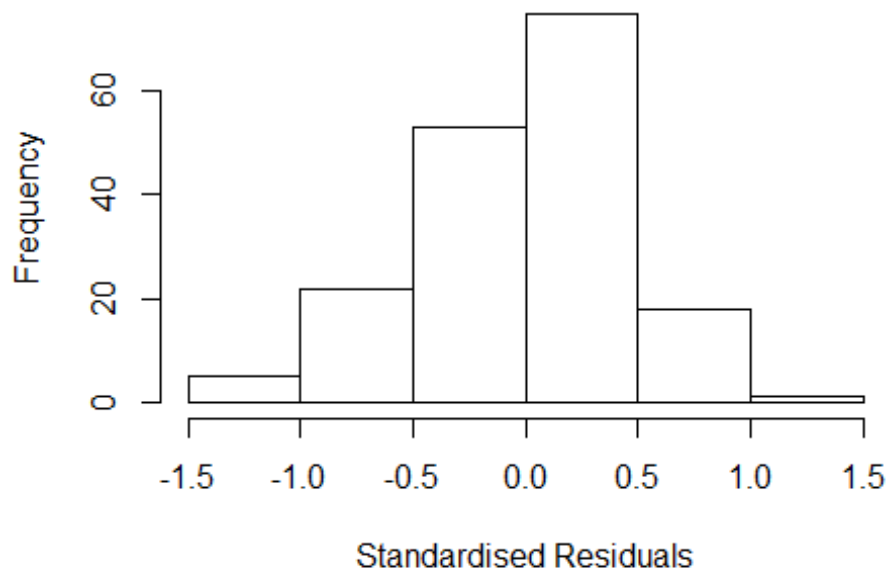
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3272 -0.2637 0.0364 0.2541 1.2989
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0065 0.1600 6.29 3e-09 ***
## FirstAuthorFemale1 -0.0198 0.0842 -0.24 0.81
## Year1997 -0.1615 0.2648 -0.61 0.54
## Year1998 0.3207 0.2481 1.29 0.20
## Year1999 0.1760 0.2280 0.77 0.44
## Year2000 -0.0698 0.1983 -0.35 0.73
## Year2001 0.2371 0.2695 0.88 0.38
## Year2002 0.2760 0.2197 1.26 0.21
## Year2003 0.3091 0.2967 1.04 0.30
## Year2004 0.2499 0.1958 1.28 0.20
## Year2005 0.1042 0.1952 0.53 0.59
## Year2006 -0.0386 0.2605 -0.15 0.88
```

```

## Year2007          0.0965      0.2248      0.43      0.67
## Year2008          0.2302      0.2521      0.91      0.36
## Year2009          0.1283      0.1853      0.69      0.49
## Year2010          0.2001      0.2051      0.98      0.33
## Year2011          0.0816      0.1759      0.46      0.64
## Year2012          0.2055      0.1882      1.09      0.28
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.426
## Multiple R-squared:  0.061, Adjusted R-squared:  -0.0413
## 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.
##  0.311  0.865   0.960   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      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 2.262 1      1.504
## Year            2.262 16      1.026

```


Residuals from last author



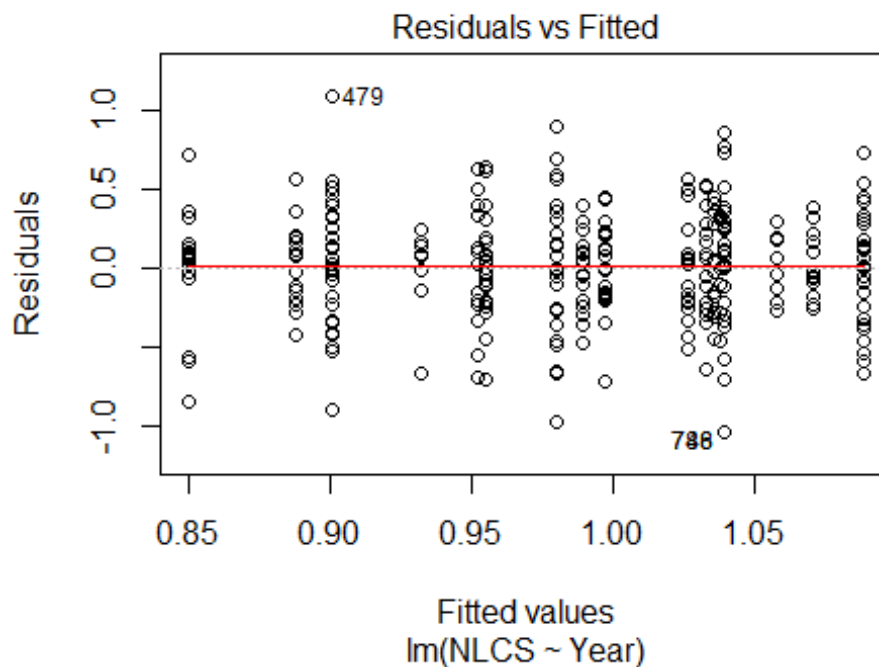
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3101 -0.2803 0.0461 0.2571 1.2364
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0695 0.1733 6.17 5.6e-09 ***
## LastAuthorFemale1 -0.1440 0.1109 -1.30 0.20
## Year1997 -0.1740 0.2549 -0.68 0.50
## Year1998 0.2405 0.2618 0.92 0.36
## Year1999 0.0925 0.2519 0.37 0.71
## Year2000 -0.0744 0.1927 -0.39 0.70
## Year2001 0.1877 0.2601 0.72 0.47
## Year2002 0.2086 0.2311 0.90 0.37
## Year2003 0.2328 0.3039 0.77 0.44
## Year2004 0.1790 0.2087 0.86 0.39
## Year2005 0.0437 0.2100 0.21 0.84
## Year2006 -0.0858 0.2575 -0.33 0.74
```

```

## Year2007          0.0415      0.2217      0.19      0.85
## Year2008          0.2073      0.2449      0.85      0.40
## Year2009          0.0726      0.2004      0.36      0.72
## Year2010          0.1511      0.2104      0.72      0.47
## Year2011          0.0478      0.1783      0.27      0.79
## Year2012          0.2050      0.1930      1.06      0.29
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.431
## Multiple R-squared:  0.0724, Adjusted R-squared:  -0.0287
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 162 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.336  0.855  0.958  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      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 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
##   25   25   23   34   34   47   35   29   27   24   46   38   26   39   30
## 2011 2012
##   63   52
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   13   10    8   17    9   19   16   16   13   14   29   23   18   29   20
## 2011 2012

```

```
## 34 30
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 12 10 8 16 8 17 14 14 12 13 26 20 14 26 16
## 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 = 29, df = 16, p-value = 0.02
```



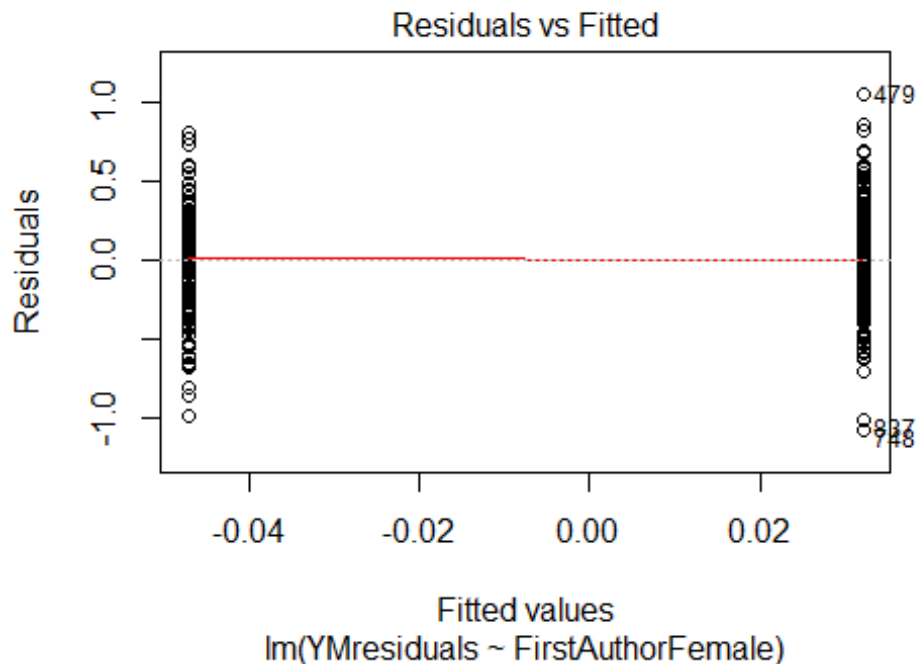
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.15, df = 1, p-value = 0.7

## [1] "Female first author team size 2018 geometric mean: 2.03393700979443"
## [1] "Male first author team size 2018 geometric mean: 3.45597892147664"

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

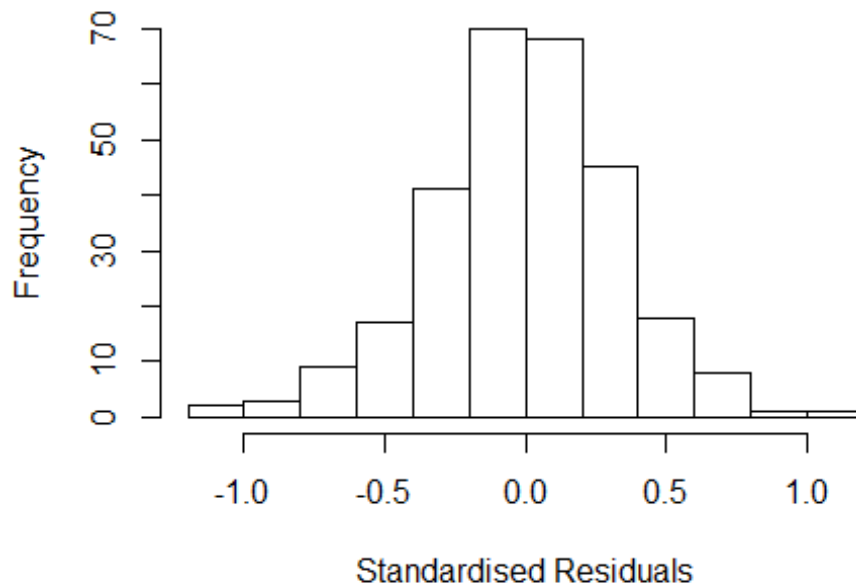
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 24, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.843527111557"
## [1] "Male last author team size 2018 geometric mean: 2.84201826191233"

## Warning in wilcox.test.default(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 = 0.7
## 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 1.409  1      1.187
## LastAuthorFemale  1.288  1      1.135
## UniqueAuthors    4.116  4      1.193
## Year              5.669 16      1.056
```

Residuals from first and last author and team size



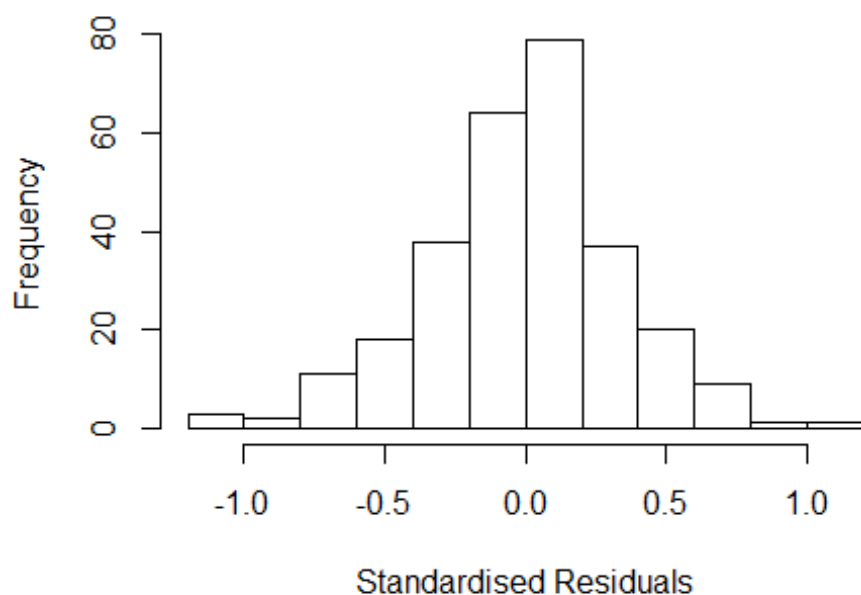
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.168142 -0.207639 -0.000427 0.216969 1.006538
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.00827 0.11928 8.45 2.1e-15 ***
## FirstAuthorFemale1 -0.08401 0.04444 -1.89 0.06 .
## LastAuthorFemale1 0.00923 0.04880 0.19 0.85
## UniqueAuthors2 0.07871 0.07655 1.03 0.30
## UniqueAuthors3 -0.02223 0.07432 -0.30 0.77
## UniqueAuthors4 0.09125 0.09760 0.93 0.35
## UniqueAuthors5 0.02097 0.09358 0.22 0.82
## Year1997 -0.03481 0.12772 -0.27 0.79
## Year1998 0.02003 0.14436 0.14 0.89
## Year1999 -0.06993 0.13471 -0.52 0.60
```

```

## Year2000      -0.01081    0.12193   -0.09    0.93
## Year2001      -0.00556    0.11737   -0.05    0.96
## Year2002      -0.04062    0.13129   -0.31    0.76
## Year2003      -0.16941    0.11853   -1.43    0.15
## Year2004       0.02432    0.12222    0.20    0.84
## Year2005      -0.00031    0.14494    0.00    1.00
## Year2006      -0.11675    0.13777   -0.85    0.40
## Year2007      -0.03559    0.11876   -0.30    0.76
## Year2008      -0.01468    0.13835   -0.11    0.92
## Year2009       0.04159    0.12502    0.33    0.74
## Year2010       0.01841    0.11701    0.16    0.88
## Year2011       0.06863    0.13191    0.52    0.60
## Year2012      -0.01164    0.13891   -0.08    0.93
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.313
## Multiple R-squared:  0.0726, Adjusted R-squared:  -0.00586
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 258 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.135  0.869  0.948  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      3.53e-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 1.354 1      1.163
## LastAuthorFemale  1.281 1      1.132
## Year              1.646 16      1.016

```

Residuals from first and last author



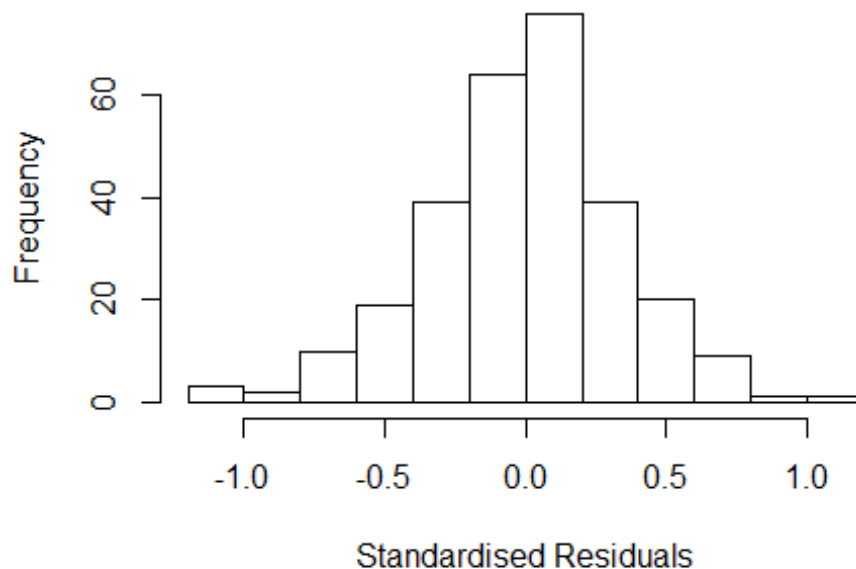
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.116 -0.215  0.012  0.196  1.068
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.03654    0.09365   11.07  <2e-16 ***
## FirstAuthorFemale1 -0.08160    0.04587   -1.78   0.076 .
## LastAuthorFemale1  0.01279    0.04915    0.26   0.795
## Year1997        -0.05394    0.11288   -0.48   0.633
## Year1998         0.02475    0.13611    0.18   0.856
## Year1999        -0.08237    0.12593   -0.65   0.514
## Year2000         0.00694    0.11091    0.06   0.950
## Year2001         0.00870    0.10674    0.08   0.935
## Year2002        -0.02811    0.12972   -0.22   0.829
## Year2003        -0.16967    0.11027   -1.54   0.125
## Year2004         0.03721    0.11101    0.34   0.738
## Year2005        -0.01702    0.13849   -0.12   0.902
```

```

## Year2006      -0.13092    0.12746   -1.03    0.305
## Year2007      -0.03763    0.10944   -0.34    0.731
## Year2008      -0.03969    0.13432   -0.30    0.768
## Year2009       0.04712    0.11776    0.40    0.689
## Year2010       0.01377    0.11088    0.12    0.901
## Year2011       0.07896    0.12209    0.65    0.518
## Year2012      -0.01537    0.12802   -0.12    0.905
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.313
## Multiple R-squared:  0.056, Adjusted R-squared:  -0.00836
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 35 weights are ~= 1. The remaining 248 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.179  0.852  0.944  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      3.53e-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 1.301 1      1.141
## Year              1.301 16      1.008

```


Residuals from first author



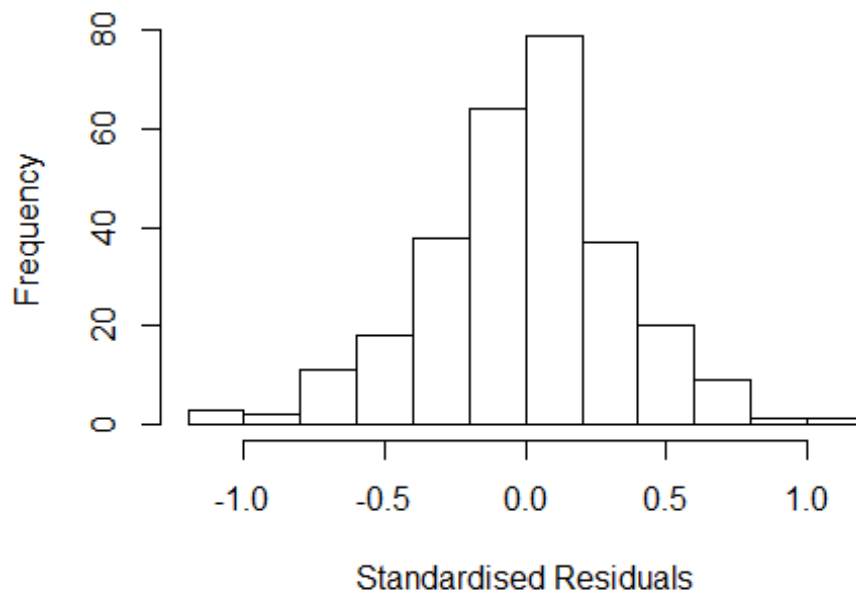
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1179 -0.2108  0.0107  0.1941  1.0794
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.03894    0.09343   11.12  <2e-16 ***
## FirstAuthorFemale1 -0.07967    0.04513   -1.77   0.079 .
## Year1997      -0.05439    0.11317   -0.48   0.631
## Year1998       0.02530    0.13685    0.18   0.853
## Year1999      -0.08543    0.12463   -0.69   0.494
## Year2000       0.00432    0.11071    0.04   0.969
## Year2001       0.00923    0.10664    0.09   0.931
## Year2002      -0.02888    0.12972   -0.22   0.824
## Year2003      -0.16959    0.11026   -1.54   0.125
## Year2004       0.03544    0.11112    0.32   0.750
## Year2005      -0.01681    0.13857   -0.12   0.904
## Year2006      -0.13238    0.12705   -1.04   0.298
```

```

## Year2007          -0.03697    0.10934   -0.34    0.736
## Year2008          -0.03889    0.13407   -0.29    0.772
## Year2009           0.04930    0.11653    0.42    0.673
## Year2010           0.01298    0.11102    0.12    0.907
## Year2011           0.07899    0.12231    0.65    0.519
## Year2012          -0.01559    0.12790   -0.12    0.903
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.314
## Multiple R-squared:  0.056, Adjusted R-squared:  -0.00455
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 34 weights are ~= 1. The remaining 249 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.180  0.852  0.943  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      3.53e-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 1.228 1          1.108
## Year              1.228 16          1.006

```

Residuals from last author



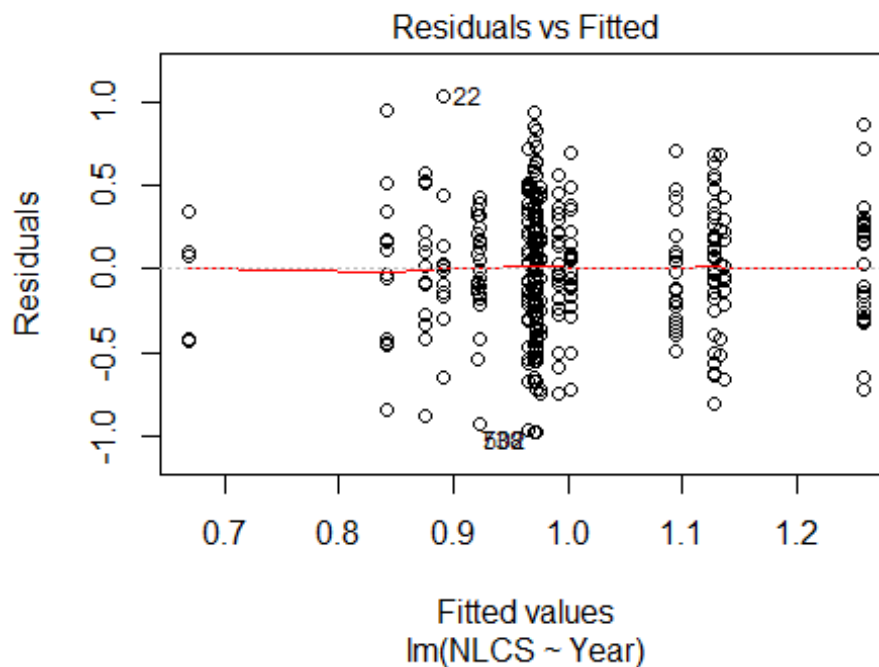
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.09083 -0.21425 -0.00614  0.20658  1.11830
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.00352    0.08972   11.18  <2e-16 ***
## LastAuthorFemale1 -0.00402    0.04857   -0.08    0.93
## Year1997        -0.04352    0.10945   -0.40    0.69
## Year1998         0.03935    0.14040    0.28    0.78
## Year1999        -0.11545    0.12390   -0.93    0.35
## Year2000         0.03015    0.11168    0.27    0.79
## Year2001         0.01551    0.10555    0.15    0.88
## Year2002        -0.02113    0.12593   -0.17    0.87
## Year2003        -0.16933    0.10757   -1.57    0.12
## Year2004         0.04508    0.10804    0.42    0.68
## Year2005         0.00527    0.14002    0.04    0.97
## Year2006        -0.13179    0.12812   -1.03    0.30
```

```

## Year2007          -0.03281      0.11055    -0.30      0.77
## Year2008          -0.02979      0.13429    -0.22      0.82
## Year2009           0.05607      0.11775     0.48      0.63
## Year2010           0.00979      0.10931     0.09      0.93
## Year2011           0.08731      0.12259     0.71      0.48
## Year2012          -0.01136      0.12737    -0.09      0.93
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.312
## Multiple R-squared:  0.0453, Adjusted R-squared:  -0.0159
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 253 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.171  0.857  0.949  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      3.53e-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: 283"
## [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
##   23   18   13   25   15   26   20   30   19   28   28   32   37   66   59
## 2011 2012
##   53   51
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   12   12    5   17    6    9   10   17   13   23   21   25   29   56   39
## 2011 2012

```

```
## 38 33
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 12 11 5 16 6 8 7 16 9 21 20 23 26 53 35
## 2011 2012
## 36 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 = 0.7
```



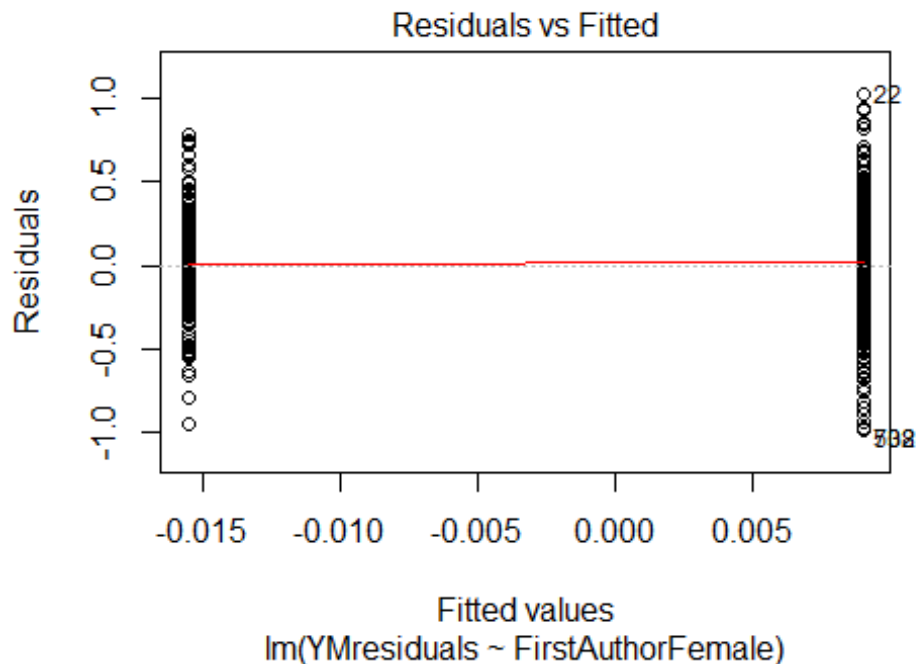
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 5, df = 1, p-value = 0.03

## [1] "Female first author team size 2018 geometric mean: 2.37956557896877"
## [1] "Male first author team size 2018 geometric mean: 2.78500635346593"

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

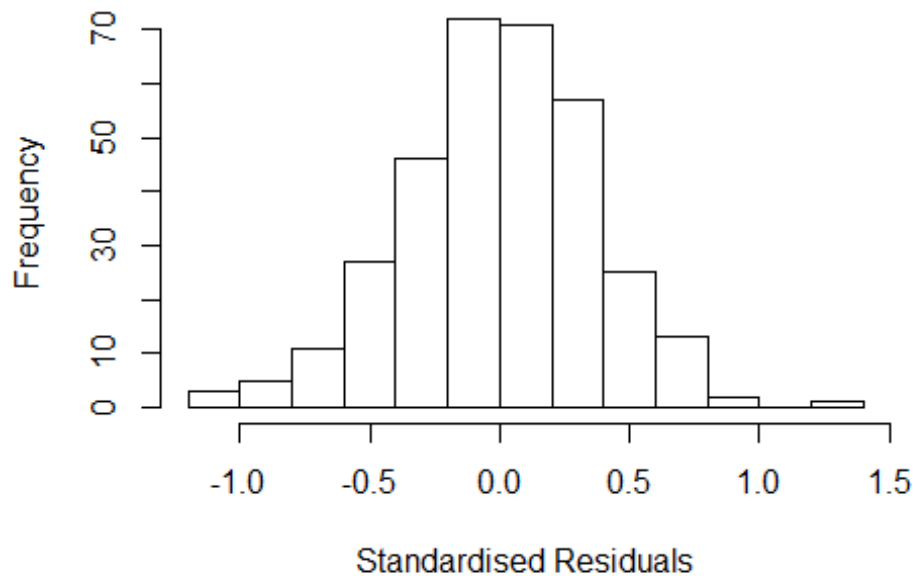
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 39, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.68575127330659"
## [1] "Male last author team size 2018 geometric mean: 2.59514398249042"

## Warning in wilcox.test.default(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 = 0.9
## 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 1.308 1      1.144
## LastAuthorFemale  1.292 1      1.137
## UniqueAuthors    2.626 4      1.128
## Year              2.971 16     1.035
```

Residuals from first and last author and team size



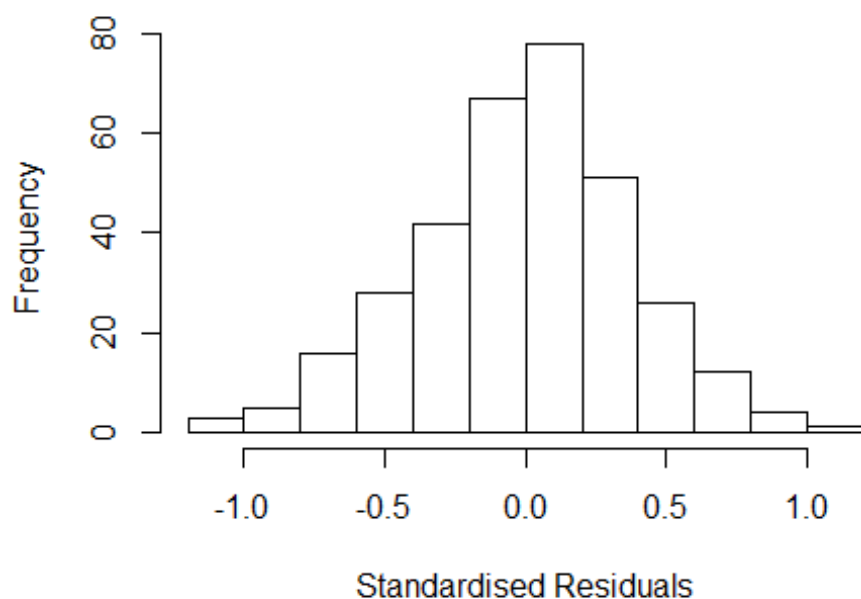
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.07433 -0.22614 0.00864 0.22564 1.24504
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6760 0.1126 6.00 5.4e-09 ***
## FirstAuthorFemale1 -0.0847 0.0455 -1.86 0.0637 .
## LastAuthorFemale1 0.0486 0.0475 1.02 0.3071
## UniqueAuthors2 0.2801 0.0641 4.37 1.7e-05 ***
## UniqueAuthors3 0.3017 0.0720 4.19 3.7e-05 ***
## UniqueAuthors4 0.2559 0.0891 2.87 0.0044 **
## UniqueAuthors5 0.2840 0.0929 3.06 0.0024 **
## Year1997 -0.0111 0.1707 -0.06 0.9484
## Year1998 0.2162 0.2237 0.97 0.3346
## Year1999 0.1373 0.1438 0.95 0.3406
```

```

## Year2000          -0.2371      0.1593    -1.49    0.1376
## Year2001           0.1569      0.1333     1.18    0.2401
## Year2002           0.1434      0.1572     0.91    0.3625
## Year2003           0.0563      0.1269     0.44    0.6577
## Year2004          -0.1959      0.1666    -1.18    0.2405
## Year2005           0.0942      0.1209     0.78    0.4366
## Year2006           0.0786      0.1246     0.63    0.5283
## Year2007           0.0976      0.1217     0.80    0.4228
## Year2008           0.3636      0.1250     2.91    0.0039 **
## Year2009           0.0967      0.1151     0.84    0.4017
## Year2010           0.1091      0.1280     0.85    0.3946
## Year2011           0.0997      0.1385     0.72    0.4721
## Year2012           0.2486      0.1299     1.91    0.0566 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.353
## Multiple R-squared:  0.164, Adjusted R-squared:  0.105
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 33 weights are ~= 1. The remaining 300 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.187  0.860  0.953   0.899   0.985   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      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 1.241 1      1.114
## LastAuthorFemale  1.162 1      1.078
## Year              1.353 16      1.009

```


Residuals from first and last author



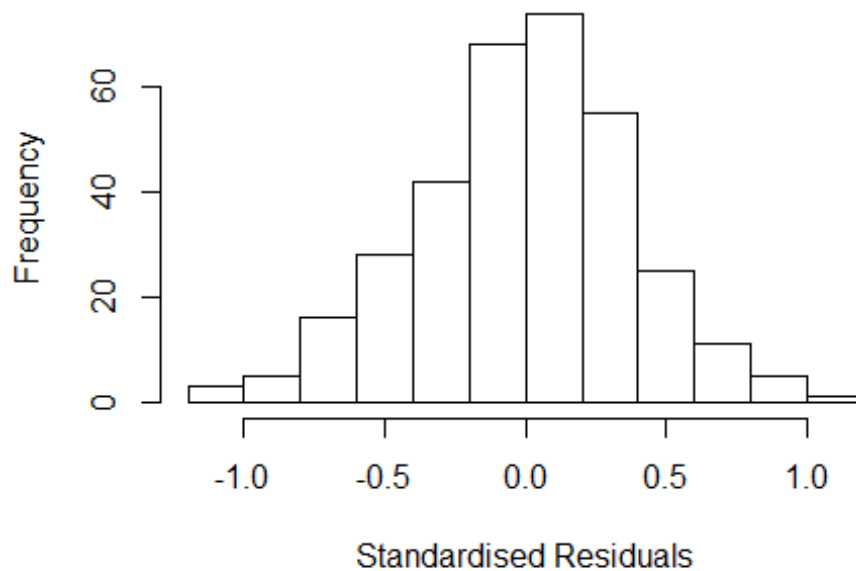
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0082 -0.2384 0.0152 0.2366 1.0866
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.83443 0.09504 8.78 < 2e-16 ***
## FirstAuthorFemale1 -0.05244 0.04514 -1.16 0.24619
## LastAuthorFemale1 0.05269 0.04521 1.17 0.24473
## Year1997 0.00676 0.19324 0.04 0.97210
## Year1998 0.26305 0.25586 1.03 0.30469
## Year1999 0.16186 0.14695 1.10 0.27155
## Year2000 -0.16760 0.17620 -0.95 0.34222
## Year2001 0.16615 0.11696 1.42 0.15645
## Year2002 0.25743 0.14927 1.72 0.08559 .
## Year2003 0.11116 0.12234 0.91 0.36426
## Year2004 -0.08231 0.15910 -0.52 0.60525
## Year2005 0.20655 0.11480 1.80 0.07295 .
```

```

## Year2006          0.15716      0.12286      1.28  0.20175
## Year2007          0.19635      0.11440      1.72  0.08709 .
## Year2008          0.44592      0.12443      3.58  0.00039 ***
## Year2009          0.13969      0.11290      1.24  0.21692
## Year2010          0.17376      0.12140      1.43  0.15336
## Year2011          0.16627      0.13154      1.26  0.20716
## Year2012          0.31314      0.12340      2.54  0.01165 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.366
## Multiple R-squared:  0.0927, Adjusted R-squared:  0.0407
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 27 weights are ~= 1. The remaining 306 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.359  0.861  0.955  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      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 1.202 1          1.097
## Year              1.202 16          1.006

```

Residuals from first author



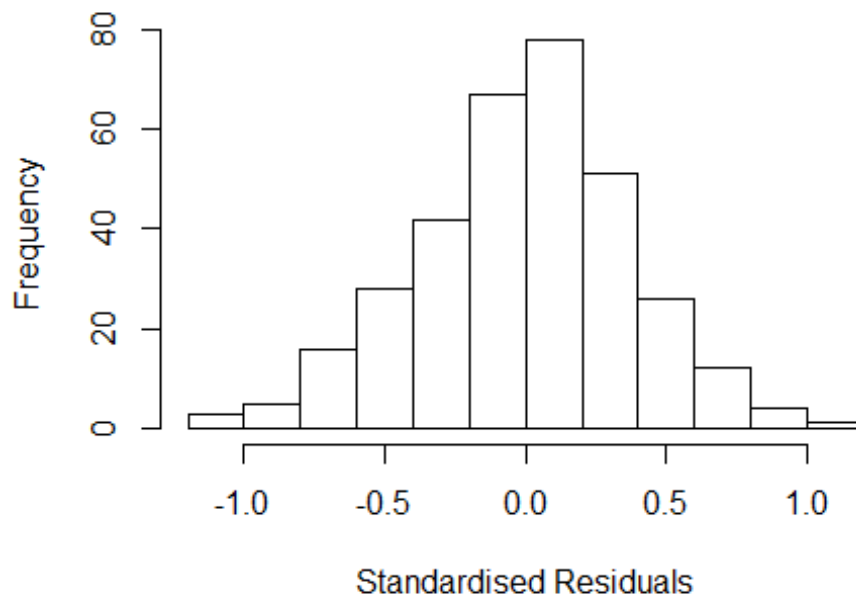
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0115 -0.2435 0.0145 0.2341 1.0719
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.84910 0.09627 8.82 < 2e-16 ***
## FirstAuthorFemale1 -0.04204 0.04444 -0.95 0.34480
## Year1997 0.00143 0.19367 0.01 0.99412
## Year1998 0.26465 0.25822 1.02 0.30621
## Year1999 0.15438 0.14890 1.04 0.30065
## Year2000 -0.16271 0.17770 -0.92 0.36055
## Year2001 0.15410 0.11729 1.31 0.18987
## Year2002 0.24726 0.14906 1.66 0.09815 .
## Year2003 0.10716 0.12319 0.87 0.38503
## Year2004 -0.07267 0.15783 -0.46 0.64553
## Year2005 0.19628 0.11723 1.67 0.09506 .
## Year2006 0.15054 0.12462 1.21 0.22796
```

```

## Year2007          0.19143      0.11468      1.67  0.09605 .
## Year2008          0.44379      0.12564      3.53  0.00047 ***
## Year2009          0.13449      0.11395      1.18  0.23882
## Year2010          0.16243      0.12268      1.32  0.18648
## Year2011          0.15521      0.13303      1.17  0.24421
## Year2012          0.30723      0.12485      2.46  0.01440 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.366
## Multiple R-squared:  0.0898, Adjusted R-squared:  0.0407
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 27 weights are ~= 1. The remaining 306 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.372  0.867   0.956   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      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 1.128 1          1.062
## Year            1.128 16          1.004

```

Residuals from last author



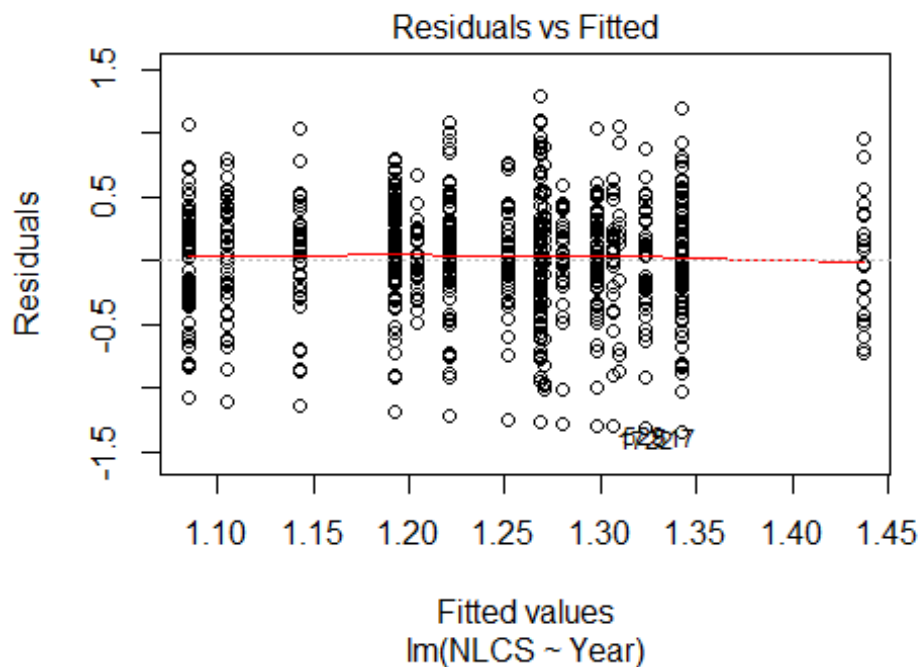
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9819 -0.2316 0.0141 0.2415 1.0934
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.82760 0.09394 8.81 < 2e-16 ***
## LastAuthorFemale1 0.03978 0.04475 0.89 0.37471
## Year1997 -0.00177 0.19788 -0.01 0.99286
## Year1998 0.27395 0.25547 1.07 0.28440
## Year1999 0.16218 0.14482 1.12 0.26362
## Year2000 -0.16628 0.17490 -0.95 0.34248
## Year2001 0.15473 0.11831 1.31 0.19188
## Year2002 0.24357 0.14327 1.70 0.09010 .
## Year2003 0.10313 0.12126 0.85 0.39569
## Year2004 -0.06957 0.15815 -0.44 0.66031
## Year2005 0.19864 0.11327 1.75 0.08045 .
## Year2006 0.14541 0.11967 1.22 0.22523
```

```

## Year2007      0.18165      0.11043      1.64  0.10098
## Year2008      0.43635      0.12304      3.55  0.00045 ***
## Year2009      0.12785      0.11025      1.16  0.24706
## Year2010      0.15325      0.11626      1.32  0.18841
## Year2011      0.15430      0.13015      1.19  0.23671
## Year2012      0.30367      0.12140      2.50  0.01288 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.367
## Multiple R-squared:  0.089, Adjusted R-squared:  0.0399
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 311 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.354  0.864  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      3.00e-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 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
##   40   68   89  108  115   99  113   75  112  104  122  114  132  181  173
## 2011 2012
##  150  166
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   12   22   28   36   38   25   35   28   47   41   41   43   58   94   88
## 2011 2012

```

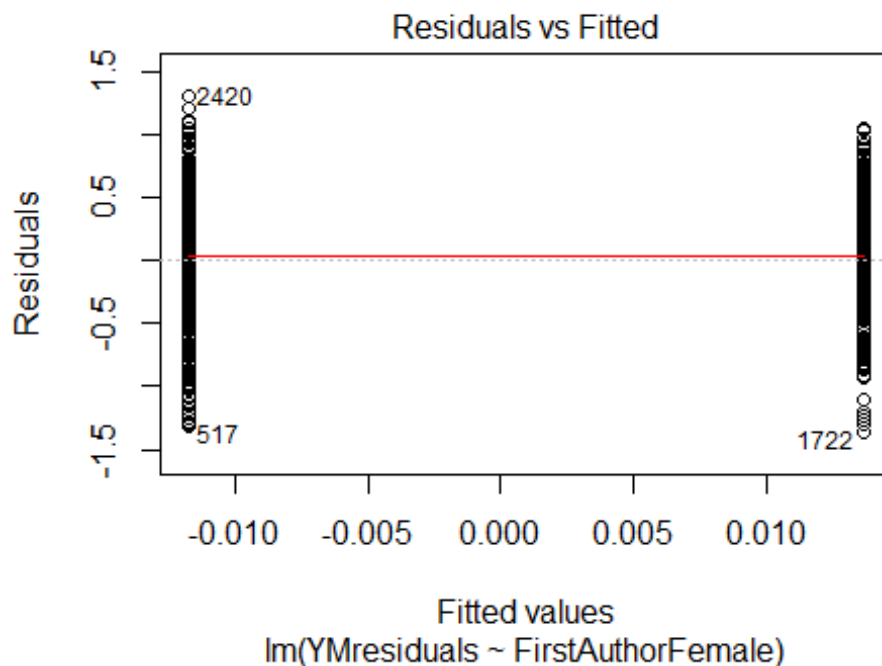
```
##      81    90
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    9   21   23   33   34   19   28   21   38   38   37   39   44   78   69
## 2011 2012
##    63   73
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett 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 = 4.7, df = 1, p-value = 0.03
## [1] "Female first author team size 2018 geometric mean: 4.35273207977458"
## [1] "Male first author team size 2018 geometric mean: 3.86278696947268"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

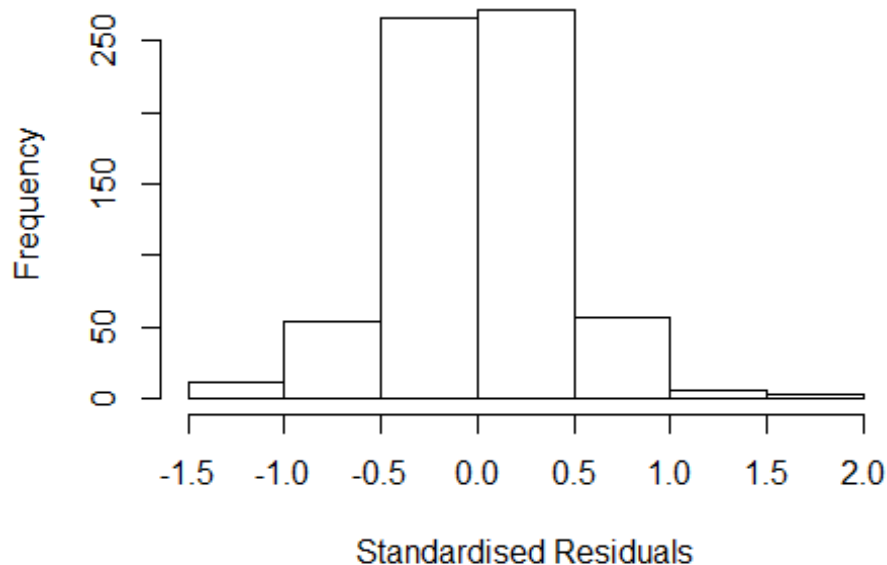
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 820, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.07216614432461"
## [1] "Male last author team size 2018 geometric mean: 4.15196180837399"

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



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 650, 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 1.193  1      1.092
## LastAuthorFemale  1.188  1      1.090
## UniqueAuthors    1.853  4      1.080
## Year              2.192 16      1.025
```


Residuals from first and last author and team size



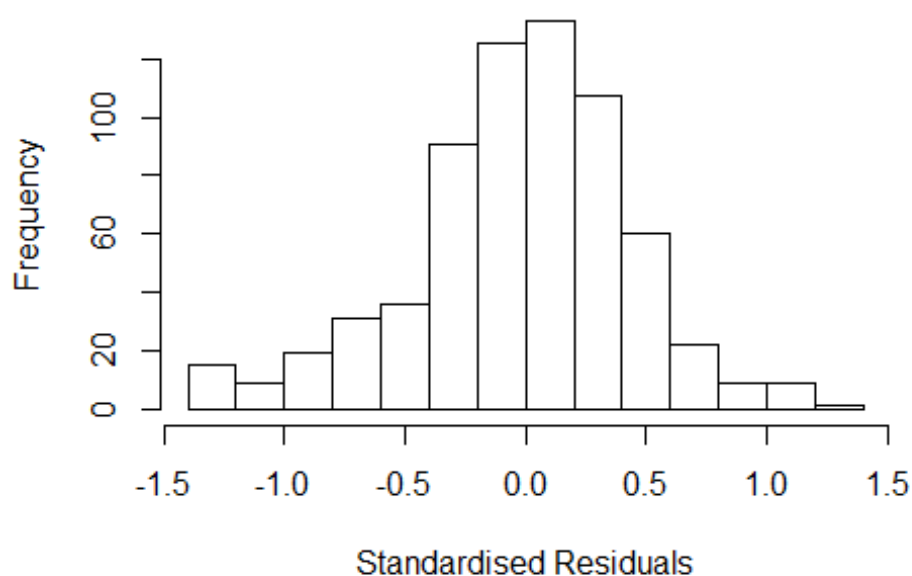
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3352 -0.2422 0.0087 0.2430 1.8642
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.62832 0.25901 2.43 0.016 *
## FirstAuthorFemale1 0.00836 0.03304 0.25 0.800
## LastAuthorFemale1 -0.01696 0.03691 -0.46 0.646
## UniqueAuthors2 0.70084 0.08912 7.86 1.6e-14 ***
## UniqueAuthors3 0.78560 0.08893 8.83 < 2e-16 ***
## UniqueAuthors4 0.80897 0.08782 9.21 < 2e-16 ***
## UniqueAuthors5 0.87690 0.08831 9.93 < 2e-16 ***
## Year1997 0.11730 0.26856 0.44 0.662
## Year1998 0.10773 0.26088 0.41 0.680
## Year1999 -0.21656 0.25970 -0.83 0.405
```

```

## Year2000      -0.01088    0.25805   -0.04    0.966
## Year2001      -0.08160    0.26264   -0.31    0.756
## Year2002      -0.07960    0.26044   -0.31    0.760
## Year2003      -0.15051    0.26100   -0.58    0.564
## Year2004      -0.18503    0.25641   -0.72    0.471
## Year2005      -0.28847    0.25350   -1.14    0.256
## Year2006      -0.33350    0.26097   -1.28    0.202
## Year2007      -0.14221    0.25326   -0.56    0.575
## Year2008      -0.07786    0.25549   -0.30    0.761
## Year2009      -0.06298    0.25094   -0.25    0.802
## Year2010      -0.18407    0.25203   -0.73    0.465
## Year2011      -0.16998    0.25437   -0.67    0.504
## Year2012      -0.17333    0.25611   -0.68    0.499
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.373
## Multiple R-squared:  0.248, Adjusted R-squared:  0.222
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## observation 270 is an outlier with |weight| = 0 ( < 0.00015);
## 49 weights are ~= 1. The remaining 617 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0466 0.8690 0.9540 0.8920 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           1.50e-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 1.130 1 1.063
## LastAuthorFemale 1.142 1 1.069
## Year 1.245 16 1.007

```

Residuals from first and last author



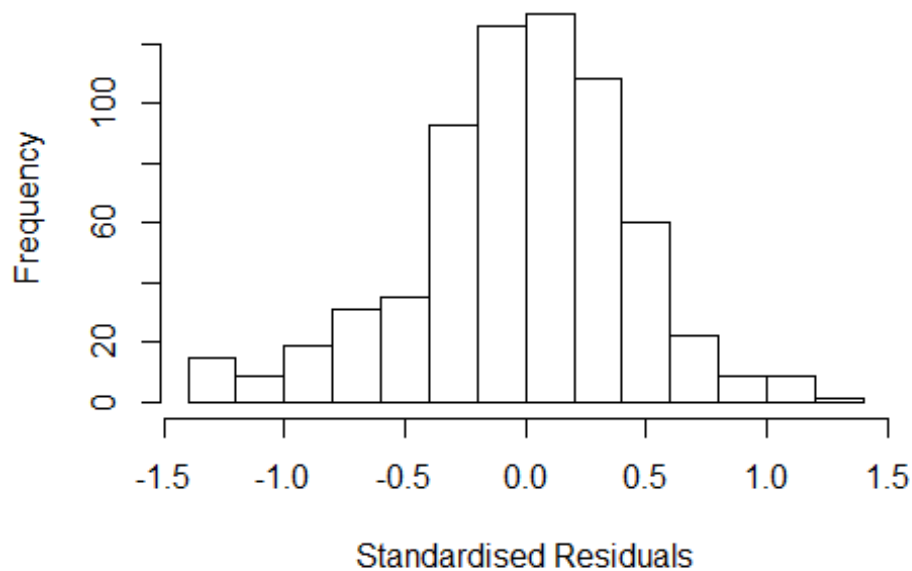
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3501 -0.2540 0.0145 0.2578 1.3199
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.23442 0.23627 5.22 2.4e-07 ***
## FirstAuthorFemale1 0.04307 0.03535 1.22 0.22
## LastAuthorFemale1 -0.02996 0.04006 -0.75 0.45
## Year1997 0.13325 0.26185 0.51 0.61
## Year1998 0.19161 0.25810 0.74 0.46
## Year1999 -0.11751 0.25057 -0.47 0.64
## Year2000 0.11565 0.24888 0.46 0.64
## Year2001 0.03659 0.25923 0.14 0.89
## Year2002 0.05150 0.24386 0.21 0.83
## Year2003 -0.02222 0.24420 -0.09 0.93
## Year2004 -0.08002 0.24897 -0.32 0.75
## Year2005 -0.12944 0.24313 -0.53 0.59
```

```

## Year2006      -0.17951    0.25158   -0.71    0.48
## Year2007      0.03558    0.24308    0.15    0.88
## Year2008      0.10093    0.24475    0.41    0.68
## Year2009      0.09776    0.24096    0.41    0.69
## Year2010     -0.01995    0.24551   -0.08    0.94
## Year2011     -0.01993    0.24349   -0.08    0.93
## Year2012      0.00268    0.24898    0.01    0.99
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.405
## Multiple R-squared:  0.0455, Adjusted R-squared:  0.019
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 55 weights are ~= 1. The remaining 612 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.243  0.865  0.956  0.888  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      1.50e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##      trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## 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
## -1.34463 -0.25855 0.00894 0.25751 1.33036
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.22950 0.23445 5.24 2.1e-07 ***
## FirstAuthorFemale1 0.03849 0.03514 1.10 0.27
## Year1997 0.12898 0.26151 0.49 0.62
## Year1998 0.19653 0.25464 0.77 0.44
## Year1999 -0.11347 0.24839 -0.46 0.65
## Year2000 0.11513 0.24715 0.47 0.64
## Year2001 0.03916 0.25717 0.15 0.88
## Year2002 0.05126 0.24202 0.21 0.83
## Year2003 -0.02351 0.24304 -0.10 0.92
## Year2004 -0.07766 0.24660 -0.31 0.75
## Year2005 -0.12566 0.24126 -0.52 0.60
## Year2006 -0.17936 0.24956 -0.72 0.47
```

```

## Year2007          0.03913    0.24158    0.16    0.87
## Year2008          0.10243    0.24265    0.42    0.67
## Year2009          0.09513    0.23959    0.40    0.69
## Year2010         -0.02224    0.24277   -0.09    0.93
## Year2011         -0.02244    0.24151   -0.09    0.93
## Year2012         -0.00286    0.24822   -0.01    0.99
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.0451, Adjusted R-squared:  0.0201
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 55 weights are ~= 1. The remaining 612 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.239  0.864  0.957  0.886  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.50e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.128 1          1.062
## Year            1.128 16          1.004
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.36448 -0.24910  0.00691  0.25912  1.30072

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.236210   0.231957   5.33 1.4e-07 ***
## LastAuthorFemale1 -0.021567   0.039896  -0.54   0.59
## Year1997       0.162509   0.255563   0.64   0.53
## Year1998       0.206143   0.251993   0.82   0.41
## Year1999      -0.098211   0.244940  -0.40   0.69
## Year2000       0.128268   0.243386   0.53   0.60
## Year2001       0.061642   0.251730   0.24   0.81
## Year2002       0.064947   0.238933   0.27   0.79
## Year2003      -0.012154   0.239581  -0.05   0.96
## Year2004      -0.061520   0.242743  -0.25   0.80
## Year2005      -0.108730   0.237151  -0.46   0.65
## Year2006      -0.162776   0.246607  -0.66   0.51
## Year2007       0.052077   0.237991   0.22   0.83
## Year2008       0.116452   0.238835   0.49   0.63
## Year2009       0.113440   0.235417   0.48   0.63
## Year2010      -0.001611   0.238632  -0.01   0.99
## Year2011      -0.000295   0.237162   0.00   1.00
## Year2012       0.020070   0.243666   0.08   0.93
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.0439, Adjusted R-squared:  0.0188
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 57 weights are ~= 1. The remaining 610 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.225  0.863   0.954   0.885   0.985   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.50e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##   nResample    max.it    best.r.s    k.fast.s    k.max maxit.scale
##      500         50         2         1        1000         200
##   trace.lev    mts    compute.rd
##      0        1000         0
##           psi          subsampling          cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 667"
## [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
##    1    2    2    5   11    9   11    7    5   10    9   18   15   16   10
## 2011 2012
##   12   13
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    0    0    4    3    6    8    1    2    9    8   15    9   11    9
## 2011 2012
##    9    6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    0    0    4    2    6    6    1    2    8    8   12    8    9    6
## 2011 2012
##    9    5
## [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: 3.17767152314644"

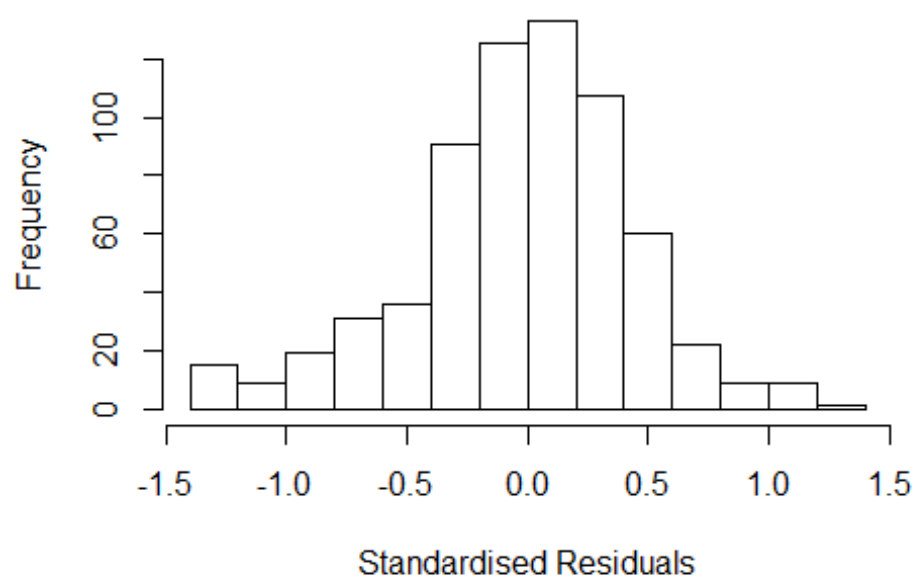
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 11, p-value = 0.3
## 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: 3.14734519026494"

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

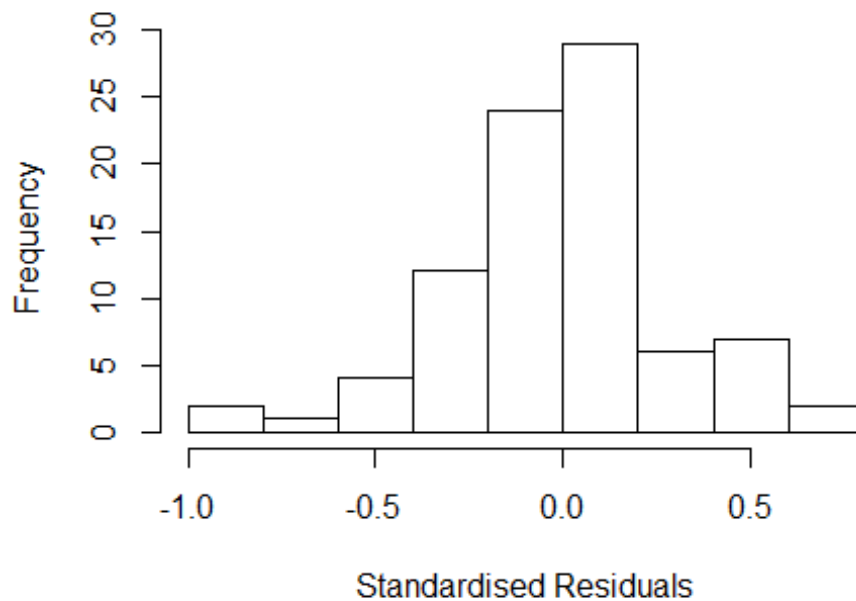
```


Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 11, 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 2.216e+13  1      4.707e+06
## LastAuthorFemale  8.353e+13  1      9.139e+06
## UniqueAuthors    2.762e+02  4      2.019e+00
## Year              6.187e+16 14      3.978e+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
## -0.920122 -0.161639 0.000932 0.132456 0.714135
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.09657 0.24432 4.49 2.9e-05 ***
## FirstAuthorFemale1 0.01858 0.09354 0.20 0.843
## LastAuthorFemale1 0.10722 0.18159 0.59 0.557
## UniqueAuthors2 -0.25370 0.15396 -1.65 0.104
## UniqueAuthors3 0.00821 0.18785 0.04 0.965
## UniqueAuthors4 -0.04080 0.13608 -0.30 0.765
## UniqueAuthors5 -0.13782 0.14708 -0.94 0.352
## Year1999 0.36161 0.78149 0.46 0.645
## Year2000 -0.37561 0.20406 -1.84 0.070 .
## Year2001 -0.14581 0.17207 -0.85 0.400
```

```

## Year2002      -0.01784    0.19558   -0.09    0.928
## Year2003      0.11064    0.24145    0.46    0.648
## Year2004     -0.03028    0.17740   -0.17    0.865
## Year2005      0.15193    0.17838    0.85    0.397
## Year2006      0.19959    0.18029    1.11    0.272
## Year2007      0.20799    0.19244    1.08    0.284
## Year2008      0.08776    0.21414    0.41    0.683
## Year2009      0.41476    0.19905    2.08    0.041 *
## Year2010      0.35023    0.16745    2.09    0.040 *
## Year2011      0.23056    0.16566    1.39    0.169
## Year2012      0.37256    0.19034    1.96    0.055 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.28
## Multiple R-squared:  0.375, Adjusted R-squared:  0.186
## Convergence in 38 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 74 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.260  0.858  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      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"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 5.704e+12  1      2.388e+06
## LastAuthorFemale  1.358e+13  1      3.684e+06
## Year              1.160e+14 14      3.179e+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:

```

```

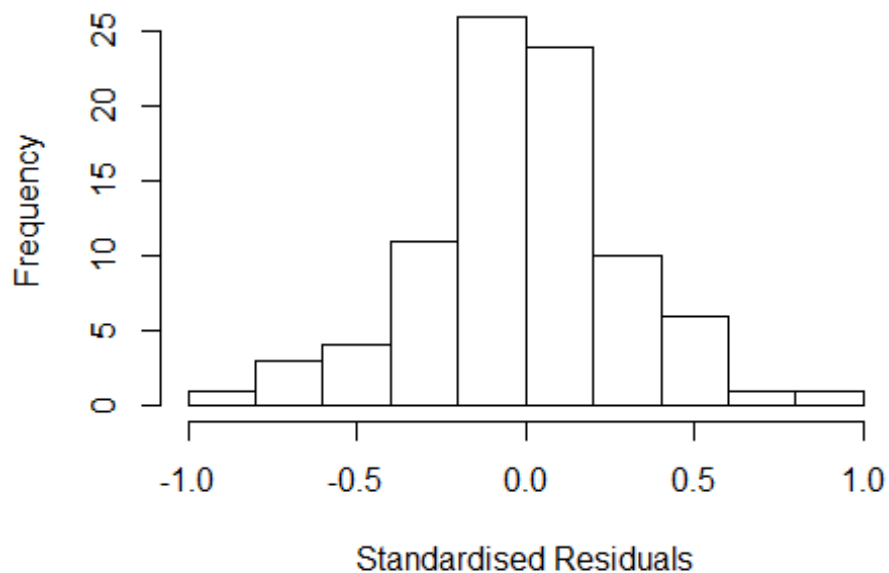
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q      Median        3Q        Max
## -8.71e-01 -1.46e-01 -3.03e-15  1.52e-01  8.92e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.01739    0.12446   8.17 8.7e-12 ***
## FirstAuthorFemale1  0.00722    0.08068   0.09  0.9289
## LastAuthorFemale1  0.19461    0.12446   1.56  0.1224
## Year1999         0.24854    0.56700   0.44  0.6625
## Year2000        -0.43801    0.14422  -3.04  0.0034 **
## Year2001        -0.22528    0.14094  -1.60  0.1145
## Year2002         0.04803    0.18156   0.26  0.7921
## Year2003         0.20938    0.16424   1.27  0.2066
## Year2004         0.03261    0.14032   0.23  0.8169
## Year2005         0.08771    0.17734   0.49  0.6224
## Year2006         0.18038    0.15719   1.15  0.2551
## Year2007         0.12339    0.17731   0.70  0.4888
## Year2008        -0.06243    0.13589  -0.46  0.6473
## Year2009         0.38955    0.14137   2.76  0.0075 **
## Year2010         0.21697    0.13779   1.57  0.1199
## Year2011         0.23444    0.15422   1.52  0.1330
## Year2012         0.42182    0.15269   2.76  0.0073 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.287
## Multiple R-squared:  0.331, Adjusted R-squared:  0.178
## Convergence in 27 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 80 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.312  0.871  0.965  0.901  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      1.15e-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"

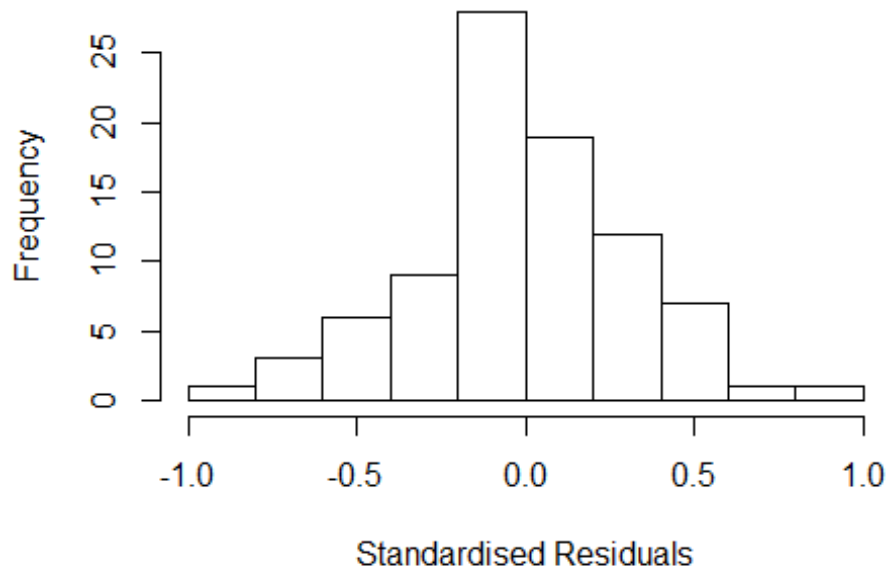
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead
```

Residuals from first and last author



```
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 4.572e+13  1      6.762e+06
## Year              4.572e+13 14      3.075e+00
```

Residuals from first author



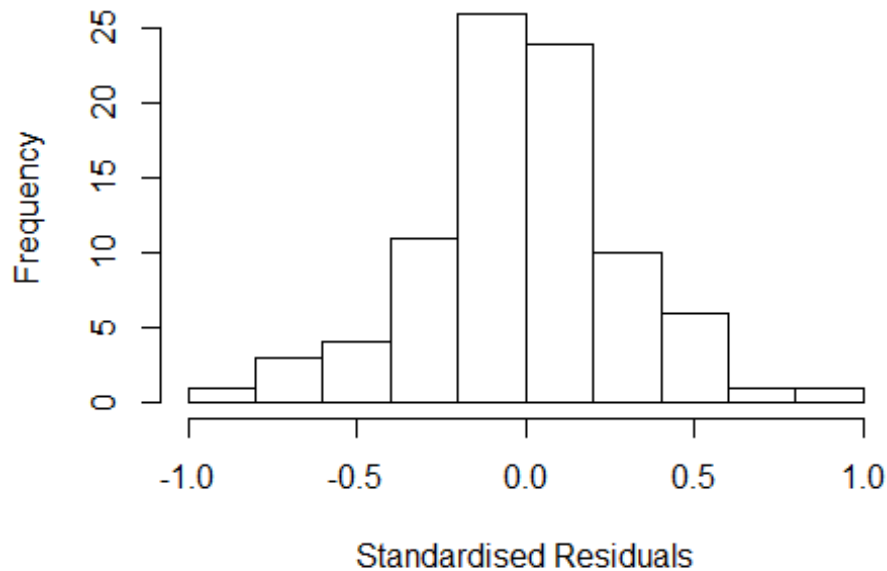
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.80767 -0.16597 -0.00583 0.19365 0.80589
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.21200 0.00000 Inf < 2e-16 ***
## FirstAuthorFemale1 0.04523 0.07716 0.59 0.5596
## Year1999 -0.00933 0.41678 -0.02 0.9822
## Year2000 -0.65162 0.06174 -10.55 3.4e-16 ***
## Year2001 -0.41917 0.06924 -6.05 6.0e-08 ***
## Year2002 -0.10779 0.16109 -0.67 0.5056
## Year2003 -0.02323 0.07716 -0.30 0.7642
## Year2004 -0.16200 0.06458 -2.51 0.0144 *
## Year2005 -0.12139 0.11822 -1.03 0.3080
## Year2006 -0.03153 0.07343 -0.43 0.6690
## Year2007 -0.06004 0.12943 -0.46 0.6441
## Year2008 -0.17089 0.18943 -0.90 0.3700
```

```

## Year2009          0.25144    0.12027    2.09    0.0402 *
## Year2010          0.04462    0.11148    0.40    0.6902
## Year2011          0.03355    0.08049    0.42    0.6781
## Year2012          0.21916    0.08195    2.67    0.0093 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.319
## Multiple R-squared:  0.29,   Adjusted R-squared:  0.14
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 76 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.500  0.867  0.962  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      1.15e-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 1.43e+14 1      1.196e+07
## Year              1.43e+14 14      3.203e+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
## -8.69e-01 -1.43e-01 6.11e-16 1.50e-01 8.86e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0153 0.1192 8.52 1.8e-12 ***
## LastAuthorFemale1 0.1967 0.1192 1.65 0.1032
## Year1999 0.2483 0.5146 0.48 0.6309
## Year2000 -0.4323 0.1234 -3.50 0.0008 ***
## Year2001 -0.2231 0.1366 -1.63 0.1069
## Year2002 0.0519 0.1713 0.30 0.7630
## Year2003 0.2187 0.1192 1.84 0.0706 .
## Year2004 0.0347 0.1356 0.26 0.7987
## Year2005 0.0924 0.1667 0.55 0.5812
## Year2006 0.1852 0.1356 1.37 0.1763
## Year2007 0.1265 0.1631 0.78 0.4408
## Year2008 -0.0539 0.1089 -0.50 0.6220
```



```

## Year2009          0.3942      0.1243      3.17      0.0022 **
## Year2010          0.2209      0.1328      1.66      0.1007
## Year2011          0.2373      0.1443      1.64      0.1045
## Year2012          0.4254      0.1427      2.98      0.0039 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.29
## Multiple R-squared:  0.328, Adjusted R-squared:  0.186
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 80 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.332  0.875  0.966  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      1.15e-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 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 2007 2008 2009 2010 2011
##    3    3    3    7    2    6    3    3    3    6    1    6    3    5    3
## 2012
##    1
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2007 2008 2009 2010 2011
##    1    0    0    2    0    1    0    1    1    1    0    2    2    4    3
## 2012
##    0
##

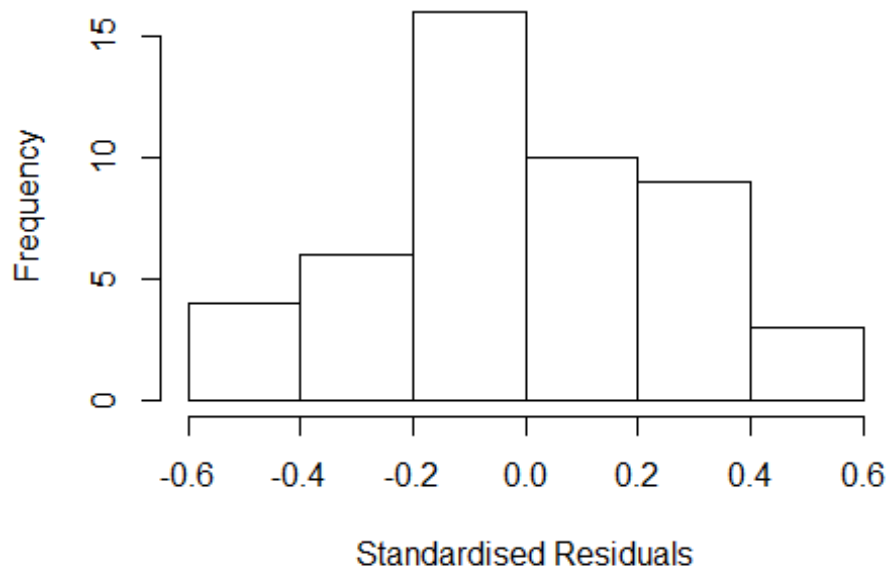
```

```

## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2007 2008 2009 2010 2011
##      1      0      0      2      0      1      0      1      1      1      0      2      1      2      3
## 2012
##      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: 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: 15"
## [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
##      4      2      1      2      3      5      4      2      5      5      7      4      5     10      9
## 2011 2012
##      3      9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      1      2      1      1      1      2      2      1      2      3      7      4      4      6      8
## 2011 2012
##      3      5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      0      2      1      0      1      2      2      1      2      3      6      4      4      5      7
## 2011 2012
##      3      5
## [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"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.165e+01 1          3.413
## LastAuthorFemale  6.948e+00 1          2.636
## UniqueAuthors    1.855e+30 4          6075.124
## Year              1.302e+31 14          12.919

```

Residuals from first and last author and team size



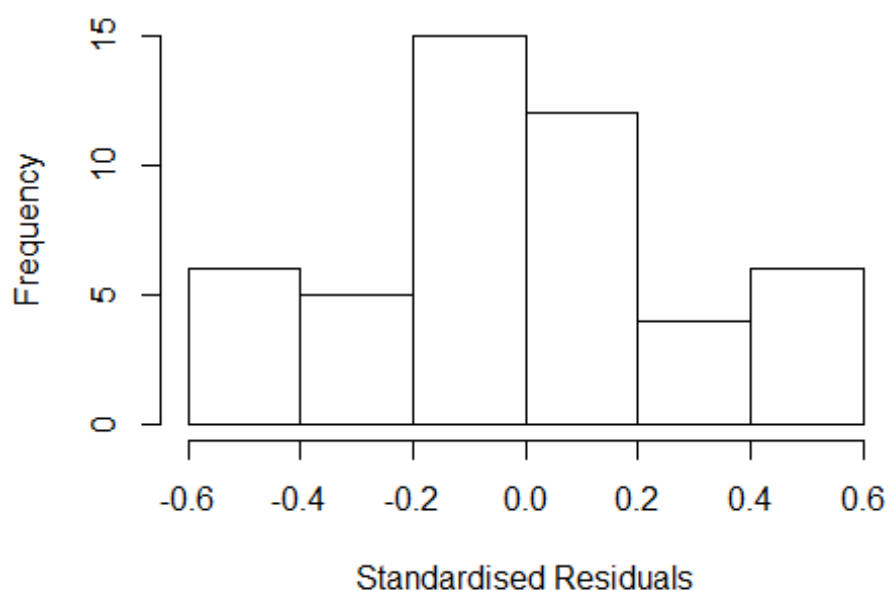
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -5.46e-01 -1.66e-01 -2.78e-16 1.98e-01 5.72e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5333 0.3860 1.38 0.178
## FirstAuthorFemale1 0.1423 0.0970 1.47 0.154
## LastAuthorFemale1 0.1136 0.1144 0.99 0.330
## UniqueAuthors2 0.2198 0.1679 1.31 0.202
## UniqueAuthors3 0.0838 0.2035 0.41 0.684
## UniqueAuthors4 0.3683 0.1869 1.97 0.059 .
## UniqueAuthors5 0.2748 0.1901 1.45 0.160
## Year1998 0.0668 0.3694 0.18 0.858
## Year2000 -0.5333 0.3860 -1.38 0.178
## Year2001 0.2523 0.3854 0.65 0.518
```

```

## Year2002          0.3985      0.6246      0.64      0.529
## Year2003          0.6763      0.3929      1.72      0.097 .
## Year2004         -0.2182      1.1479     -0.19      0.851
## Year2005          0.3779      0.3625      1.04      0.307
## Year2006          0.1376      0.3842      0.36      0.723
## Year2007          0.2523      0.3721      0.68      0.504
## Year2008          0.2352      0.3769      0.62      0.538
## Year2009          0.3408      0.3954      0.86      0.396
## Year2010         -0.2322      0.3717     -0.62      0.537
## Year2011          0.3480      0.4317      0.81      0.427
## Year2012          0.2852      0.3829      0.74      0.463
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.305
## Multiple R-squared:  0.533, Adjusted R-squared:  0.187
## Convergence in 31 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 41 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.704  0.882  0.956  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      2.08e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "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      1.359
## LastAuthorFemale  9.187 1      3.031
## Year              20.609 14      1.114

```

Residuals from first and last author



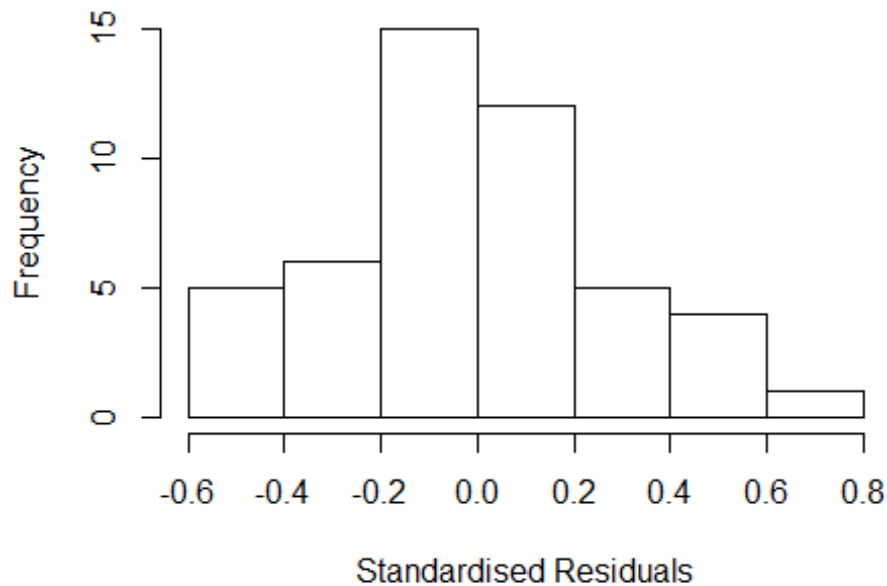
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -5.83e-01 -1.71e-01 -5.27e-16 1.38e-01 5.53e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5480 0.5179 1.06 0.2982
## FirstAuthorFemale1 0.2795 0.0922 3.03 0.0049 **
## LastAuthorFemale1 0.1574 0.1031 1.53 0.1367
## Year1998 0.2720 0.5179 0.53 0.6032
## Year2000 -0.5480 0.5179 -1.06 0.2982
## Year2001 0.4418 0.5233 0.84 0.4051
## Year2002 0.4458 0.7633 0.58 0.5635
## Year2003 1.0300 0.5179 1.99 0.0556 .
## Year2004 -0.0130 1.2032 -0.01 0.9914
## Year2005 0.4983 0.5281 0.94 0.3526
## Year2006 0.3294 0.5316 0.62 0.5400
## Year2007 0.2271 0.5249 0.43 0.6683
```

```

## Year2008          0.4016      0.5261      0.76      0.4511
## Year2009          0.5214      0.5396      0.97      0.3414
## Year2010         -0.1388      0.5237     -0.27      0.7927
## Year2011          0.3703      0.5766      0.64      0.5255
## Year2012          0.3746      0.5264      0.71      0.4821
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.305
## Multiple R-squared:  0.497, Adjusted R-squared:  0.237
## Convergence in 31 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 41 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.695  0.838  0.955  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      2.08e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.456 1          1.859
## Year              3.456 14          1.045

```

Residuals from first author



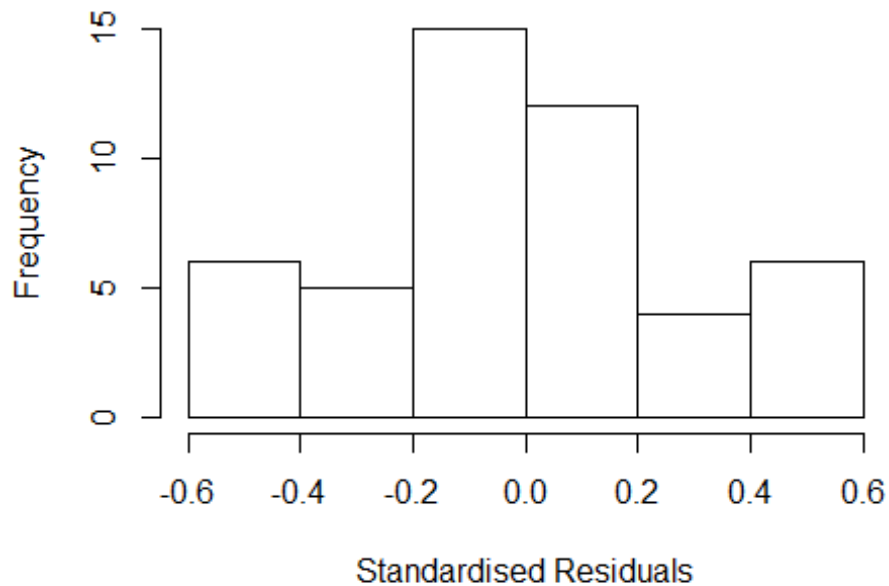
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -5.73e-01 -1.56e-01 -6.66e-16 1.88e-01 6.36e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5247 0.4727 1.11 0.2752
## FirstAuthorFemale1 0.3028 0.1086 2.79 0.0088 **
## Year1998 0.2953 0.4727 0.62 0.5366
## Year2000 -0.5247 0.4727 -1.11 0.2752
## Year2001 0.5438 0.4789 1.14 0.2645
## Year2002 0.5478 0.5828 0.94 0.3543
## Year2003 1.0533 0.4727 2.23 0.0330 *
## Year2004 0.0103 0.9553 0.01 0.9915
## Year2005 0.5685 0.4888 1.16 0.2535
## Year2006 0.4036 0.4879 0.83 0.4142
## Year2007 0.3369 0.4771 0.71 0.4852
## Year2008 0.4420 0.4923 0.90 0.3760
```

```

## Year2009          0.6188      0.5104      1.21      0.2342
## Year2010         -0.1052      0.4726     -0.22      0.8252
## Year2011          0.3832      0.5374      0.71      0.4810
## Year2012          0.5181      0.4700      1.10      0.2785
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.33
## Multiple R-squared:  0.464, Adjusted R-squared:  0.213
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 40 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.690  0.884  0.958  0.922  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      2.08e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##  nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##  trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 20.05 1          4.478
## Year            20.05 14          1.113

```


Residuals from last author



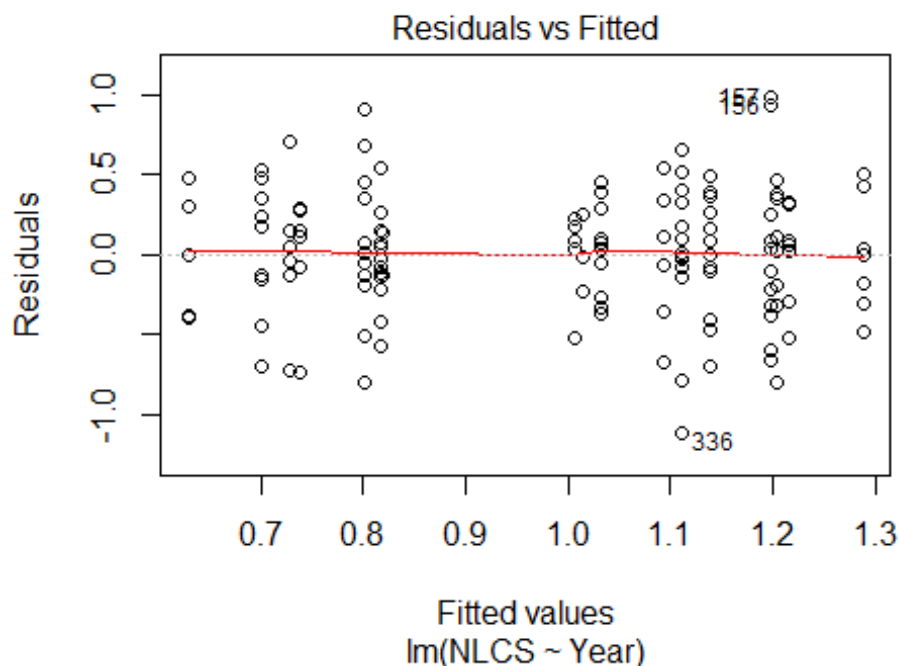
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -5.77e-01 -2.01e-01 -8.54e-17 1.87e-01 5.68e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8275 0.4793 1.73 0.094 .
## LastAuthorFemale1 0.1760 0.1011 1.74 0.091 .
## Year1998 -0.0075 0.4793 -0.02 0.988
## Year2000 -0.8275 0.4793 -1.73 0.094 .
## Year2001 0.1530 0.4823 0.32 0.753
## Year2002 0.1570 0.7276 0.22 0.831
## Year2003 0.7505 0.4793 1.57 0.127
## Year2004 -0.2925 1.0344 -0.28 0.779
## Year2005 0.3031 0.4810 0.63 0.533
## Year2006 0.0793 0.5022 0.16 0.876
## Year2007 0.2127 0.4955 0.43 0.671
## Year2008 0.2002 0.5011 0.40 0.692
```

```

## Year2009          0.2343      0.5013      0.47      0.643
## Year2010         -0.2319      0.4977     -0.47      0.644
## Year2011          0.0845      0.5423      0.16      0.877
## Year2012          0.1378      0.4905      0.28      0.781
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.319
## Multiple R-squared:  0.449, Adjusted R-squared:  0.191
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 40 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.724  0.840  0.947  0.913  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      2.08e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 48"
## [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
##   16   13    9   24   13   12   19   11   15   25    8   16   15    8   13
## 2011 2012
##   13   17
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    6    3   11    5    2   12    5   11   15    8   12   12    7    7
## 2011 2012
##    7   14
##

```

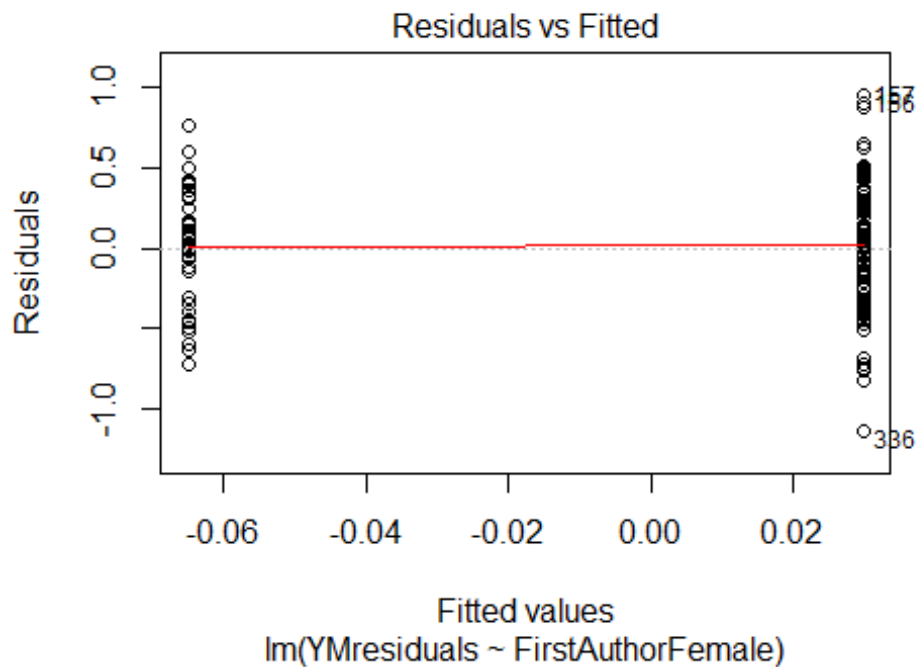
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      6      6      3     11      5      1     10      5      9     13      7      9     11      7      6
## 2011 2012
##      4     13
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"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



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.34, df = 1, p-value = 0.6
##
## [1] "Female first author team size 2018 geometric mean: 3"
## [1] "Male first author team size 2018 geometric mean: 2.95761686116672"
## Warning in wilcox.test.default(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] "Female last author team size 2018 geometric mean: 4.16016764610381"
## [1] "Male last author team size 2018 geometric mean: 2.73967946137285"

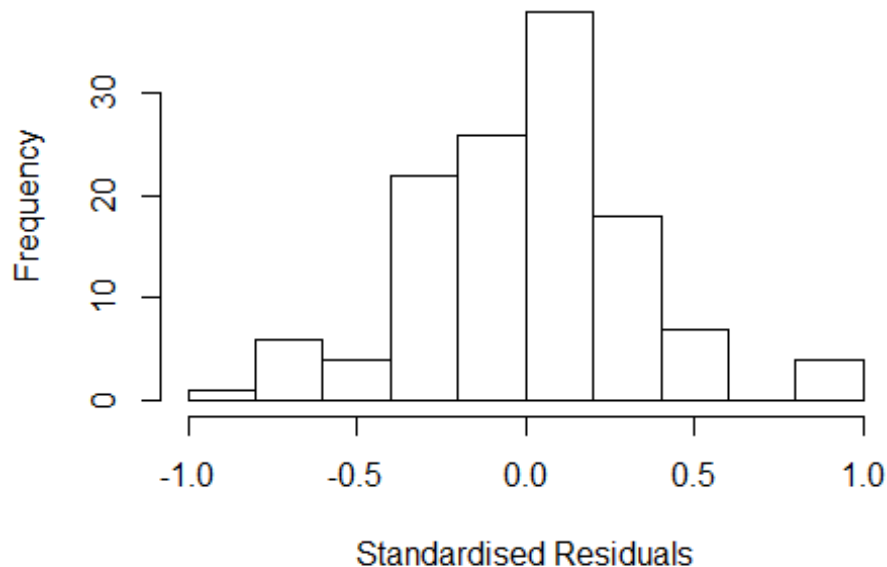
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 29, p-value = 0.2
## 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 | 2.551 | 1 | 1.597 |
| LastAuthorFemale | 3.612 | 1 | 1.900 |
| UniqueAuthors | 215.601 | 4 | 1.958 |
| Year | 287.373 | 16 | 1.194 |

Residuals from first and last author and team size



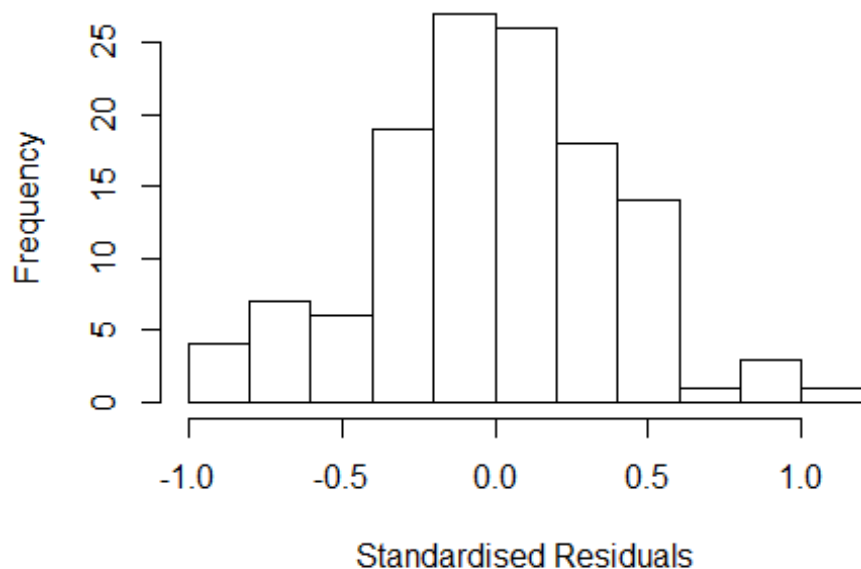
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8894 -0.2124 0.0122 0.1905 0.9601
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5685 0.1479 3.84 0.00021 ***
## FirstAuthorFemale1 -0.0371 0.0861 -0.43 0.66789
## LastAuthorFemale1 0.0488 0.0861 0.57 0.57176
## UniqueAuthors2 0.5171 0.1015 5.09 1.6e-06 ***
## UniqueAuthors3 0.2860 0.1261 2.27 0.02535 *
## UniqueAuthors4 0.5320 0.1647 3.23 0.00166 **
## UniqueAuthors5 0.6898 0.1960 3.52 0.00065 ***
## Year1997 -0.2606 0.1826 -1.43 0.15655
## Year1998 0.2550 0.1862 1.37 0.17388
## Year1999 -0.1560 0.1783 -0.87 0.38374
```

```

## Year2000          -0.2622      0.2092    -1.25   0.21306
## Year2001           0.3845      0.1479     2.60   0.01072 *
## Year2002           0.0882      0.1802     0.49   0.62537
## Year2003           0.0830      0.1901     0.44   0.66344
## Year2004           0.1825      0.3068     0.59   0.55326
## Year2005          -0.0823      0.1925    -0.43   0.66991
## Year2006           0.2048      0.2585     0.79   0.42985
## Year2007           0.1416      0.2448     0.58   0.56417
## Year2008          -0.2676      0.2641    -1.01   0.31329
## Year2009           0.0106      0.2486     0.04   0.96600
## Year2010           0.2286      0.1848     1.24   0.21898
## Year2011           0.2433      0.1677     1.45   0.14980
## Year2012           0.2551      0.1913     1.33   0.18538
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.336
## Multiple R-squared:  0.418, Adjusted R-squared:  0.294
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 116 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.395  0.899  0.961  0.907  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      7.94e-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 2.256 1      1.502
## LastAuthorFemale  3.171 1      1.781
## Year              5.631 16      1.055

```

Residuals from first and last author



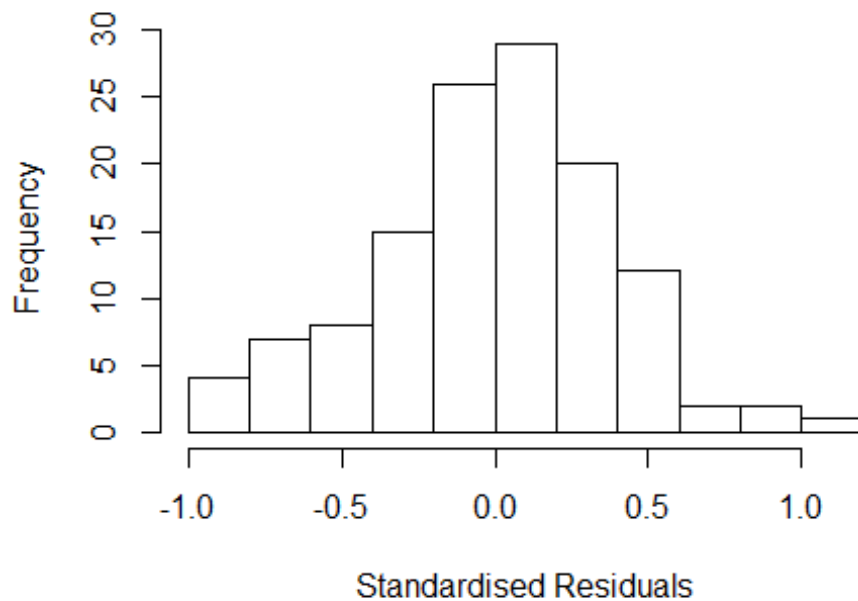
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q      Median        3Q       Max
## -0.840163 -0.230454  0.000462  0.287775  1.002419
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.7046      0.1941   3.63  0.00043 ***
## FirstAuthorFemale1 -0.0876      0.0904  -0.97  0.33454
## LastAuthorFemale1  0.1093      0.0916   1.19  0.23531
## Year1997          0.0681      0.2379   0.29  0.77518
## Year1998          0.3096      0.2266   1.37  0.17462
## Year1999          0.0324      0.2438   0.13  0.89465
## Year2000         -0.0787      0.2646  -0.30  0.76670
## Year2001          0.2484      0.1941   1.28  0.20335
## Year2002          0.3600      0.2192   1.64  0.10350
## Year2003          0.3445      0.2215   1.56  0.12279
## Year2004          0.4730      0.2795   1.69  0.09353 .
## Year2005          0.1813      0.2107   0.86  0.39149
```

```

## Year2006          0.5280      0.2575      2.05  0.04279 *
## Year2007          0.4325      0.2358      1.83  0.06942 .
## Year2008          0.1355      0.2805      0.48  0.63001
## Year2009          0.4249      0.2543      1.67  0.09767 .
## Year2010          0.4669      0.2205      2.12  0.03657 *
## Year2011          0.4561      0.1961      2.33  0.02191 *
## Year2012          0.5133      0.2068      2.48  0.01462 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.394
## Multiple R-squared:  0.216, Adjusted R-squared:  0.0837
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 113 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.498  0.884  0.949  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      7.94e-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 1.95 1      1.396
## Year              1.95 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
## -0.8513 -0.2157 0.0165 0.2736 1.0041
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.75342 0.18159 4.15 6.7e-05 ***
## FirstAuthorFemale1 -0.07687 0.08949 -0.86 0.392
## Year1997 0.03874 0.22715 0.17 0.865
## Year1998 0.26083 0.21601 1.21 0.230
## Year1999 0.00696 0.23966 0.03 0.977
## Year2000 -0.12751 0.25626 -0.50 0.620
## Year2001 0.19958 0.18159 1.10 0.274
## Year2002 0.31920 0.20582 1.55 0.124
## Year2003 0.31393 0.21617 1.45 0.149
## Year2004 0.42244 0.26884 1.57 0.119
## Year2005 0.13014 0.19774 0.66 0.512
## Year2006 0.49220 0.25802 1.91 0.059 .
```

```

## Year2007          0.42479    0.23732    1.79    0.076 .
## Year2008          0.09793    0.27595    0.35    0.723
## Year2009          0.39076    0.24517    1.59    0.114
## Year2010          0.46597    0.22737    2.05    0.043 *
## Year2011          0.49138    0.21510    2.28    0.024 *
## Year2012          0.48637    0.20327    2.39    0.018 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.394
## Multiple R-squared:  0.207, Adjusted R-squared:  0.0823
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 112 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.496  0.891  0.951  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      7.94e-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 3.083 1          1.756
## Year          3.083 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
## -0.86846 -0.21603  0.00353  0.26011  1.01915

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.6880    0.1880    3.66 0.00039 ***
## LastAuthorFemale1 0.0973    0.0889    1.09 0.27604
## Year1997        0.0702    0.2294    0.31 0.76031
## Year1998        0.3263    0.2215    1.47 0.14367
## Year1999        0.0175    0.2374    0.07 0.94126
## Year2000       -0.0621    0.2614   -0.24 0.81256
## Year2001        0.2650    0.1880    1.41 0.16158
## Year2002        0.3690    0.2112    1.75 0.08343 .
## Year2003        0.3137    0.2166    1.45 0.15046
## Year2004        0.4729    0.2790    1.70 0.09295 .
## Year2005        0.1791    0.2044    0.88 0.38290
## Year2006        0.5179    0.2499    2.07 0.04056 *
## Year2007        0.4030    0.2248    1.79 0.07584 .
## Year2008        0.1041    0.2732    0.38 0.70393
## Year2009        0.4012    0.2396    1.67 0.09687 .
## Year2010        0.4624    0.2076    2.23 0.02801 *
## Year2011        0.4819    0.1941    2.48 0.01456 *
## Year2012        0.5025    0.1992    2.52 0.01309 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.39
## Multiple R-squared:  0.21, Adjusted R-squared:  0.0857
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 113 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.475  0.873   0.954   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      7.94e-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: 126"
## [1] ""

```

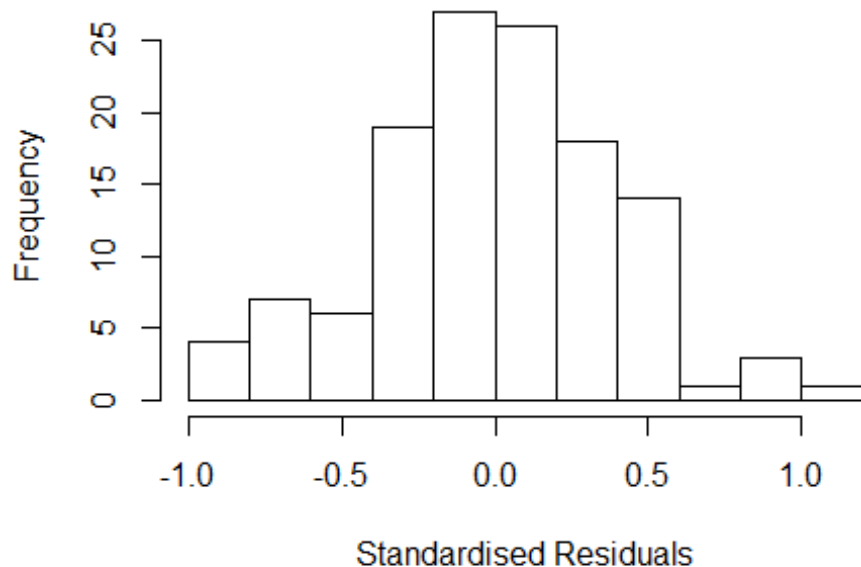
```

## [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
##    7    5    9   10    8    9    8    3    5    3   15   16   12   14   12
## 2011 2012
##   10   15
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    2    4    2    5    4    2    0    1    8    8    4    6    5
## 2011 2012
##    2    7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    2    4    2    5    2    1    0    1    7    8    4    3    4
## 2011 2012
##    2    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: 2.44948974278318"

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

```

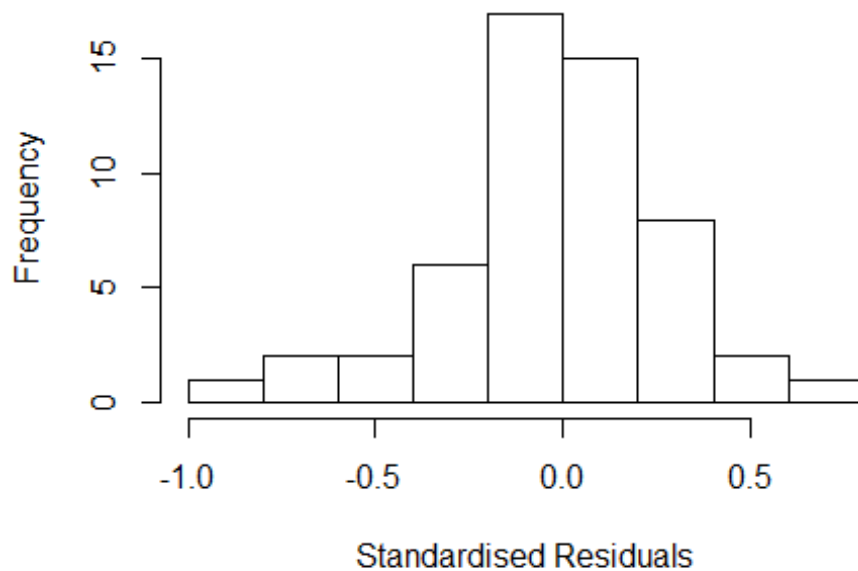
Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3.5, p-value = 0.4
## 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: 3.08007028824102"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
```

| | GVIF | Df | GVIF ^{1/(2*Df)} |
|-------------------|-----------|----|--------------------------|
| FirstAuthorFemale | 5.064e+14 | 1 | 2.250e+07 |
| LastAuthorFemale | 2.282e+01 | 1 | 4.777e+00 |
| UniqueAuthors | 7.610e+31 | 4 | 9.664e+03 |
| Year | 5.001e+33 | 15 | 1.328e+01 |

Residuals from first and last author and team size



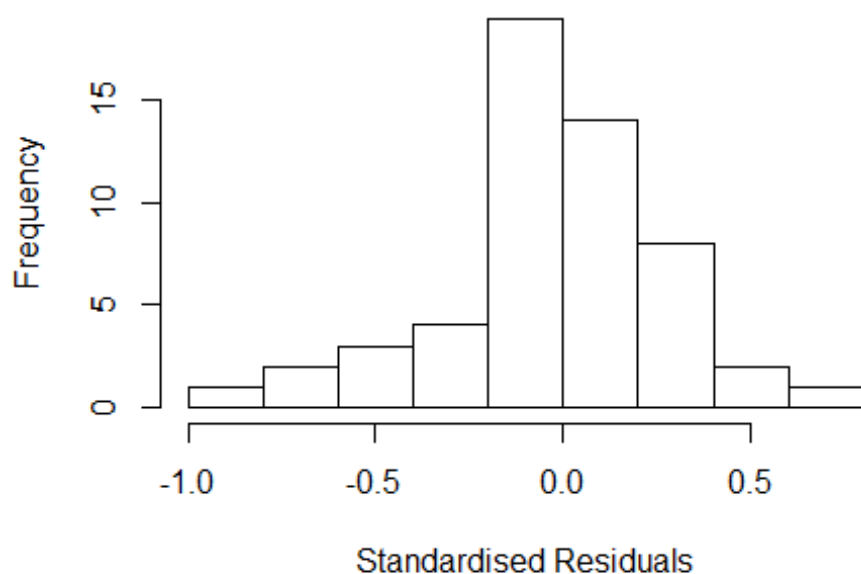
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -8.13e-01 -1.55e-01 2.22e-16 1.69e-01 6.17e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9687 0.1541 6.29 4.8e-07 ***
## FirstAuthorFemale1 -0.0400 0.1431 -0.28 0.7819
## LastAuthorFemale1 -0.2603 0.1519 -1.71 0.0963 .
## UniqueAuthors2 -0.0442 0.1829 -0.24 0.8107
## UniqueAuthors3 -0.1056 0.2282 -0.46 0.6467
## UniqueAuthors4 0.0750 0.4508 0.17 0.8690
## UniqueAuthors5 -0.1784 0.2579 -0.69 0.4941
## Year1997 1.0083 0.1541 6.54 2.3e-07 ***
## Year1998 0.3337 0.1549 2.15 0.0389 *
## Year1999 -0.2435 0.2311 -1.05 0.2999
```

```

## Year2000          0.3226      0.0911      3.54      0.0012 **
## Year2001          0.0530      0.1408      0.38      0.7089
## Year2002          0.4742      0.1400      3.39      0.0019 **
## Year2003         -0.2535      0.1246     -2.03      0.0503 .
## Year2005          0.7707      0.1681      4.58     6.6e-05 ***
## Year2006          0.3933      0.1801      2.18      0.0365 *
## Year2007          0.3239      0.1627      1.99      0.0552 .
## Year2008          0.2467      0.3685      0.67      0.5080
## Year2009          0.0188      0.3131      0.06      0.9525
## Year2010          0.3959      0.1447      2.74      0.0100 *
## Year2011          0.3528      0.2016      1.75      0.0897 .
## Year2012          0.5917      0.2194      2.70      0.0111 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.33
## Multiple R-squared:  0.427, Adjusted R-squared:  0.0517
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 42 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.524  0.887   0.961   0.912   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      1.85e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 7.184e+15  1      8.476e+07
## LastAuthorFemale  1.910e+01  1      4.371e+00
## Year              1.241e+17 15      3.714e+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
## -9.10e-01 -1.39e-01 -9.71e-16 1.63e-01 6.89e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9245 0.1013 9.13 6.7e-11 ***
## FirstAuthorFemale1 -0.0418 0.1326 -0.31 0.75460
## LastAuthorFemale1 -0.2143 0.1172 -1.83 0.07566 .
## Year1997 1.0525 0.1013 10.39 2.2e-12 ***
## Year1998 0.3030 0.1267 2.39 0.02211 *
## Year1999 -0.2199 0.2142 -1.03 0.31161
## Year2000 0.2928 0.0721 4.06 0.00025 ***
## Year2001 0.0656 0.1358 0.48 0.63209
## Year2002 0.4435 0.1136 3.90 0.00040 ***
## Year2003 -0.2535 0.1013 -2.50 0.01698 *
## Year2005 0.6382 0.0763 8.37 5.7e-10 ***
## Year2006 0.3653 0.1578 2.31 0.02643 *
```

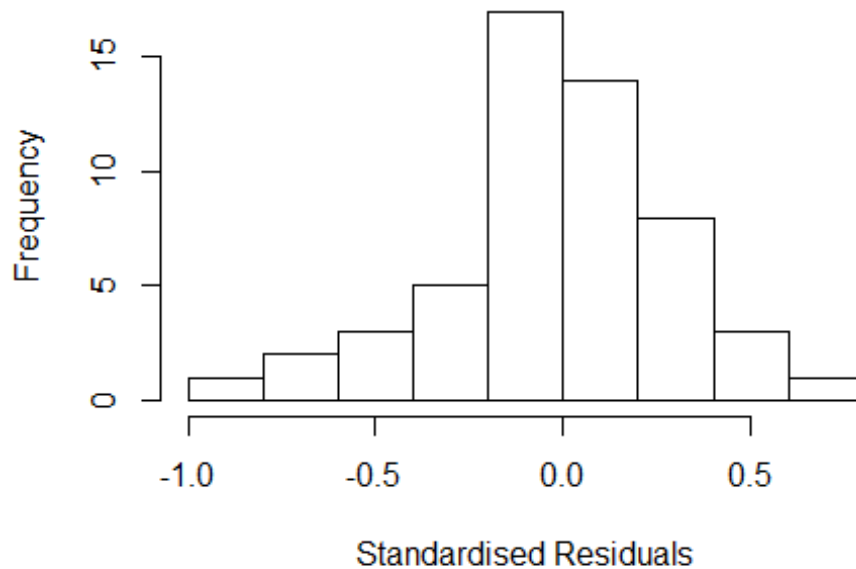


```

## Year2007          0.2597      0.1153      2.25  0.03053 *
## Year2008          0.2821      0.3557      0.79  0.43290
## Year2009          0.0396      0.3435      0.12  0.90890
## Year2010          0.3306      0.1010      3.27  0.00236 **
## Year2011          0.2550      0.1773      1.44  0.15908
## Year2012          0.5328      0.1924      2.77  0.00881 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.326
## Multiple R-squared:  0.412, Adjusted R-squared:  0.134
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 44 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.415  0.892  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      1.85e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -8.043e+14 1      NaN
## Year -8.043e+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
## -9.04e-01 -1.38e-01 -4.16e-17 1.71e-01 7.40e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8163 0.0925 8.83 1.2e-10 ***
## FirstAuthorFemale1 -0.0396 0.1356 -0.29 0.7722
## Year1997 1.1607 0.0925 12.55 6.6e-15 ***
## Year1998 0.4112 0.1197 3.44 0.0015 **
## Year1999 -0.1109 0.2075 -0.53 0.5964
## Year2000 0.4000 0.0737 5.43 3.8e-06 ***
## Year2001 0.1737 0.1300 1.34 0.1898
## Year2002 0.5517 0.1058 5.22 7.2e-06 ***
## Year2003 -0.1453 0.0925 -1.57 0.1246
## Year2005 0.7443 0.0925 8.05 1.2e-09 ***
## Year2006 0.4231 0.1843 2.30 0.0275 *
## Year2007 0.3407 0.1344 2.53 0.0156 *
```

```

## Year2008          0.3850      0.3465      1.11      0.2738
## Year2009          0.1443      0.3379      0.43      0.6718
## Year2010          0.4383      0.0970      4.52      6.2e-05 ***
## Year2011          0.3632      0.1722      2.11      0.0417 *
## Year2012          0.6402      0.1918      3.34      0.0019 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.33
## Multiple R-squared:  0.394, Adjusted R-squared:  0.132
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 45 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.434  0.888  0.964  0.913  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          1.85e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 12825  1          113.248
## Year              12825 15          1.371

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q    Median      3Q      Max
## -9.10e-01 -1.38e-01 -1.11e-16  1.73e-01  7.09e-01
##

```

```

## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.9033    0.0591  15.29 < 2e-16 ***
## LastAuthorFemale1 -0.2135    0.1181  -1.81  0.07878 .
## Year1997          1.0737    0.0591  18.18 < 2e-16 ***
## Year1998          0.3242    0.0963   3.37  0.00178 **
## Year1999         -0.1986    0.1952  -1.02  0.31551
## Year2000          0.2932    0.0639   4.59  5.0e-05 ***
## Year2001          0.0868    0.1090   0.80  0.43070
## Year2002          0.4647    0.0783   5.94  7.6e-07 ***
## Year2003         -0.2323    0.0591  -3.93  0.00036 ***
## Year2005          0.6177    0.0591  10.46  1.3e-12 ***
## Year2006          0.3671    0.1660   2.21  0.03328 *
## Year2007          0.2766    0.0837   3.30  0.00213 **
## Year2008          0.3034    0.3371   0.90  0.37398
## Year2009          0.0438    0.3226   0.14  0.89264
## Year2010          0.3410    0.0930   3.66  0.00077 ***
## Year2011          0.2762    0.1571   1.76  0.08688 .
## Year2012          0.5395    0.2003   2.69  0.01057 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.326
## Multiple R-squared:  0.41, Adjusted R-squared:  0.154
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 44 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.415  0.891  0.968  0.911  0.986  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.85e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500            50          2            1            1000      200
##      trace.lev      mts      compute.rd
##      0              1000      0
##      psi            subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 54"
## [1] ""
## [1] ""
## [1] "#####"

```

```

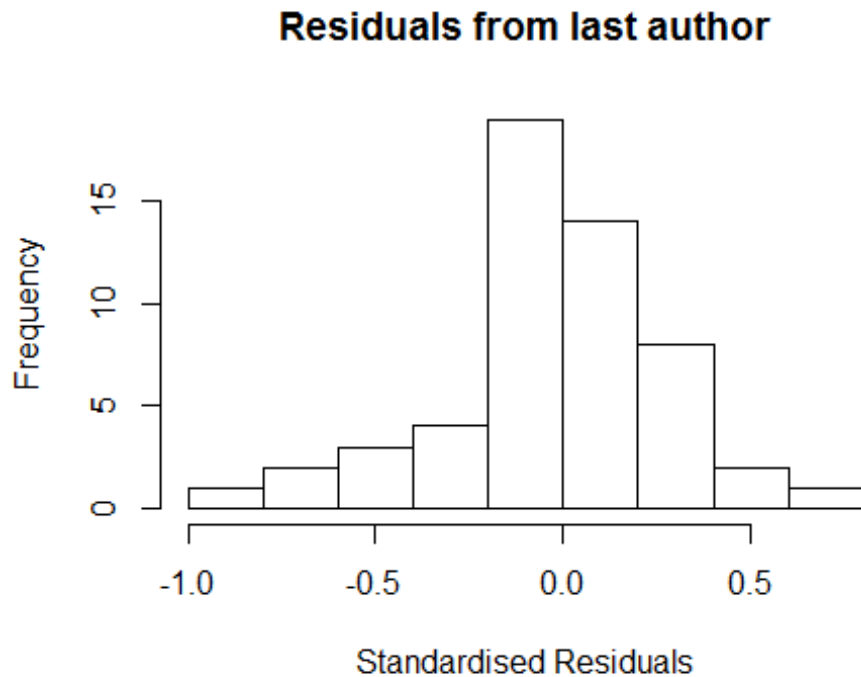
## [1] "Analysis of AJSC 1200"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    1    4    2    6    8    6    7   11    4    3    4   10
##
## 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    0    3    0    4    3    5    5    9    3    1    3    8
##
## 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    0    3    0    4    3    5    5    9    3    1    3    8
## [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"

## Warning in wilcox.test.default(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 = NA
## 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"

## Warning in wilcox.test.default(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 = NA
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -1.313e+15  1      NaN
## LastAuthorFemale  2.010e+15  1    44837113
## UniqueAuthors    -2.547e+29  3      NaN
## Year              -4.203e+44 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 + LastAuthorFemale +
UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
```

```

##      Min      1Q Median      3Q      Max
## -1.563 -0.270  0.000  0.284  1.711
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)  -1.93e-15   2.22e-08   0.00e+00  1.00000
## FirstAuthorFemale1  6.84e-01   4.34e-01   1.58e+00  0.12573
## LastAuthorFemale1 -1.49e+00   6.08e-01  -2.46e+00  0.02011 *
## UniqueAuthors2    1.38e+00   3.22e-01   4.30e+00  0.00017 ***
## UniqueAuthors3    8.28e-01   7.14e-01   1.16e+00  0.25583
## UniqueAuthors4    1.13e+00   2.87e-01   3.96e+00  0.00045 ***
## Year2000         2.58e-15   3.48e-08   0.00e+00  1.00000
## Year2002         3.77e-01   3.77e-01   1.00e+00  0.32548
## Year2004         1.85e+00   1.63e-01   1.14e+01  3.3e-12 ***
## Year2005         1.74e+00   5.96e-01   2.93e+00  0.00659 **
## Year2006         1.30e+00   1.78e-01   7.31e+00  4.8e-08 ***
## Year2007         2.12e+00   2.15e-01   9.86e+00  9.1e-11 ***
## Year2008         1.56e+00   4.08e-01   3.83e+00  0.00063 ***
## Year2009         1.27e+00   3.02e-01   4.19e+00  0.00024 ***
## Year2010         1.39e+00   3.74e-08   3.71e+07  < 2e-16 ***
## Year2011         1.05e+00   2.47e-01   4.24e+00  0.00021 ***
## Year2012         9.62e-01   5.03e-01   1.91e+00  0.06552 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.68
## Multiple R-squared:  0.6,   Adjusted R-squared:  0.38
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 36 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.506  0.906  0.973  0.917  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      2.17e-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)

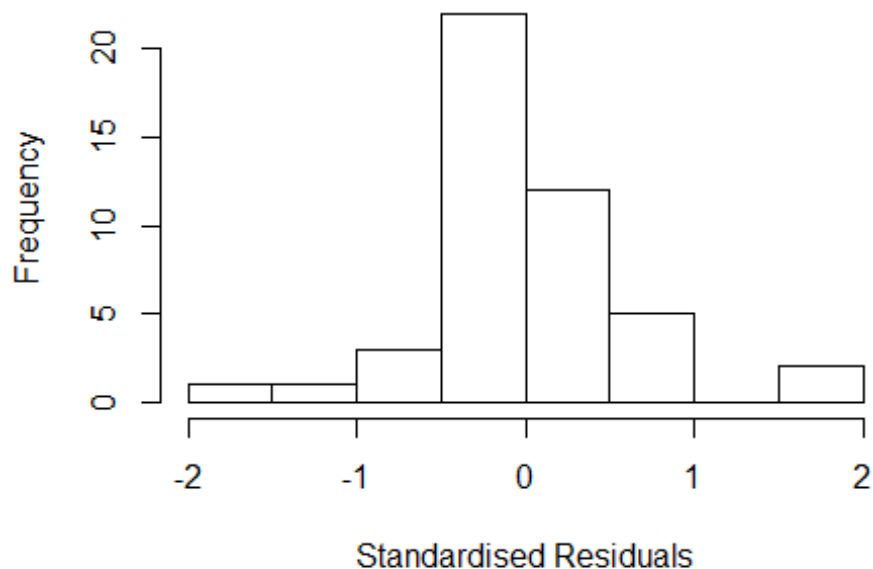
```

```
## [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 cov2cor(v): 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 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 + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.56428 -0.38859 -0.00914  0.48269  1.44839
##
```



```

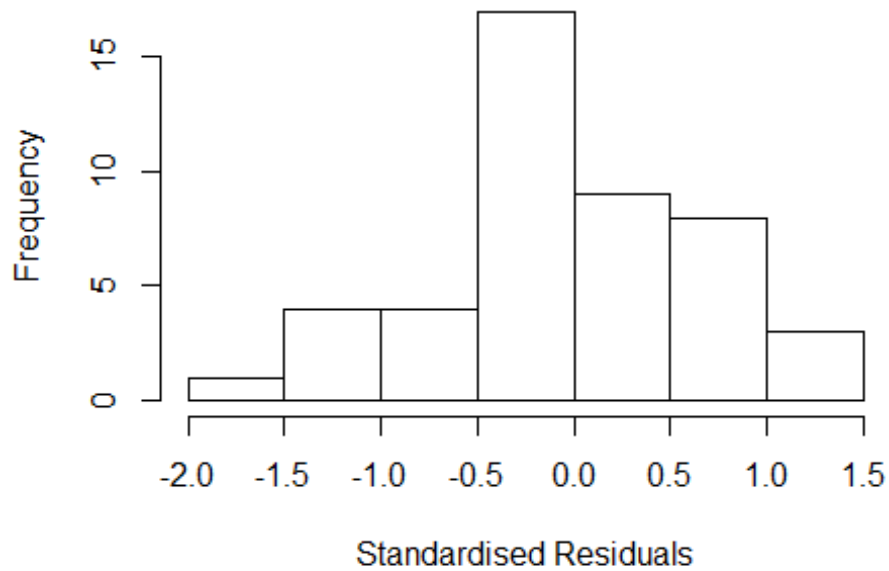
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    8.09e-16  0.00e+00   Inf    < 2e-16 ***
## FirstAuthorFemale1 4.35e-01  8.52e-01   0.51  0.61310
## LastAuthorFemale1 -9.85e-01  8.35e-01  -1.18  0.24709
## Year2000        -2.19e-16  0.00e+00  -Inf    < 2e-16 ***
## Year2002         3.89e-01  3.62e-01   1.07  0.29090
## Year2004         1.79e+00  1.93e-01   9.27  1.4e-10 ***
## Year2005         1.55e+00  6.63e-01   2.34  0.02553 *
## Year2006         1.25e+00  2.06e-01   6.10  8.2e-07 ***
## Year2007         2.52e+00  3.41e-01   7.38  2.1e-08 ***
## Year2008         1.56e+00  3.64e-01   4.30  0.00015 ***
## Year2009         1.27e+00  2.94e-01   4.32  0.00014 ***
## Year2010         1.39e+00  0.00e+00   Inf    < 2e-16 ***
## Year2011         8.77e-01  2.41e-01   3.64  0.00096 ***
## Year2012         1.17e+00  4.68e-01   2.50  0.01795 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.821
## Multiple R-squared:  0.441, Adjusted R-squared:  0.214
## 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.
##  0.697  0.903  0.966  0.933  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"

## 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

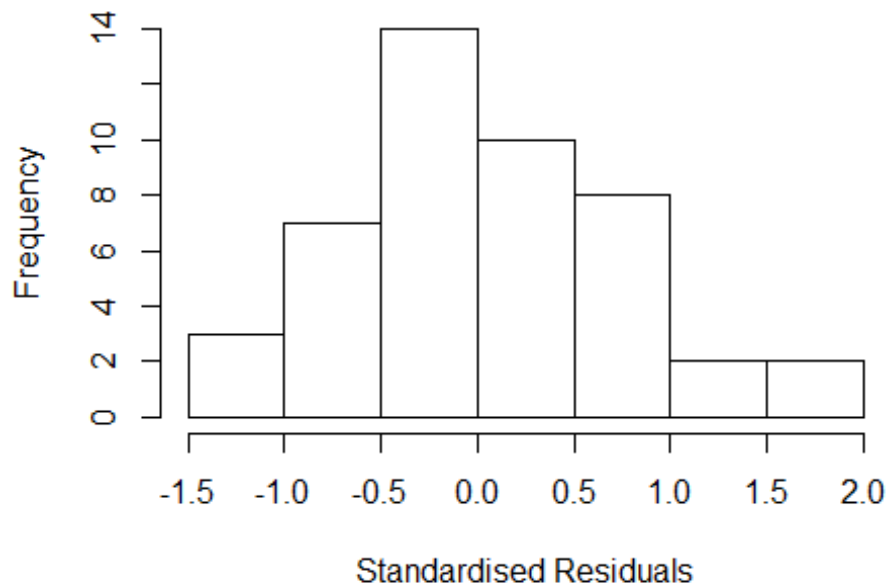
```

Residuals from first and last author



| | | | | |
|----|-------------------|------|----|---------------------|
| ## | | GVIF | Df | $GVIF^{(1/(2*Df))}$ |
| ## | FirstAuthorFemale | NaN | 1 | NaN |
| ## | Year | NaN | 11 | NaN |

Residuals from first author



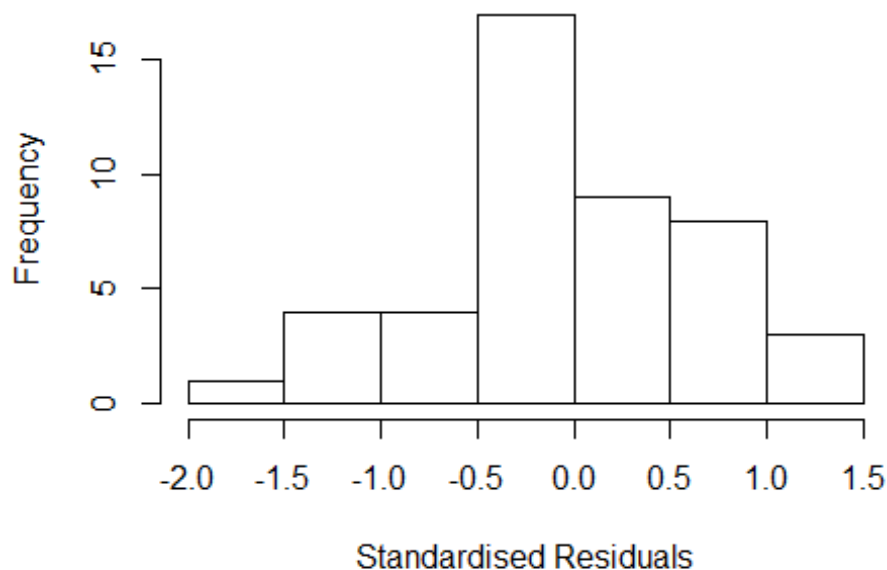
```

## [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
## -1.40e+00 -3.88e-01 3.16e-16 4.63e-01 1.77e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 3.38e-17 2.12e-08 0.00 1.00000
## FirstAuthorFemale1 -4.71e-01 3.12e-01 -1.51 0.14055
## Year2000 2.40e-16 0.00e+00 Inf < 2e-16 ***
## Year2002 3.88e-01 3.64e-01 1.07 0.29410
## Year2004 1.77e+00 2.06e-01 8.60 6.1e-10 ***
## Year2005 1.49e+00 7.08e-01 2.11 0.04282 *
## Year2006 1.24e+00 2.18e-01 5.70 2.3e-06 ***
## Year2007 2.49e+00 3.26e-01 7.62 8.9e-09 ***
## Year2008 1.40e+00 2.88e-01 4.86 2.8e-05 ***
## Year2009 1.27e+00 2.95e-01 4.30 0.00014 ***
## Year2010 1.39e+00 0.00e+00 Inf < 2e-16 ***
## Year2011 8.24e-01 2.48e-01 3.33 0.00215 **
## Year2012 1.17e+00 5.51e-01 2.12 0.04132 *
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.807
## Multiple R-squared: 0.416, Adjusted R-squared: 0.203
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~ = 1. The remaining 40 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.608 0.915 0.963 0.928 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 2.17e-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 4.925 1          2.219
## Year              4.925 11         1.075
```

Residuals from last author



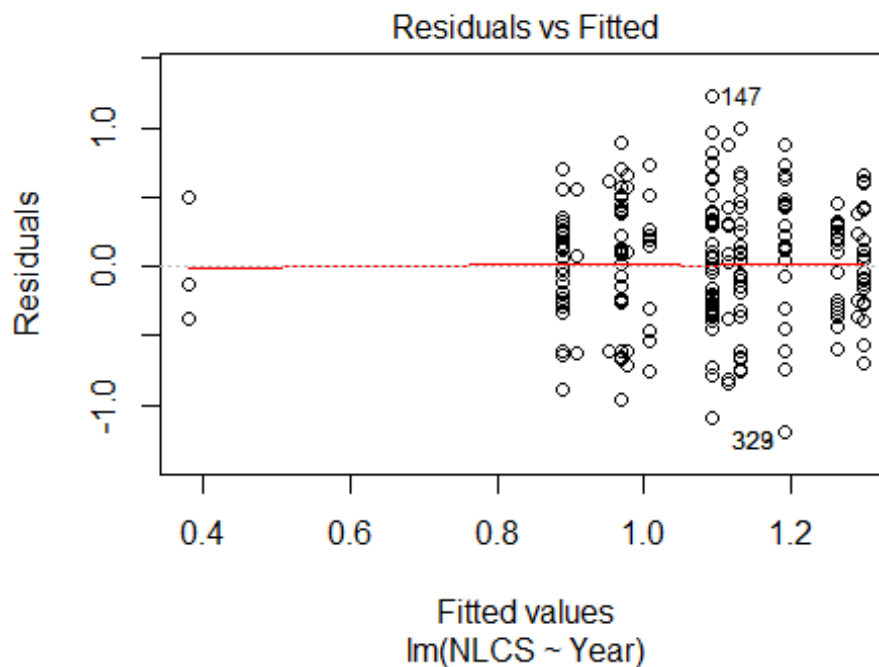
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.54e+00 -3.87e-01 -2.27e-15  4.74e-01  1.52e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   3.75e-15   2.29e-08  0.00e+00  1.00000
## LastAuthorFemale1 -5.99e-01   3.09e-01 -1.94e+00  0.06076 .
```

```

## Year2000      -2.97e-15   4.16e-08   0.00e+00   1.00000
## Year2002      3.87e-01   3.64e-01   1.06e+00   0.29647
## Year2004      1.80e+00   1.87e-01   9.65e+00   3.9e-11 ***
## Year2005      1.59e+00   6.50e-01   2.44e+00   0.02000 *
## Year2006      1.26e+00   2.00e-01   6.31e+00   3.9e-07 ***
## Year2007      2.54e+00   3.56e-01   7.13e+00   3.6e-08 ***
## Year2008      1.54e+00   3.58e-01   4.30e+00   0.00014 ***
## Year2009      1.27e+00   2.95e-01   4.29e+00   0.00015 ***
## Year2010      1.39e+00   2.29e-08   6.06e+07   < 2e-16 ***
## Year2011      9.09e-01   2.40e-01   3.79e+00   0.00061 ***
## Year2012      1.24e+00   5.03e-01   2.46e+00   0.01929 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.789
## Multiple R-squared:  0.438, Adjusted R-squared:  0.234
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.682  0.905   0.964   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      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 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
##    1    3    7    4   12    6    8    7    6   12   19   22   37   26   30
## 2011 2012

```

```
## 33 31
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 0 3 4 3 9 2 7 7 5 10 17 20 36 20 27
## 2011 2012
## 28 23
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 0 3 3 3 8 2 6 7 4 10 17 15 34 19 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 = 19, df = 15, p-value = 0.2
```



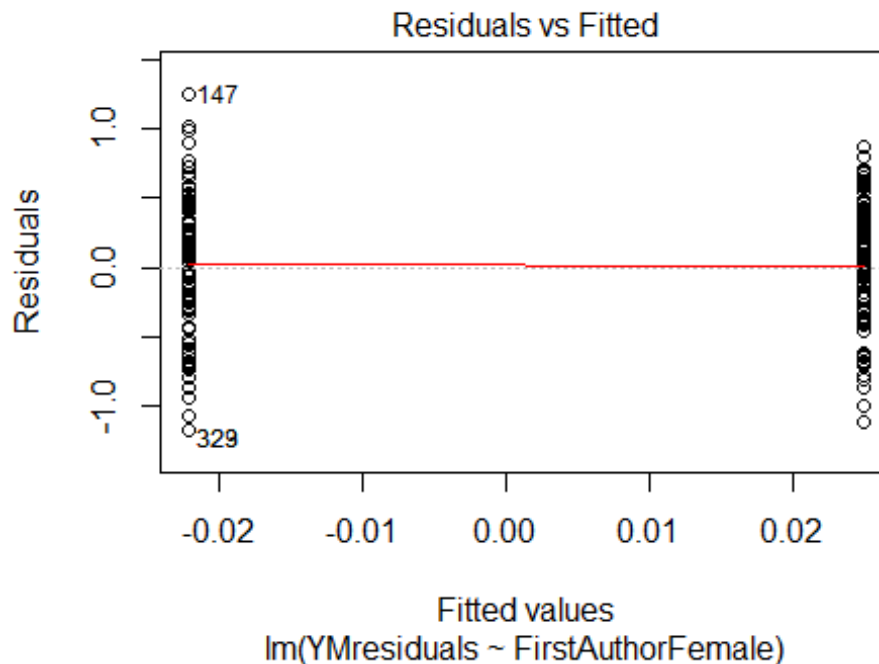
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.2, df = 1, p-value = 0.1

## [1] "Female first author team size 2018 geometric mean: 2.21587509423541"
## [1] "Male first author team size 2018 geometric mean: 1.43096908110526"
```

```
## Warning in wilcox.test.default(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 = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.20294266123408"
## [1] "Male last author team size 2018 geometric mean: 1.44612554959192"

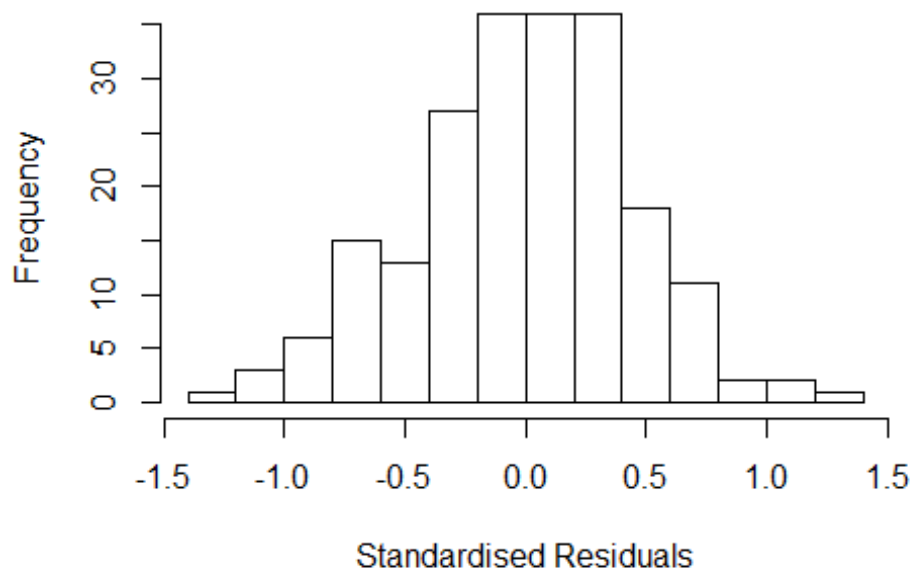
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## 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 1.899 1 1.378
```

| | | | |
|---------------------|-------|----|-------|
| ## LastAuthorFemale | 2.150 | 1 | 1.466 |
| ## UniqueAuthors | 4.400 | 4 | 1.203 |
| ## Year | 5.311 | 15 | 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
## -1.2292 -0.2557 0.0232 0.2956 1.2222
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.29193 0.18752 1.56 0.12123
## FirstAuthorFemale1 0.04822 0.08376 0.58 0.56549
## LastAuthorFemale1 0.00452 0.08627 0.05 0.95830
## UniqueAuthors2 0.20026 0.09363 2.14 0.03376 *
## UniqueAuthors3 0.15790 0.16300 0.97 0.33396
## UniqueAuthors4 0.41270 0.09160 4.51 1.2e-05 ***
```

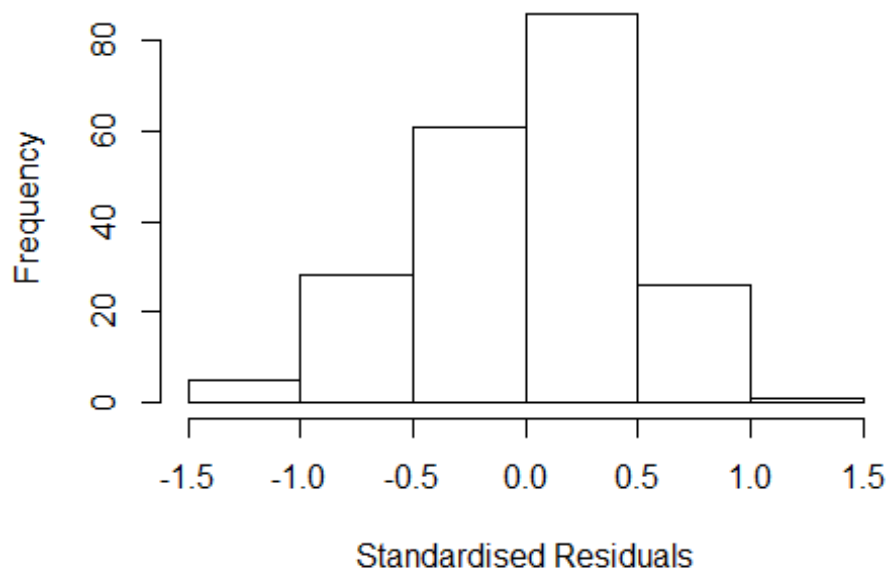


```

## UniqueAuthors5      0.48799      0.12645      3.86 0.00016 ***
## Year1998             1.07635      0.28841      3.73 0.00025 ***
## Year1999             0.59381      0.40765      1.46 0.14690
## Year2000             0.64332      0.26411      2.44 0.01581 *
## Year2001             0.63470      0.67455      0.94 0.34797
## Year2002             0.47662      0.21552      2.21 0.02823 *
## Year2003             0.73184      0.22465      3.26 0.00134 **
## Year2004             0.76959      0.34932      2.20 0.02882 *
## Year2005             0.59358      0.23737      2.50 0.01327 *
## Year2006             0.50456      0.20528      2.46 0.01490 *
## Year2007             0.82466      0.21656      3.81 0.00019 ***
## Year2008             0.59707      0.21457      2.78 0.00595 **
## Year2009             0.79691      0.20534      3.88 0.00014 ***
## Year2010             0.73177      0.21260      3.44 0.00071 ***
## Year2011             0.51749      0.21982      2.35 0.01962 *
## Year2012             0.77941      0.23076      3.38 0.00089 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.444
## Multiple R-squared:  0.193, Adjusted R-squared:  0.102
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 193 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.423  0.873  0.957  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      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 2.185 1          1.478
## LastAuthorFemale  1.925 1          1.387
## Year              1.384 15          1.011

```

Residuals from first and last author



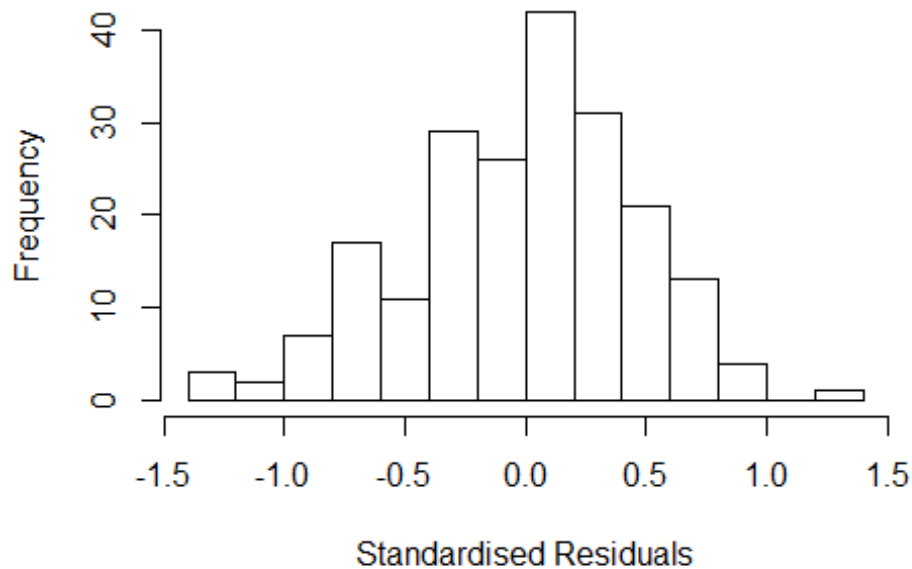
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2747 -0.2976 0.0584 0.3230 1.3003
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.3661 0.2360 1.55 0.12252
## FirstAuthorFemale1 0.0793 0.0933 0.85 0.39667
## LastAuthorFemale1 -0.0722 0.0907 -0.80 0.42751
## Year1998 1.0127 0.3100 3.27 0.00129 **
## Year1999 0.5494 0.4070 1.35 0.17859
## Year2000 0.6502 0.3055 2.13 0.03461 *
## Year2001 0.5834 0.6807 0.86 0.39252
## Year2002 0.6539 0.2620 2.50 0.01342 *
## Year2003 0.7248 0.2645 2.74 0.00673 **
## Year2004 0.8209 0.3882 2.11 0.03576 *
## Year2005 0.6339 0.2870 2.21 0.02837 *
## Year2006 0.5289 0.2494 2.12 0.03526 *
```

```

## Year2007          0.9026      0.2614      3.45  0.00068 ***
## Year2008          0.7218      0.2563      2.82  0.00538 **
## Year2009          0.9037      0.2464      3.67  0.00032 ***
## Year2010          0.7854      0.2533      3.10  0.00222 **
## Year2011          0.5967      0.2595      2.30  0.02257 *
## Year2012          0.9086      0.2772      3.28  0.00125 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.483
## Multiple R-squared:  0.106, Adjusted R-squared:  0.026
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 197 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.449  0.879  0.958  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      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 1.248 1          1.117
## Year              1.248 15          1.007

```

Residuals from first author



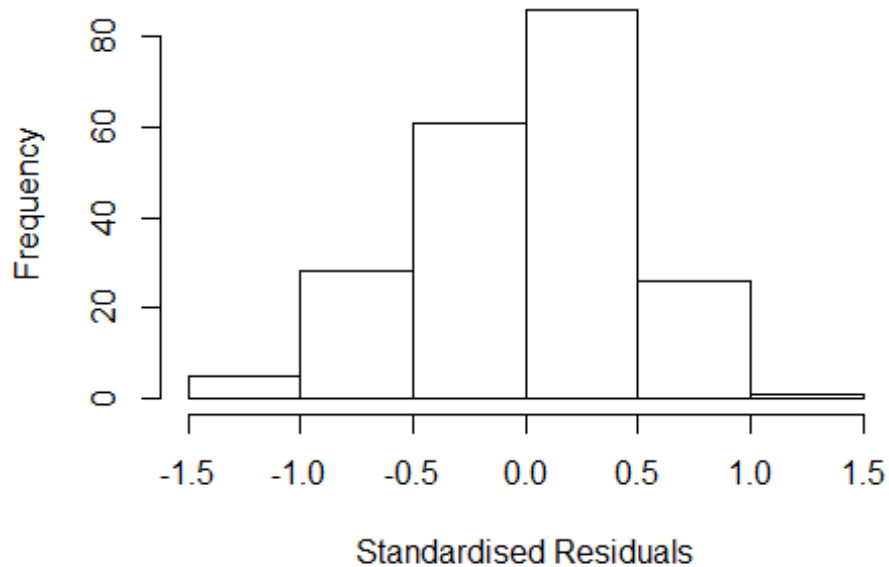
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2624 -0.3180 0.0452 0.3255 1.2401
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.3562 0.2439 1.46 0.14580
## FirstAuthorFemale1 0.0360 0.0716 0.50 0.61578
## Year1998 1.0146 0.3226 3.15 0.00193 **
## Year1999 0.5393 0.4214 1.28 0.20213
## Year2000 0.6497 0.3133 2.07 0.03943 *
## Year2001 0.5788 0.6630 0.87 0.38381
## Year2002 0.6570 0.2722 2.41 0.01676 *
## Year2003 0.7223 0.2726 2.65 0.00874 **
## Year2004 0.8279 0.3787 2.19 0.03002 *
## Year2005 0.6446 0.2935 2.20 0.02930 *
## Year2006 0.5262 0.2570 2.05 0.04200 *
## Year2007 0.9046 0.2710 3.34 0.00102 **
```

```

## Year2008          0.7196      0.2644      2.72  0.00709 **
## Year2009          0.9074      0.2547      3.56  0.00047 ***
## Year2010          0.7779      0.2601      2.99  0.00315 **
## Year2011          0.5985      0.2673      2.24  0.02630 *
## Year2012          0.9062      0.2830      3.20  0.00160 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.477
## Multiple R-squared:  0.103, Adjusted R-squared:  0.0276
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 192 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.464  0.862  0.952  0.911  0.986  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      4.83e-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 1.115 1          1.056
## Year            1.115 15          1.004

```

Residuals from last author



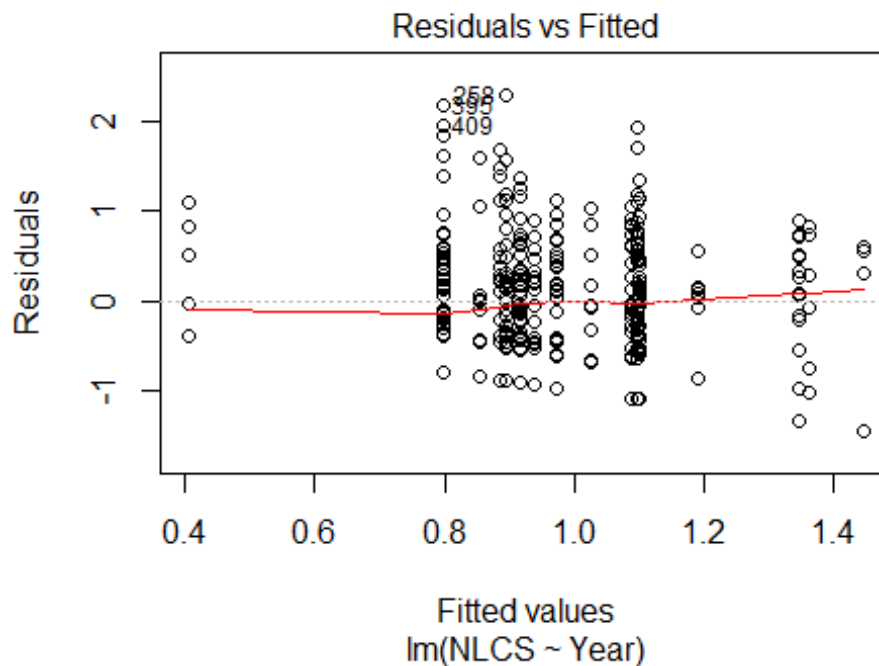
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2915 -0.3175 0.0565 0.3112 1.2322
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.3770 0.2279 1.65 0.09968 .
## LastAuthorFemale1 -0.0261 0.0688 -0.38 0.70515
## Year1998 1.0115 0.2970 3.41 0.00080 ***
## Year1999 0.5619 0.3917 1.43 0.15303
## Year2000 0.6735 0.2985 2.26 0.02519 *
## Year2001 0.5890 0.7173 0.82 0.41256
## Year2002 0.6734 0.2535 2.66 0.00856 **
## Year2003 0.7283 0.2566 2.84 0.00503 **
## Year2004 0.8259 0.3687 2.24 0.02625 *
## Year2005 0.6453 0.2819 2.29 0.02320 *
## Year2006 0.5248 0.2427 2.16 0.03186 *
## Year2007 0.9030 0.2554 3.53 0.00051 ***
```

```

## Year2008          0.7329      0.2486      2.95  0.00360 **
## Year2009          0.9212      0.2371      3.89  0.00014 ***
## Year2010          0.7897      0.2449      3.22  0.00148 **
## Year2011          0.6093      0.2527      2.41  0.01686 *
## Year2012          0.9145      0.2718      3.36  0.00093 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.48
## Multiple R-squared:  0.103, Adjusted R-squared:  0.0277
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 196 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.449  0.867  0.957   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      4.83e-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 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
##    7    9   11   12   11   17   16   34   21   29   23   31   46   35   51
## 2011 2012
##   69   59
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    6   10   10    7   11   13   25   17   25   18   27   40   31   44
## 2011 2012
##   57   48

```

```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    6   10   10    7   11   13   25   17   24   18   27   39   31   42
## 2011 2012
##   55   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 = 14, df = 16, p-value = 0.6
```

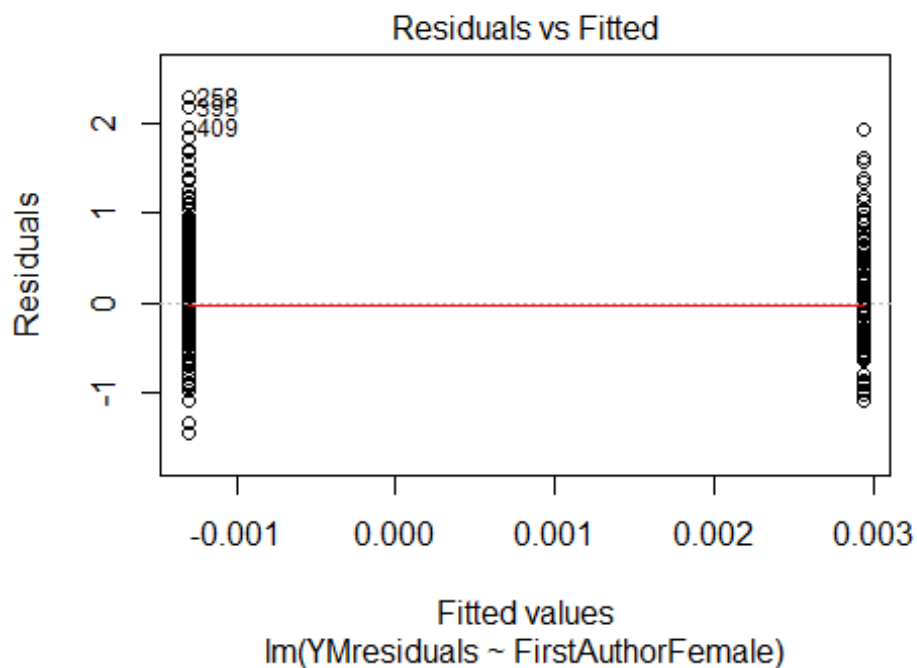


```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.11, df = 1, p-value = 0.7
## [1] "Female first author team size 2018 geometric mean: 1.02811382665607"
## [1] "Male first author team size 2018 geometric mean: 1.03715504444619"
## Warning in wilcox.test.default(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 = 0.8
## 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: 1.05476607648165"

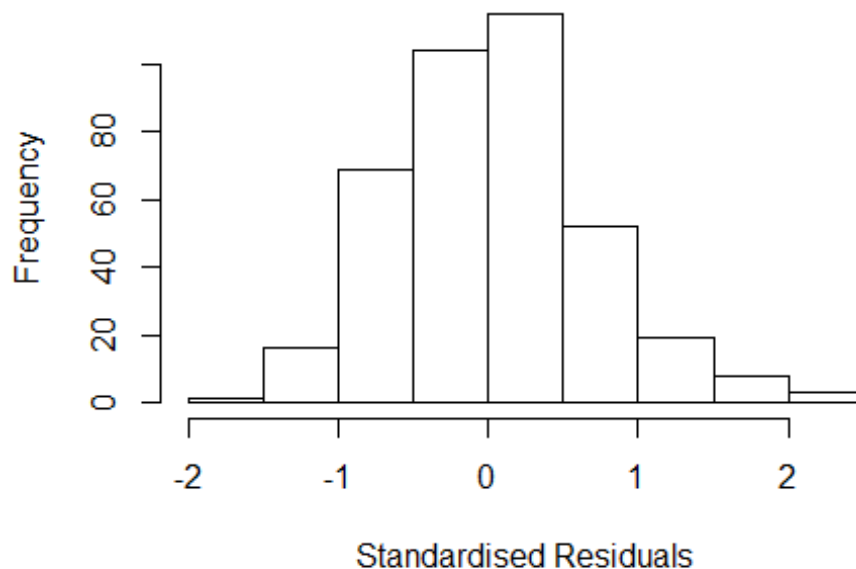
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 430, 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"
##
```

| | GVI | F | Df | GVI ^{1/(2*Df)} |
|-------------------|--------|----|----|-------------------------|
| FirstAuthorFemale | 16.377 | 1 | 1 | 4.047 |
| LastAuthorFemale | 7.257 | 1 | 1 | 2.694 |
| UniqueAuthors | 45.852 | 3 | 3 | 1.892 |
| Year | 19.166 | 16 | 16 | 1.097 |

Residuals from first and last author and team size



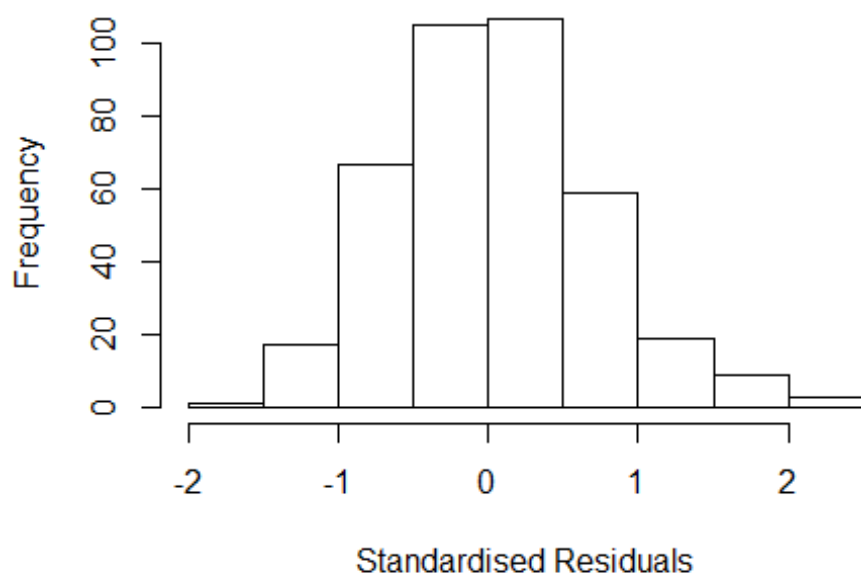
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.6419 -0.4087 0.0207 0.4521 2.4298
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.642 0.401 4.10 5.2e-05 ***
## FirstAuthorFemale1 -0.097 0.159 -0.61 0.54229
## LastAuthorFemale1 0.147 0.160 0.92 0.35999
## UniqueAuthors2 0.547 0.148 3.69 0.00026 ***
## UniqueAuthors3 0.144 0.276 0.52 0.60135
## UniqueAuthors4 0.686 0.193 3.55 0.00044 ***
## Year1997 -0.324 0.506 -0.64 0.52258
## Year1998 -0.891 0.465 -1.92 0.05603 .
## Year1999 -1.349 0.436 -3.10 0.00212 **
## Year2000 -0.462 0.426 -1.08 0.27918
```

```

## Year2001          -0.715      0.451   -1.58   0.11422
## Year2002          -0.748      0.432   -1.73   0.08454 .
## Year2003          -0.846      0.447   -1.89   0.05903 .
## Year2004          -0.266      0.442   -0.60   0.54821
## Year2005          -0.683      0.416   -1.64   0.10161
## Year2006          -0.757      0.423   -1.79   0.07443 .
## Year2007          -0.581      0.422   -1.38   0.16926
## Year2008          -0.884      0.416   -2.12   0.03453 *
## Year2009          -0.759      0.419   -1.81   0.07058 .
## Year2010          -0.857      0.411   -2.08   0.03805 *
## Year2011          -1.026      0.421   -2.44   0.01522 *
## Year2012          -0.614      0.418   -1.47   0.14312
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.657
## Multiple R-squared:  0.135, Adjusted R-squared:  0.0853
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 31 weights are ~= 1. The remaining 356 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.142  0.880  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      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 8.058 1      2.839
## LastAuthorFemale  8.068 1      2.840
## Year              1.170 16      1.005

```

Residuals from first and last author



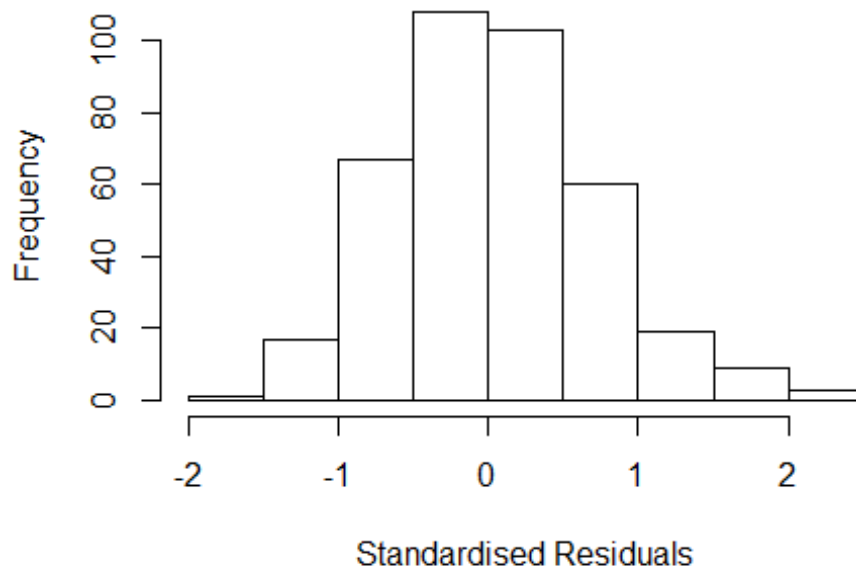
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.6332 -0.4336 0.0061 0.4518 2.4226
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.633 0.406 4.02 6.9e-05 ***
## FirstAuthorFemale1 -0.249 0.209 -1.19 0.2345
## LastAuthorFemale1 0.307 0.206 1.49 0.1374
## Year1997 -0.322 0.504 -0.64 0.5227
## Year1998 -0.879 0.469 -1.87 0.0617 .
## Year1999 -1.279 0.447 -2.86 0.0045 **
## Year2000 -0.458 0.431 -1.06 0.2888
## Year2001 -0.669 0.450 -1.49 0.1380
## Year2002 -0.739 0.437 -1.69 0.0916 .
## Year2003 -0.836 0.450 -1.86 0.0641 .
## Year2004 -0.239 0.441 -0.54 0.5883
## Year2005 -0.632 0.422 -1.50 0.1351
```

```

## Year2006          -0.727      0.427   -1.70   0.0893 .
## Year2007          -0.553      0.430   -1.29   0.1986
## Year2008          -0.868      0.422   -2.05   0.0406 *
## Year2009          -0.657      0.422   -1.56   0.1208
## Year2010          -0.769      0.417   -1.85   0.0656 .
## Year2011          -0.980      0.426   -2.30   0.0219 *
## Year2012          -0.597      0.425   -1.41   0.1607
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.669
## Multiple R-squared:  0.0998, Adjusted R-squared:  0.0557
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 37 weights are ~= 1. The remaining 350 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##   0.162  0.875  0.949  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      2.58e-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 1.07 1      1.034
## Year              1.07 16      1.002

```

Residuals from first author



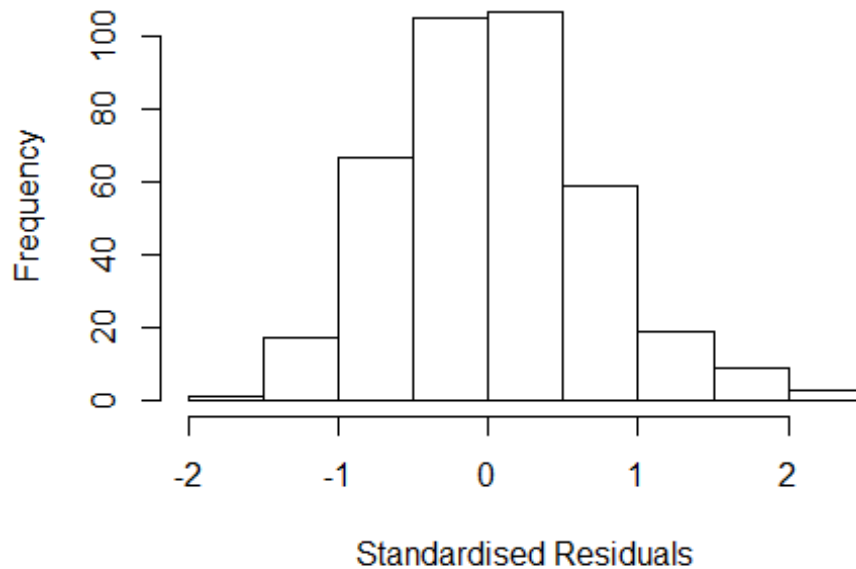
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.6344 -0.4355 0.0038 0.4562 2.4122
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6344 0.4052 4.03 6.7e-05 ***
## FirstAuthorFemale1 0.0349 0.0760 0.46 0.6466
## Year1997 -0.2641 0.5261 -0.50 0.6160
## Year1998 -0.8780 0.4690 -1.87 0.0620 .
## Year1999 -1.2759 0.4474 -2.85 0.0046 **
## Year2000 -0.4471 0.4306 -1.04 0.2998
## Year2001 -0.6335 0.4466 -1.42 0.1570
## Year2002 -0.7386 0.4367 -1.69 0.0916 .
## Year2003 -0.8306 0.4501 -1.85 0.0658 .
## Year2004 -0.2352 0.4400 -0.53 0.5934
## Year2005 -0.6097 0.4218 -1.45 0.1492
## Year2006 -0.7004 0.4257 -1.65 0.1007
```

```

## Year2007          -0.5539      0.4277   -1.30    0.1961
## Year2008          -0.8585      0.4226   -2.03    0.0429 *
## Year2009          -0.6605      0.4212   -1.57    0.1177
## Year2010          -0.7632      0.4170   -1.83    0.0681 .
## Year2011          -0.9620      0.4251   -2.26    0.0242 *
## Year2012          -0.5894      0.4245   -1.39    0.1659
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.667
## Multiple R-squared:  0.0938, Adjusted R-squared:  0.052
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 32 weights are ~= 1. The remaining 355 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.163  0.872  0.951  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] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.079 1          1.039
## Year            1.079 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
## -1.6351 -0.4374 0.0106 0.4505 2.4286
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6351 0.4048 4.04 6.5e-05 ***
## LastAuthorFemale1 0.0801 0.0762 1.05 0.2938
## Year1997 -0.2975 0.5177 -0.57 0.5659
## Year1998 -0.8854 0.4680 -1.89 0.0593 .
## Year1999 -1.2847 0.4449 -2.89 0.0041 **
## Year2000 -0.4719 0.4304 -1.10 0.2736
## Year2001 -0.6504 0.4465 -1.46 0.1461
## Year2002 -0.7437 0.4361 -1.71 0.0890 .
## Year2003 -0.8464 0.4495 -1.88 0.0605 .
## Year2004 -0.2446 0.4401 -0.56 0.5786
## Year2005 -0.6317 0.4212 -1.50 0.1345
## Year2006 -0.7218 0.4257 -1.70 0.0908 .
```

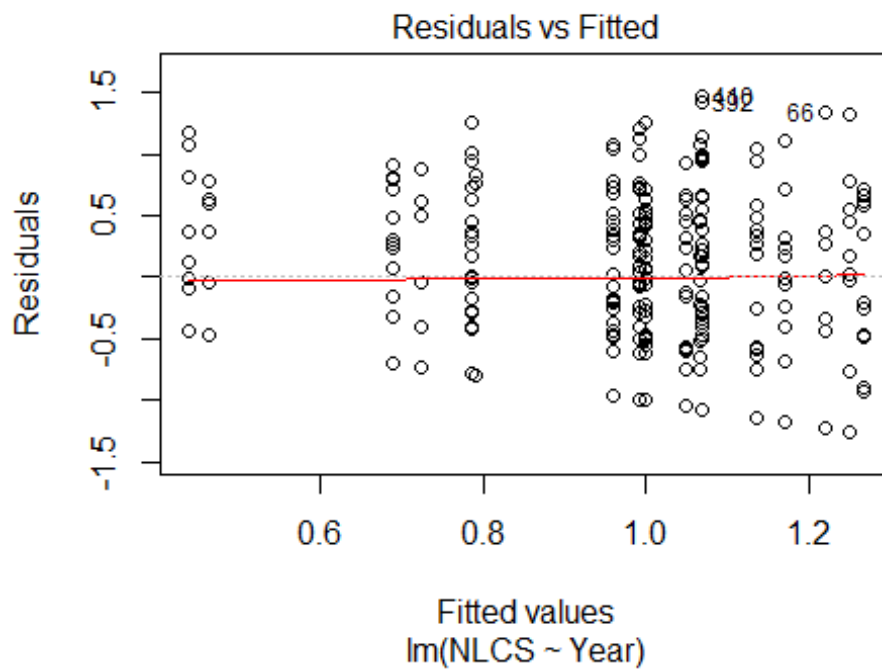


```

## Year2007          -0.5678      0.4277   -1.33   0.1851
## Year2008          -0.8758      0.4217   -2.08   0.0385 *
## Year2009          -0.6713      0.4210   -1.59   0.1116
## Year2010          -0.7757      0.4159   -1.87   0.0629 .
## Year2011          -0.9828      0.4251   -2.31   0.0213 *
## Year2012          -0.6088      0.4240   -1.44   0.1519
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.666
## Multiple R-squared:  0.0964, Adjusted R-squared:  0.0547
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 34 weights are ~= 1. The remaining 353 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.155  0.876  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      2.58e-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: 387"
## [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
##    8   28   11   11    7   15   12   17   17   22   18   35   33   31   22
## 2011 2012
##   43   36
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4   18    9   10    7    9   11   10   13   17   15   29   27   27   17
## 2011 2012

```

```
## 34 35
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4 18 8 10 7 8 9 9 12 15 15 29 26 27 16
## 2011 2012
## 34 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 = 9.2, df = 16, p-value = 0.9
```



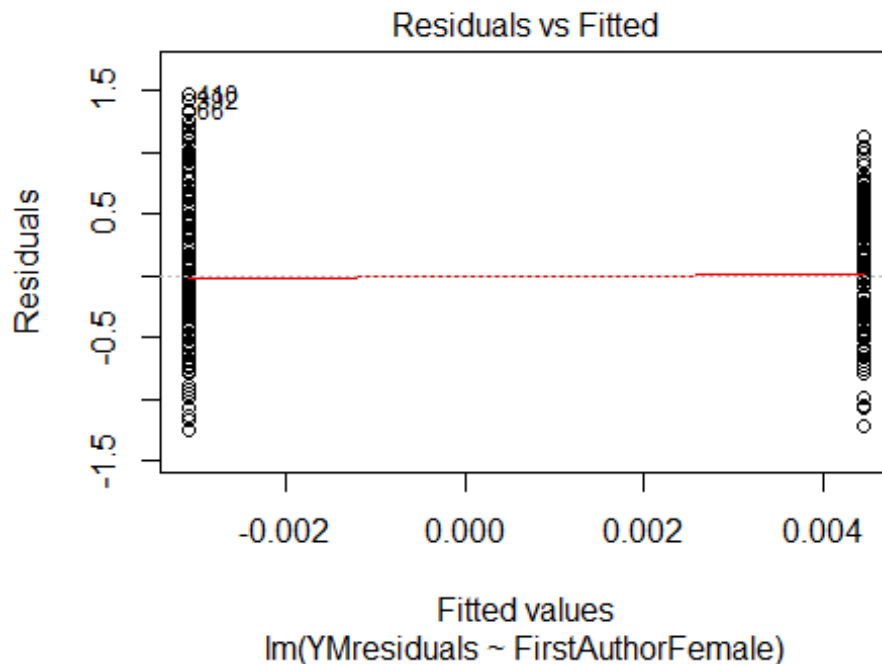
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 4.1, df = 1, p-value = 0.04

## [1] "Female first author team size 2018 geometric mean: 1.61133163576468"
## [1] "Male first author team size 2018 geometric mean: 1.529819374737"

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

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 80, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.69094368754578"
## [1] "Male last author team size 2018 geometric mean: 1.36426160182137"

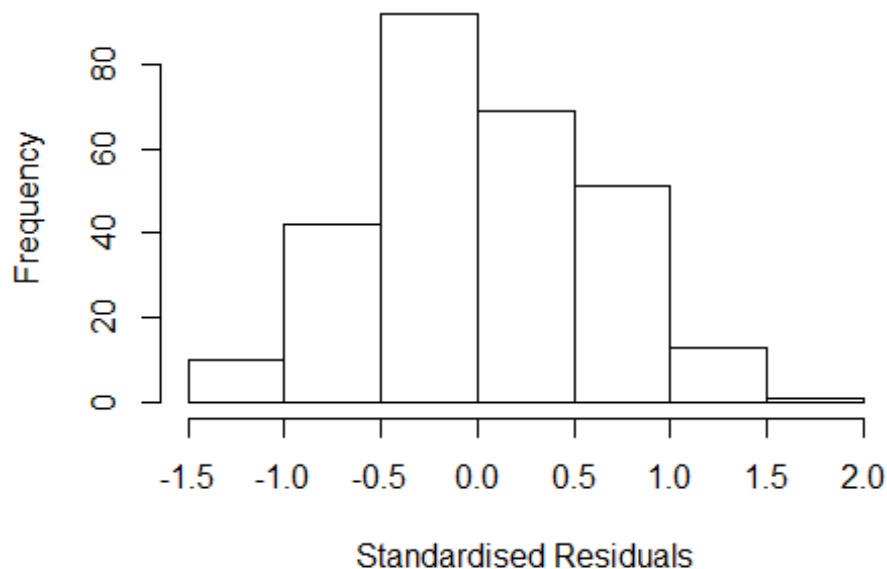
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 86, p-value = 0.6
## 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 | 5.637 | 1 | 2.374 |
| LastAuthorFemale | 5.532 | 1 | 2.352 |
| UniqueAuthors | 3.453 | 4 | 1.168 |
| Year | 3.679 | 16 | 1.042 |

Residuals from first and last author and team size



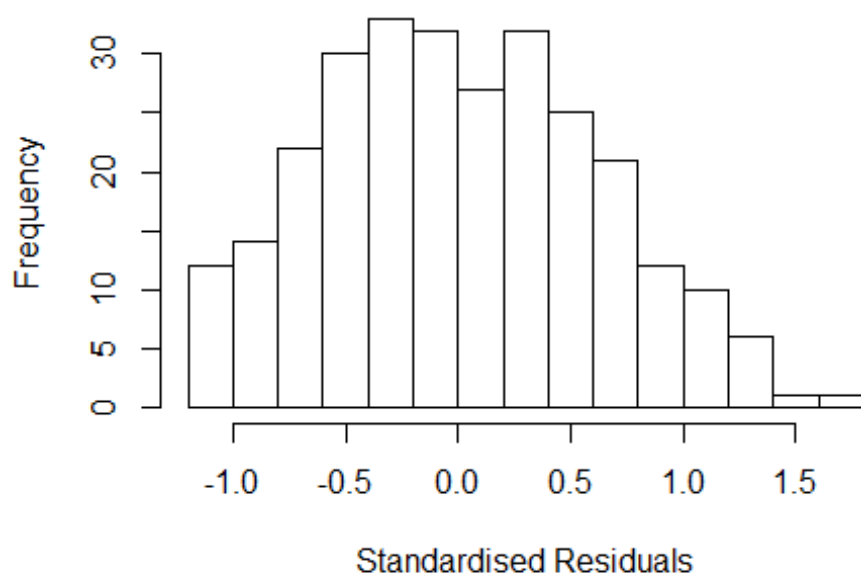
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3094 -0.4205 -0.0262 0.4430 1.6984
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.76491 0.53756 1.42 0.156
## FirstAuthorFemale1 -0.20544 0.17109 -1.20 0.231
## LastAuthorFemale1 0.30504 0.17129 1.78 0.076 .
## UniqueAuthors2 0.26056 0.10432 2.50 0.013 *
## UniqueAuthors3 0.36645 0.20968 1.75 0.082 .
## UniqueAuthors4 0.43450 0.34086 1.27 0.204
## UniqueAuthors5 -0.00906 0.25096 -0.04 0.971
## Year1997 -0.39694 0.55122 -0.72 0.472
## Year1998 0.08540 0.56961 0.15 0.881
## Year1999 -0.34964 0.55855 -0.63 0.532
```

```

## Year2000      0.31544      0.60366      0.52      0.602
## Year2001     -0.22401      0.58746     -0.38      0.703
## Year2002      0.36220      0.59036      0.61      0.540
## Year2003      0.31398      0.59629      0.53      0.599
## Year2004      0.39285      0.56533      0.69      0.488
## Year2005      0.35034      0.58107      0.60      0.547
## Year2006      0.20761      0.55830      0.37      0.710
## Year2007      0.10033      0.54788      0.18      0.855
## Year2008     -0.13436      0.54747     -0.25      0.806
## Year2009      0.05278      0.55156      0.10      0.924
## Year2010     -0.17707      0.55601     -0.32      0.750
## Year2011      0.07299      0.54920      0.13      0.894
## Year2012      0.07066      0.55280      0.13      0.898
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.637
## Multiple R-squared:  0.149, Adjusted R-squared:  0.0756
## 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.
##  0.457  0.873  0.954  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          3.60e-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 4.640 1      2.154
## LastAuthorFemale  4.747 1      2.179
## Year              1.469 16      1.012

```

Residuals from first and last author



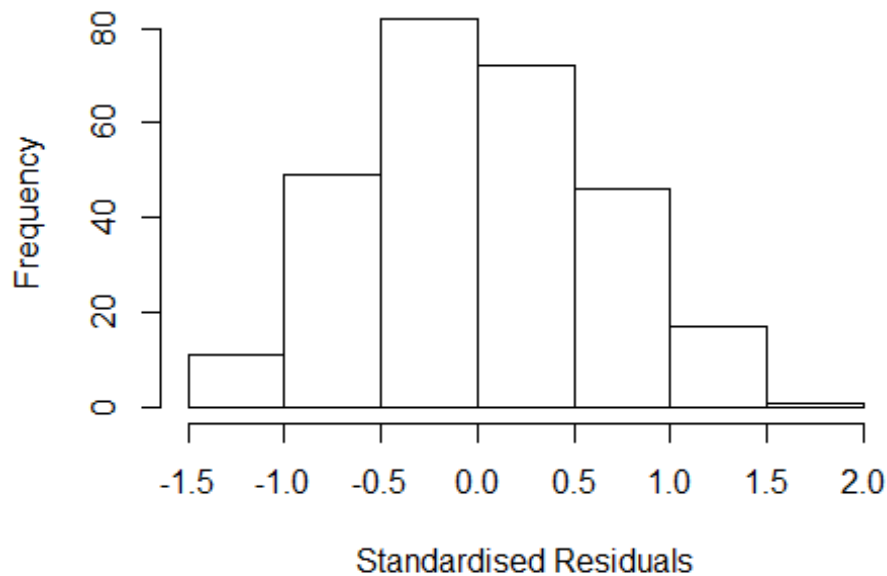
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1467 -0.4415 -0.0119 0.4609 1.6228
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7631 0.5335 1.43 0.15
## FirstAuthorFemale1 -0.1645 0.1619 -1.02 0.31
## LastAuthorFemale1 0.2707 0.1647 1.64 0.10
## Year1997 -0.3828 0.5468 -0.70 0.48
## Year1998 0.1405 0.5722 0.25 0.81
## Year1999 -0.3502 0.5543 -0.63 0.53
## Year2000 0.3828 0.5902 0.65 0.52
## Year2001 -0.1340 0.5793 -0.23 0.82
## Year2002 0.3615 0.5870 0.62 0.54
## Year2003 0.3395 0.5995 0.57 0.57
## Year2004 0.3912 0.5614 0.70 0.49
## Year2005 0.3836 0.5756 0.67 0.51
```

```

## Year2006          0.2593      0.5556      0.47      0.64
## Year2007          0.2109      0.5420      0.39      0.70
## Year2008         -0.0958      0.5430     -0.18      0.86
## Year2009          0.1264      0.5462      0.23      0.82
## Year2010         -0.0546      0.5538     -0.10      0.92
## Year2011          0.1570      0.5463      0.29      0.77
## Year2012          0.1480      0.5509      0.27      0.79
##
## Robust residual standard error: 0.64
## Multiple R-squared:  0.119, Adjusted R-squared:  0.0577
## 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.
##  0.500  0.875  0.950  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      3.60e-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 1.196 1          1.094
## 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
## -1.2234 -0.4418 -0.0185 0.4558 1.6120
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7760 0.5220 1.49 0.14
## FirstAuthorFemale1 0.0603 0.0823 0.73 0.46
## Year1997 -0.3915 0.5364 -0.73 0.47
## Year1998 0.1870 0.5631 0.33 0.74
## Year1999 -0.3449 0.5457 -0.63 0.53
## Year2000 0.3871 0.5939 0.65 0.52
## Year2001 -0.1899 0.5673 -0.33 0.74
## Year2002 0.3661 0.5732 0.64 0.52
## Year2003 0.3415 0.5885 0.58 0.56
## Year2004 0.4020 0.5509 0.73 0.47
## Year2005 0.3454 0.5625 0.61 0.54
## Year2006 0.2638 0.5472 0.48 0.63
```



```

## Year2007          0.1974      0.5308      0.37      0.71
## Year2008         -0.0701      0.5318     -0.13      0.90
## Year2009          0.1428      0.5352      0.27      0.79
## Year2010         -0.0648      0.5448     -0.12      0.91
## Year2011          0.1769      0.5350      0.33      0.74
## Year2012          0.1461      0.5400      0.27      0.79
##
## Robust residual standard error: 0.68
## Multiple R-squared:  0.107, Adjusted R-squared:  0.0481
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 259 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.553  0.892  0.956  0.924  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] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.205  1          1.098
## Year              1.205 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
## -1.144 -0.448 -0.016  0.468  1.636
##
## Coefficients:

```

```

##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.7580     0.5118   1.48    0.14
## LastAuthorFemale1 0.1255     0.0831   1.51    0.13
## Year1997         -0.3906     0.5280  -0.74    0.46
## Year1998          0.1659     0.5550   0.30    0.77
## Year1999         -0.3516     0.5330  -0.66    0.51
## Year2000          0.3692     0.5805   0.64    0.53
## Year2001         -0.1669     0.5578  -0.30    0.76
## Year2002          0.3594     0.5681   0.63    0.53
## Year2003          0.3310     0.5765   0.57    0.57
## Year2004          0.3863     0.5406   0.71    0.48
## Year2005          0.3517     0.5515   0.64    0.52
## Year2006          0.2608     0.5363   0.49    0.63
## Year2007          0.1980     0.5209   0.38    0.70
## Year2008         -0.0874     0.5225  -0.17    0.87
## Year2009          0.1208     0.5262   0.23    0.82
## Year2010         -0.0635     0.5331  -0.12    0.91
## Year2011          0.1562     0.5260   0.30    0.77
## Year2012          0.1396     0.5321   0.26    0.79
##
## Robust residual standard error: 0.665
## Multiple R-squared:  0.113, Adjusted R-squared:  0.0551
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 261 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.524  0.885   0.951   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      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 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 2006 2007 2008 2009 2010 2011
##    2    2    2    2    3    3    3    8    6    7   10   17    7   11   22
## 2012
##    10
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2006 2007 2008 2009 2010 2011
##    2    2    2    2    2    1    2    5    4    6    9   13    5   11   18
## 2012
##     8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2006 2007 2008 2009 2010 2011
##    2    2    2    2    2    1    2    5    4    6    9   13    5   11   17
## 2012
##     8
## [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"

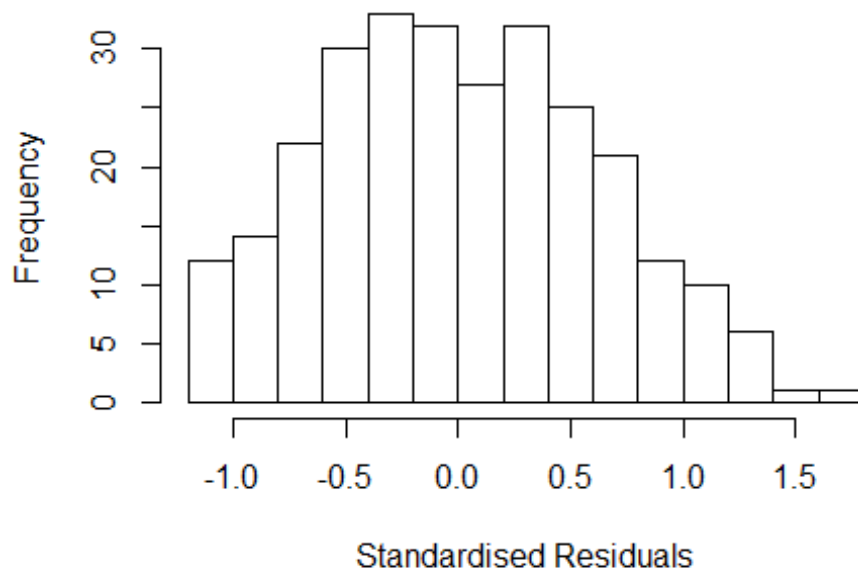
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 28, 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: 1.16652903957612"

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

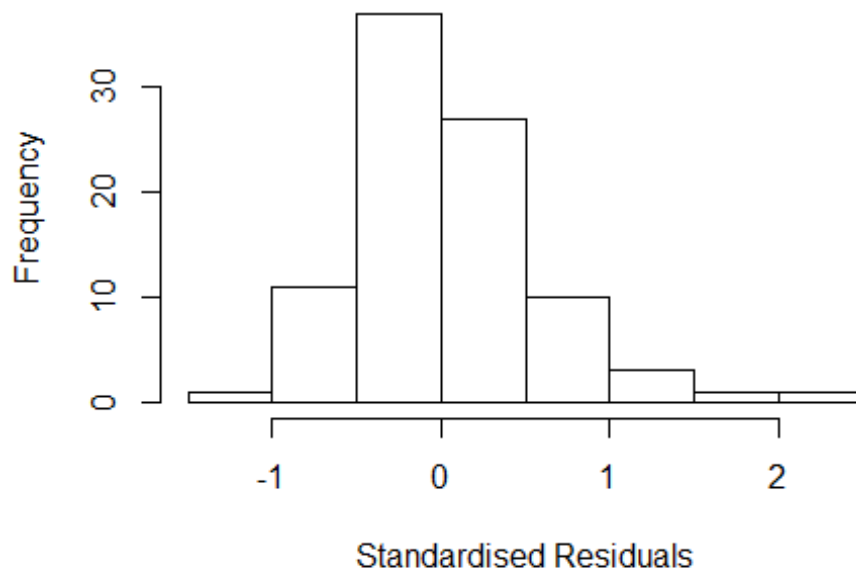
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 20, 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 16.23  1      4.029
## LastAuthorFemale  15.55  1      3.943
## UniqueAuthors    19.91  3      1.646
## Year              82.98 15      1.159
```

Residuals from first and last author and team size



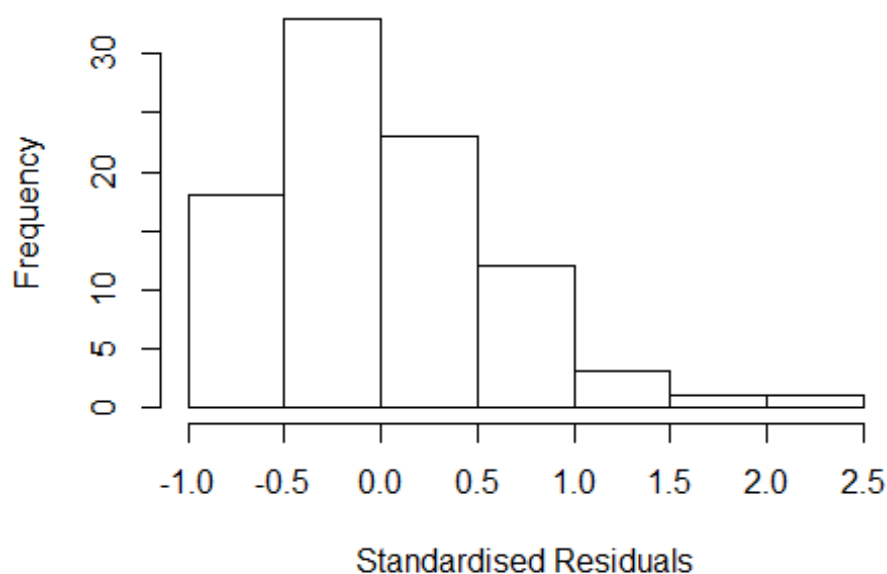
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0154 -0.3362 -0.0237 0.4016 2.1456
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2260 0.1437 8.53 1.9e-12 ***
## FirstAuthorFemale1 -0.1711 0.3254 -0.53 0.60073
## LastAuthorFemale1 0.2917 0.3272 0.89 0.37559
## UniqueAuthors2 0.4259 0.2350 1.81 0.07421 .
## UniqueAuthors3 0.2004 0.6346 0.32 0.75314
## UniqueAuthors4 -0.3528 0.3780 -0.93 0.35380
## Year1997 0.0504 0.4945 0.10 0.91913
## Year1998 -0.1378 0.2669 -0.52 0.60720
## Year1999 0.2382 0.3519 0.68 0.50071
## Year2000 -0.3228 0.3189 -1.01 0.31484
```

```

## Year2001          -0.5350      0.1437    -3.72  0.00040 ***
## Year2002           0.4250      0.6329      0.67  0.50409
## Year2003          -0.3856      0.4050     -0.95  0.34424
## Year2004           0.1544      0.6192      0.25  0.80383
## Year2006          -0.5181      0.2515     -2.06  0.04314 *
## Year2007          -0.1454      0.1884     -0.77  0.44285
## Year2008          -0.8044      0.2161     -3.72  0.00039 ***
## Year2009          -0.2250      0.3305     -0.68  0.49835
## Year2010          -0.7677      0.1791     -4.29  5.7e-05 ***
## Year2011          -0.7286      0.2212     -3.29  0.00155 **
## Year2012          -0.4183      0.2752     -1.52  0.13299
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.526
## Multiple R-squared:  0.306, Adjusted R-squared:  0.107
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 82 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.059  0.879  0.942  0.898  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      1.10e-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 13.190 1          3.632
## LastAuthorFemale  13.855 1          3.722
## Year              5.489 15          1.058

```

Residuals from first and last author



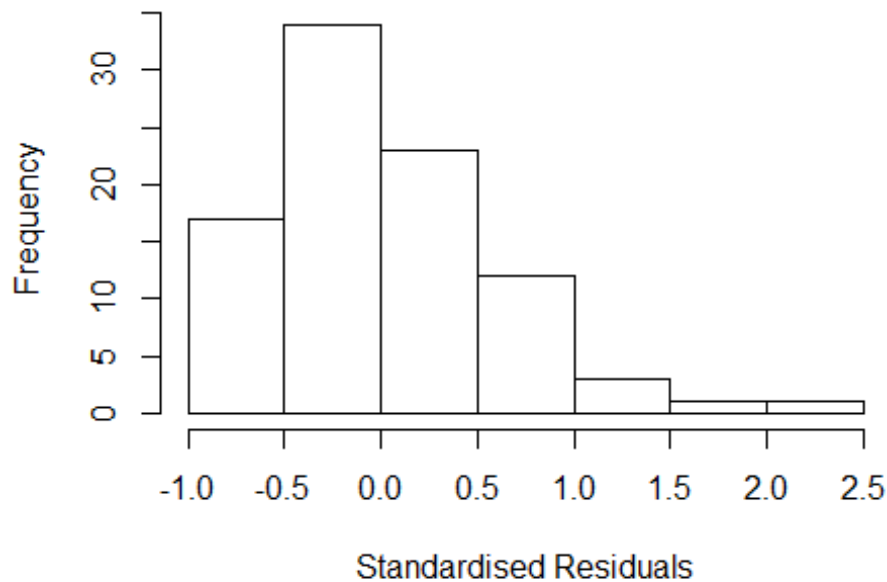
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9248 -0.3943 -0.0365 0.3664 2.0271
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2260 0.1433 8.55 1.3e-12 ***
## FirstAuthorFemale1 -0.0295 0.3105 -0.10 0.92446
## LastAuthorFemale1 0.1160 0.3087 0.38 0.70817
## Year1997 0.2928 0.2930 1.00 0.32096
## Year1998 -0.1207 0.2550 -0.47 0.63726
## Year1999 0.2553 0.3650 0.70 0.48662
## Year2000 -0.3057 0.3281 -0.93 0.35451
## Year2001 -0.5350 0.1433 -3.73 0.00037 ***
## Year2002 0.4250 0.6074 0.70 0.48633
## Year2003 -0.3012 0.3870 -0.78 0.43898
## Year2004 -0.0190 0.5817 -0.03 0.97408
## Year2006 -0.4117 0.2468 -1.67 0.09952 .
```

```

## Year2007          -0.1543      0.1867   -0.83   0.41148
## Year2008          -0.7888      0.2148   -3.67   0.00045 ***
## Year2009          -0.1193      0.3648   -0.33   0.74461
## Year2010          -0.6975      0.1861   -3.75   0.00035 ***
## Year2011          -0.6101      0.2303   -2.65   0.00988 **
## Year2012          -0.3916      0.2790   -1.40   0.16464
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.557
## Multiple R-squared:  0.254, Adjusted R-squared:  0.0804
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 82 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.158  0.886  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      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 2.468 1          1.571
## Year              2.468 15          1.031

```


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
## -0.9025 -0.4038 -0.0466 0.3746 2.0064
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2260 0.1433 8.56 1.1e-12 ***
## FirstAuthorFemale1 0.0675 0.1366 0.49 0.62283
## Year1997 0.2443 0.2903 0.84 0.40284
## Year1998 -0.1112 0.2488 -0.45 0.65613
## Year1999 0.2648 0.3709 0.71 0.47756
## Year2000 -0.2962 0.3351 -0.88 0.37954
## Year2001 -0.5350 0.1433 -3.73 0.00037 ***
## Year2002 0.4250 0.6047 0.70 0.48434
## Year2003 -0.3235 0.3614 -0.90 0.37361
## Year2004 -0.0477 0.5965 -0.08 0.93650
## Year2006 -0.4071 0.2472 -1.65 0.10380
## Year2007 -0.1342 0.1807 -0.74 0.45997
```

```

## Year2008          -0.7820      0.2137   -3.66  0.00047 ***
## Year2009          -0.1111      0.3614   -0.31  0.75935
## Year2010          -0.6975      0.1843   -3.78  0.00031 ***
## Year2011          -0.5894      0.2677   -2.20  0.03078 *
## Year2012          -0.3821      0.2787   -1.37  0.17445
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.561
## Multiple R-squared:  0.25,   Adjusted R-squared:  0.0879
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 83 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.174  0.881  0.948  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          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 2.592 1          1.610
## Year              2.592 15          1.032

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -0.9150 -0.3923 -0.0339  0.3658  2.0201
##

```

```

## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2260    0.1432   8.56 1.1e-12 ***
## LastAuthorFemale1 0.0905    0.1342   0.67 0.50239
## Year1997          0.2780    0.2581   1.08 0.28490
## Year1998         -0.1227    0.2555  -0.48 0.63241
## Year1999          0.2533    0.3615   0.70 0.48573
## Year2000         -0.3077    0.3259  -0.94 0.34810
## Year2001         -0.5350    0.1432  -3.74 0.00037 ***
## Year2002          0.4250    0.6016   0.71 0.48216
## Year2003         -0.3110    0.3521  -0.88 0.37994
## Year2004         -0.0280    0.5631  -0.05 0.96051
## Year2006         -0.4104    0.2470  -1.66 0.10083
## Year2007         -0.1528    0.1856  -0.82 0.41306
## Year2008         -0.7891    0.2142  -3.68 0.00043 ***
## Year2009         -0.1217    0.3617  -0.34 0.73744
## Year2010         -0.7001    0.1875  -3.73 0.00037 ***
## Year2011         -0.6031    0.2640  -2.28 0.02522 *
## Year2012         -0.3906    0.2783  -1.40 0.16466
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.566
## Multiple R-squared:  0.251, Adjusted R-squared:  0.089
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 82 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.175  0.887  0.946  0.908  0.984  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.10e-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] "#####"

```

```

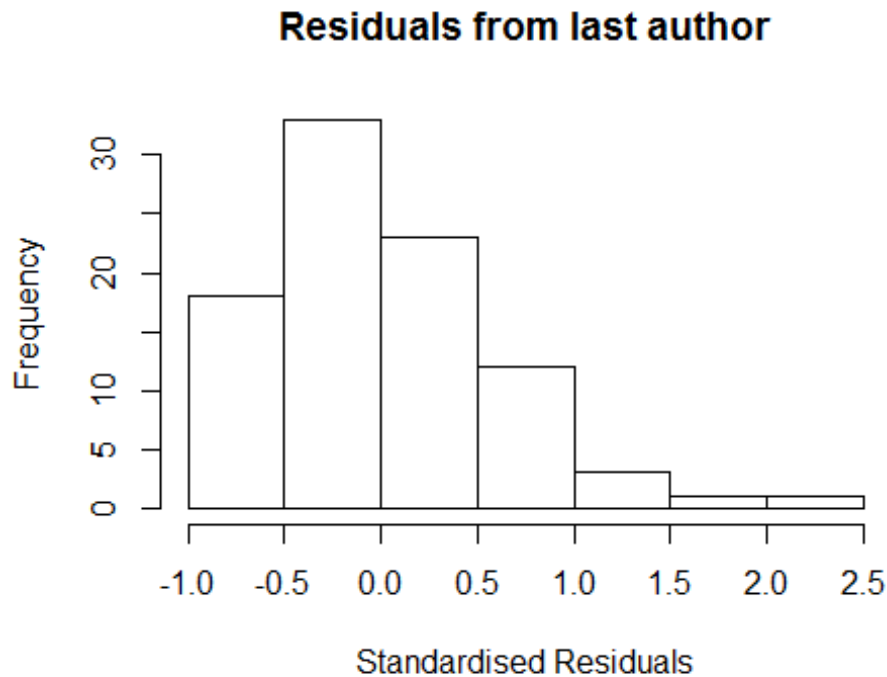
## [1] "Analysis of AJSC 1205"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    2    2    1    5    3    1    3    1    6    6    2    4    2    7
## 2012
##    5
##
## 1996 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    1    2    0    4    3    1    1    1    3    5    1    3    2    6
## 2012
##    5
##
## 1996 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    1    2    0    4    3    1    1    1    3    5    1    3    2    6
## 2012
##    5
## [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"

## Warning in wilcox.test.default(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 = NA
## 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"

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

```

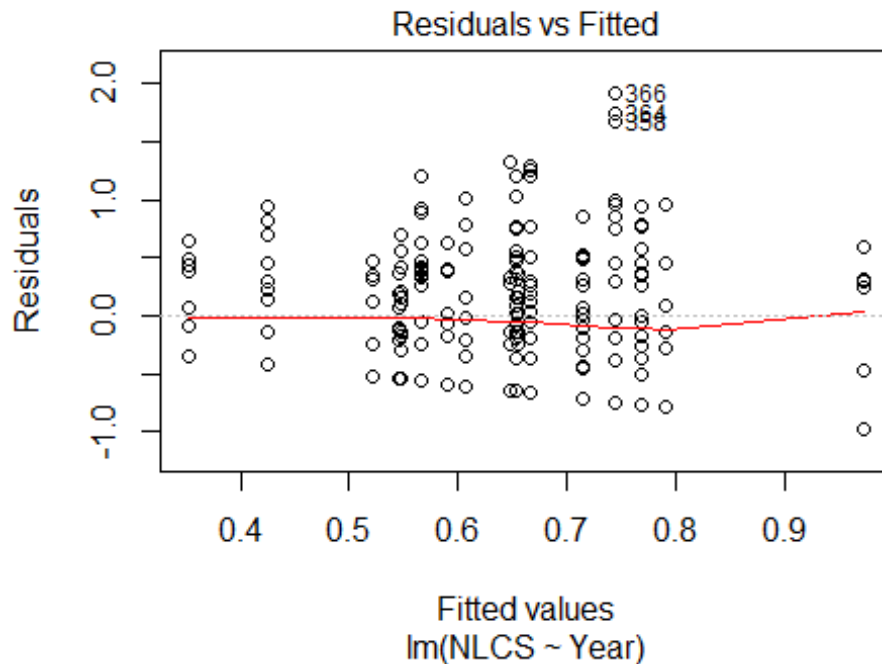


```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3, p-value = NA
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
## [1] "Regression 2: 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: 39"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1206"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
## [1] "gendered] [check that these decrease]"
##
## 2004 2005 2006 2007 2008 2009 2010
##    1    1    1    1    1    1    1
##
## 2004 2005 2006 2007 2008 2009 2010
##    1    1    1    1    1    1    0
```

```

##
## 2004 2005 2006 2007 2008 2009 2010
## 1 1 1 1 0 1 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: 5"
## [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
## 10 12 16 12 11 11 15 12 13 12 21 26 21 40 34
## 2011 2012
## 33 35
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 8 8 13 10 7 8 13 8 12 6 19 19 18 34 29
## 2011 2012
## 25 30
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 7 8 13 9 7 7 13 7 12 5 17 16 18 33 25
## 2011 2012
## 24 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 = 34, df = 16, p-value = 0.005

```



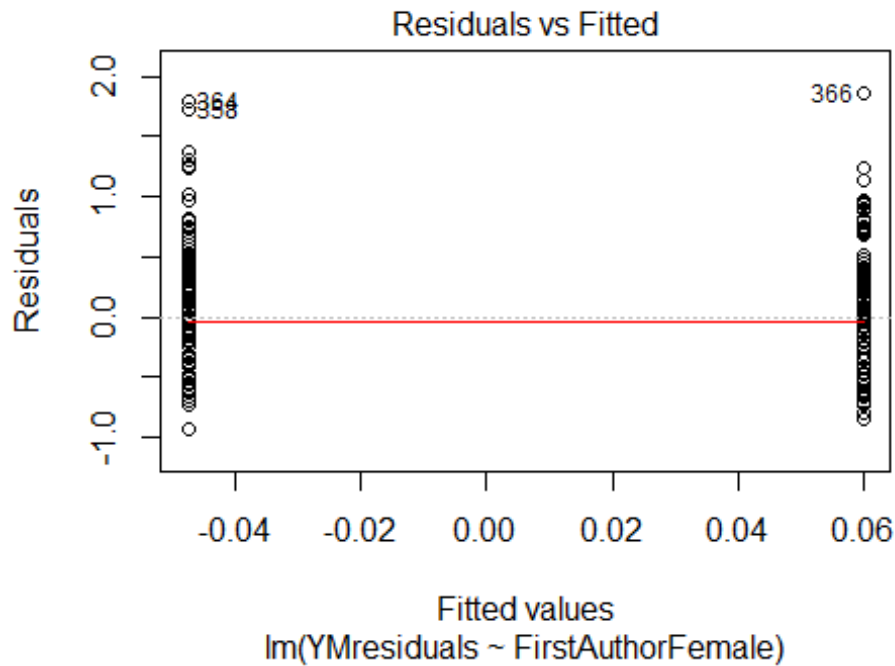
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.035, df = 1, p-value = 0.9

## [1] "Female first author team size 2018 geometric mean: 1.12246204830937"
## [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

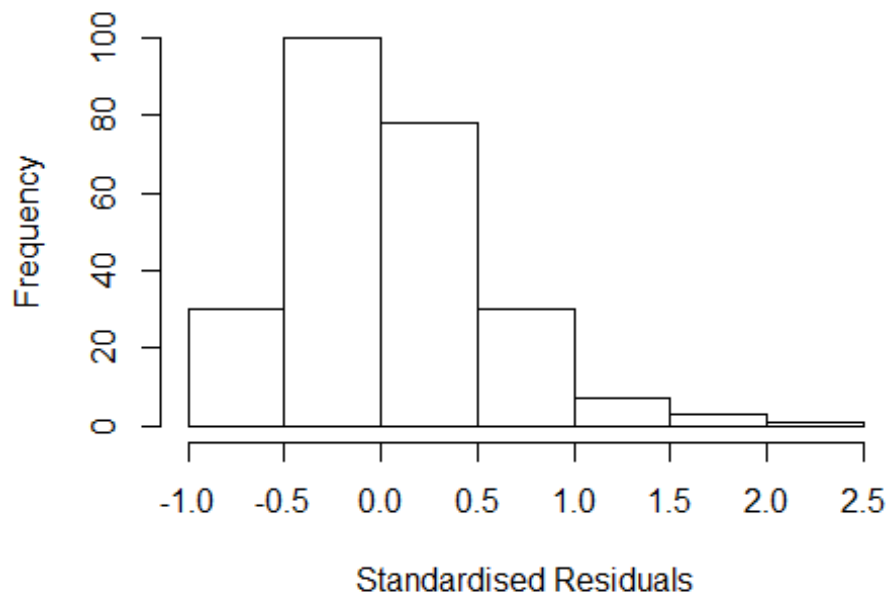
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 49, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1"
## [1] "Male last author team size 2018 geometric mean: 1.04729412282063"

## Warning in wilcox.test.default(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 = 0.6
## 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 1.342 1          1.158
## LastAuthorFemale  1.418 1          1.191
## UniqueAuthors    3.347 4          1.163
## Year              4.201 16         1.046
```


Residuals from first and last author and team size



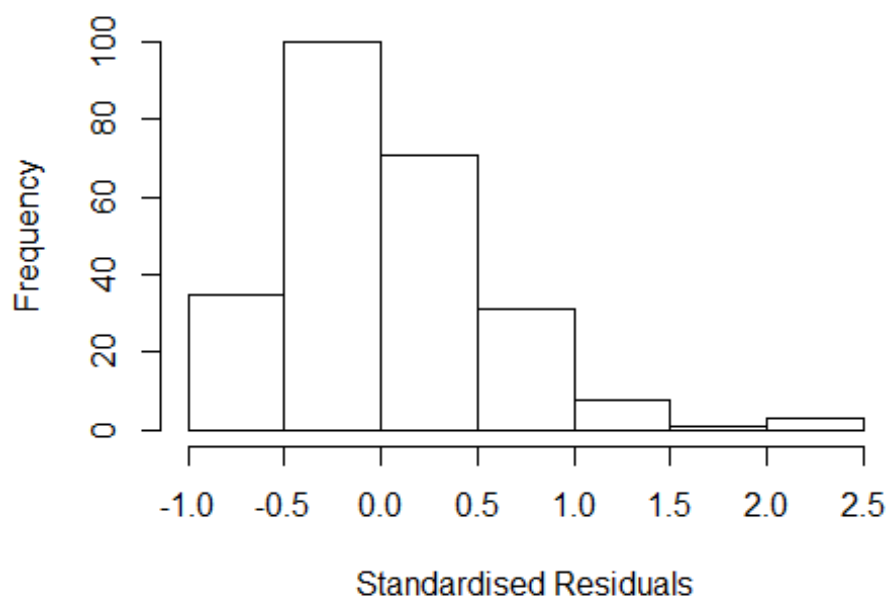
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.914 -0.331 -0.048 0.349 2.037
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.49945 0.16038 3.11 0.0021 **
## FirstAuthorFemale1 0.14003 0.06839 2.05 0.0418 *
## LastAuthorFemale1 0.07691 0.07588 1.01 0.3119
## UniqueAuthors2 -0.12563 0.09111 -1.38 0.1693
## UniqueAuthors3 -0.23585 0.08694 -2.71 0.0072 **
## UniqueAuthors4 -0.20516 0.13452 -1.53 0.1286
## UniqueAuthors5 -0.10483 0.12711 -0.82 0.4104
## Year1997 0.17566 0.24079 0.73 0.4664
## Year1998 -0.10816 0.19777 -0.55 0.5850
## Year1999 0.17373 0.15493 1.12 0.2633
```

```

## Year2000          0.30229    0.25166    1.20    0.2309
## Year2001          0.13883    0.17339    0.80    0.4241
## Year2002          0.05980    0.18794    0.32    0.7506
## Year2003          0.00671    0.25275    0.03    0.9789
## Year2004          0.05943    0.21957    0.27    0.7869
## Year2005          0.63126    0.22273    2.83    0.0050 **
## Year2006          0.36188    0.21161    1.71    0.0886 .
## Year2007          0.27623    0.18252    1.51    0.1316
## Year2008         -0.07220    0.19496   -0.37    0.7115
## Year2009          0.16899    0.17392    0.97    0.3323
## Year2010          0.08630    0.17406    0.50    0.6205
## Year2011          0.05819    0.19144    0.30    0.7614
## Year2012          0.01421    0.20514    0.07    0.9449
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.486
## Multiple R-squared:  0.142, Adjusted R-squared:  0.0582
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 232 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0398 0.8900 0.9480 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          4.02e-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 1.248 1          1.117
## LastAuthorFemale 1.315 1          1.147
## Year              1.551 16          1.014

```

Residuals from first and last author



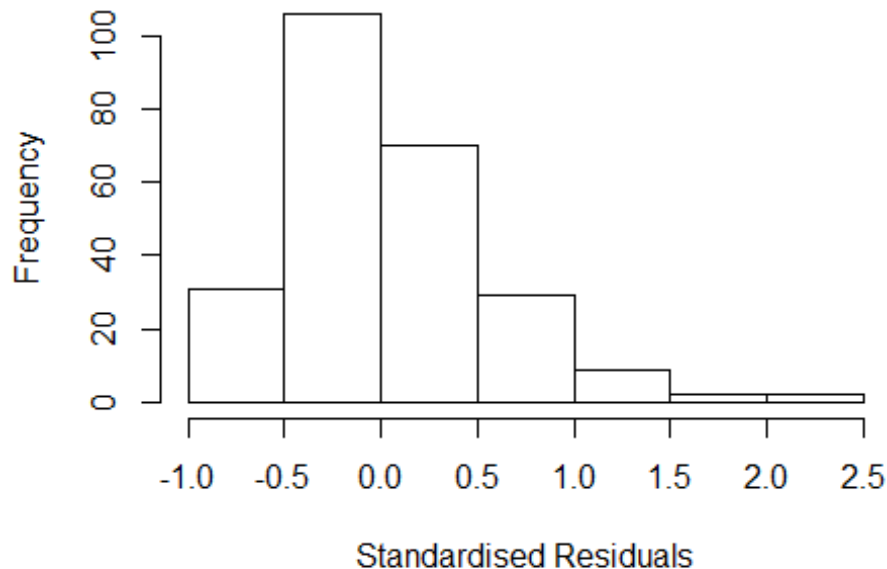
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8715 -0.3192 -0.0572 0.3832 2.0733
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.3944 0.1249 3.16 0.00180 **
## FirstAuthorFemale1 0.1016 0.0678 1.50 0.13550
## LastAuthorFemale1 0.0980 0.0817 1.20 0.23173
## Year1997 0.1894 0.2143 0.88 0.37763
## Year1998 -0.1088 0.1714 -0.63 0.52615
## Year1999 0.1879 0.1320 1.42 0.15593
## Year2000 0.2775 0.2479 1.12 0.26411
## Year2001 0.1415 0.1413 1.00 0.31770
## Year2002 0.0744 0.1622 0.46 0.64661
## Year2003 0.0135 0.2285 0.06 0.95295
## Year2004 0.0678 0.1885 0.36 0.71968
## Year2005 0.7090 0.2116 3.35 0.00094 ***
```

```

## Year2006          0.3663      0.1948      1.88  0.06125 .
## Year2007          0.2929      0.1615      1.81  0.07106 .
## Year2008         -0.0752      0.1670     -0.45  0.65277
## Year2009          0.1690      0.1513      1.12  0.26523
## Year2010          0.1081      0.1591      0.68  0.49726
## Year2011          0.0406      0.1825      0.22  0.82398
## Year2012          0.0233      0.1807      0.13  0.89753
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.49
## Multiple R-squared:  0.111, Adjusted R-squared:  0.0411
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 231 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0338 0.8860 0.9440 0.9000 0.9850 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      4.02e-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 1.193 1      1.092
## Year              1.193 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
## -0.8108 -0.3322 -0.0595 0.3815 2.0976
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.39483 0.12867 3.07 0.0024 **
## FirstAuthorFemale1 0.12747 0.06778 1.88 0.0613 .
## Year1997 0.18709 0.21845 0.86 0.3926
## Year1998 -0.09378 0.17111 -0.55 0.5842
## Year1999 0.21621 0.13502 1.60 0.1107
## Year2000 0.28848 0.24589 1.17 0.2419
## Year2001 0.15477 0.14516 1.07 0.2874
## Year2002 0.06424 0.16623 0.39 0.6995
## Year2003 -0.00235 0.22799 -0.01 0.9918
## Year2004 0.07666 0.19303 0.40 0.6916
## Year2005 0.71693 0.21617 3.32 0.0011 **
## Year2006 0.39912 0.19489 2.05 0.0417 *
```

```

## Year2007          0.31119      0.16311      1.91      0.0577 .
## Year2008         -0.06259      0.16887     -0.37      0.7113
## Year2009          0.18700      0.15613      1.20      0.2323
## Year2010          0.12325      0.16203      0.76      0.4476
## Year2011          0.07110      0.18619      0.38      0.7029
## Year2012          0.04114      0.18550      0.22      0.8247
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.488
## Multiple R-squared:  0.105, Adjusted R-squared:  0.0389
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 235 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0248 0.8820 0.9420 0.9010 0.9860 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      4.02e-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 1.255 1      1.120
## Year      1.255 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
## -0.9138 -0.3472 -0.0676  0.3742  2.0713

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.4458    0.1291    3.45 0.00066 ***
## LastAuthorFemale1 0.1307    0.0820    1.59 0.11243
## Year1997        0.1470    0.2091    0.70 0.48271
## Year1998       -0.1355    0.1711   -0.79 0.42912
## Year1999        0.1787    0.1425    1.25 0.21116
## Year2000        0.2808    0.2559    1.10 0.27357
## Year2001        0.1383    0.1498    0.92 0.35676
## Year2002        0.0855    0.1675    0.51 0.61022
## Year2003        0.0172    0.2387    0.07 0.94274
## Year2004        0.0785    0.1973    0.40 0.69106
## Year2005        0.7097    0.2026    3.50 0.00055 ***
## Year2006        0.3373    0.2030    1.66 0.09796 .
## Year2007        0.2689    0.1677    1.60 0.11023
## Year2008       -0.0871    0.1729   -0.50 0.61498
## Year2009        0.1409    0.1555    0.91 0.36579
## Year2010        0.1018    0.1659    0.61 0.53990
## Year2011        0.0085    0.1860    0.05 0.96360
## Year2012        0.0132    0.1862    0.07 0.94370
## ---
## 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:  0.0353
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 239 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0346 0.8920 0.9480 0.9030 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      4.02e-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: 249"
## [1] ""

```

```

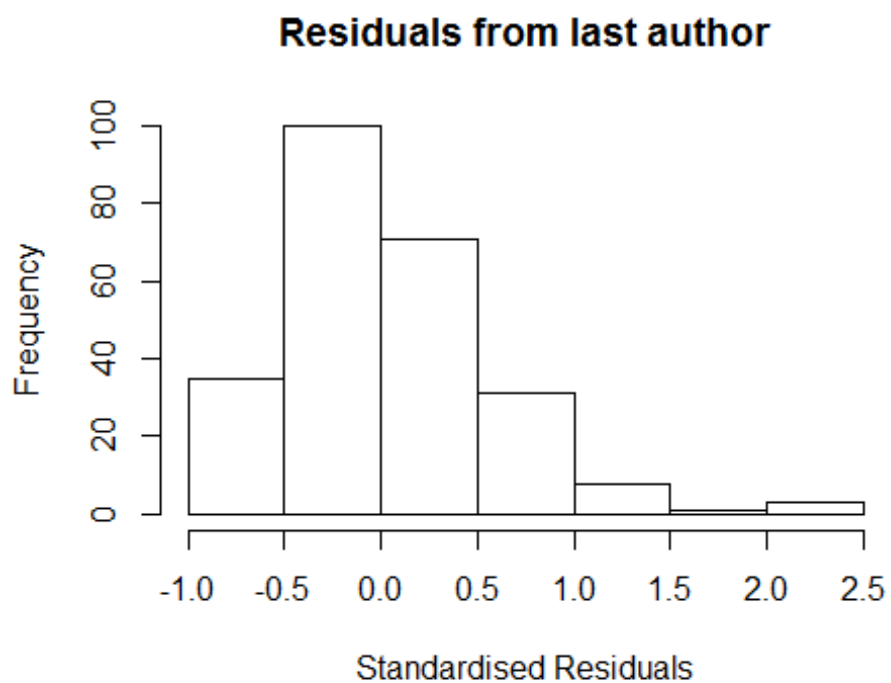
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1208"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    3    8    9    4   10   14   26   23   13   21   16   26   14   22
## 2011 2012
##   61   47
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    2    8    8    1    6   13   19   18   11   14   14   21   11   22
## 2011 2012
##   46   42
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    2    8    8    1    6   13   19   18   11   14   14   21   11   22
## 2011 2012
##   45   41
## [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: 1.03200827973421"

## Warning in wilcox.test.default(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 = 0.4
## 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: 1.03200827973421"

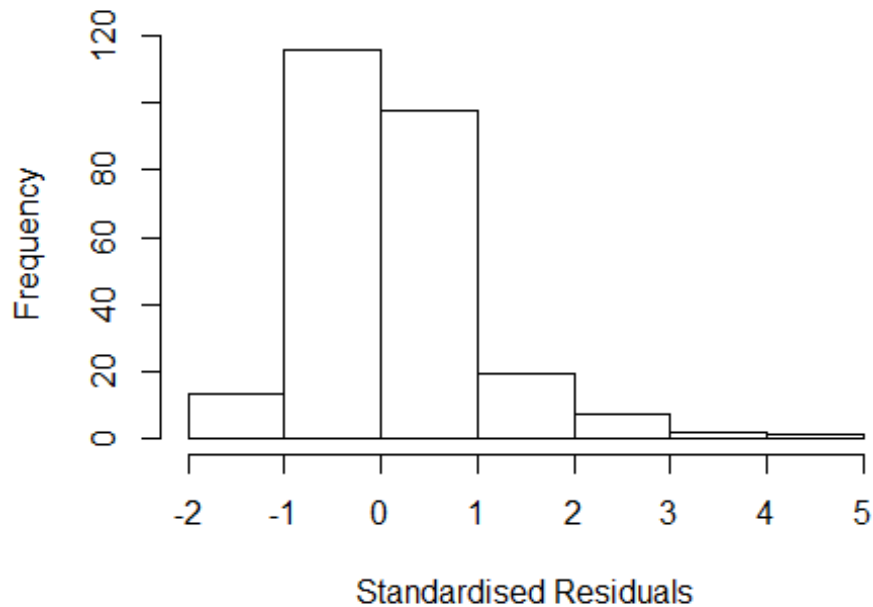
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 190, p-value = 0.4
## 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 3.686e+14  1      1.920e+07
## LastAuthorFemale  1.055e+02  1      1.027e+01
## UniqueAuthors    3.817e+14  2      4.420e+03
## Year              2.303e+13 16      2.616e+00
```

Residuals from first and last author and team size



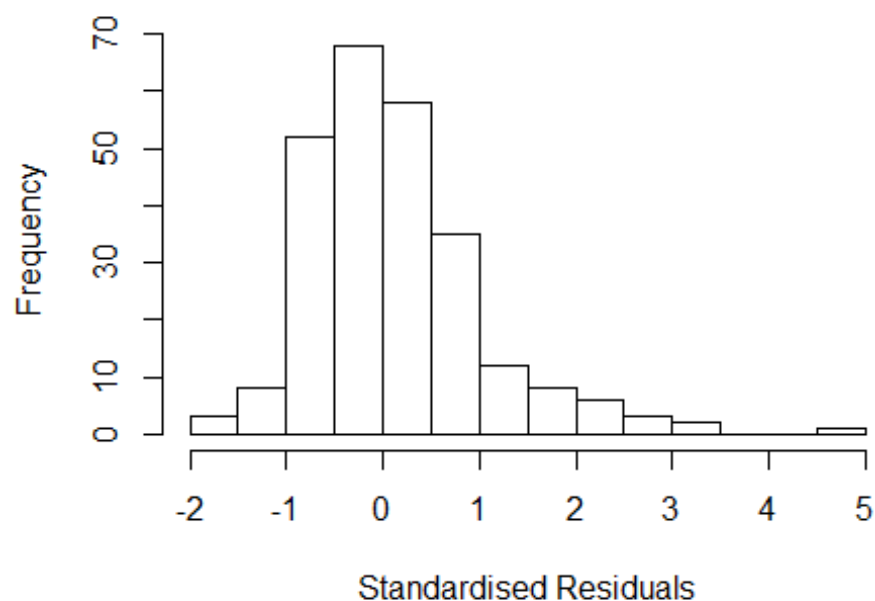
```
## [1] "List of 6 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 41  33846301457 4.222 2002    1208     1    3.253
## 42  60949945511 3.502 2002    1208     1    2.552
## 191 70350028698 3.741 2009    1208     2    3.047
## 193 84870504805 3.306 2009    1208     1    2.631
## 280 84856946247 3.391 2012    1208     1    2.740
## 303 84870415474 5.374 2012    1208     1    4.742
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q    Median        3Q        Max
## -1.72e+00 -4.35e-01  6.11e-16  5.13e-01  4.74e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.390      0.546   2.54   0.012 *
## FirstAuthorFemale1 -0.594      0.729  -0.81   0.416
## LastAuthorFemale1  0.613      0.718   0.85   0.394
## UniqueAuthors2     1.431      0.711   2.01   0.045 *
## UniqueAuthors3    -0.204      0.940  -0.22   0.829
```

```

## Year1997          -0.526      0.635   -0.83    0.408
## Year1998          -1.065      0.566   -1.88    0.061 .
## Year1999          -1.043      0.567   -1.84    0.067 .
## Year2000          -0.600      0.546   -1.10    0.273
## Year2001          -0.687      0.685   -1.00    0.317
## Year2002          -0.440      0.647   -0.68    0.497
## Year2003          -0.336      0.599   -0.56    0.575
## Year2004           0.143      0.590    0.24    0.808
## Year2005          -0.592      0.573   -1.03    0.303
## Year2006          -0.629      0.590   -1.07    0.288
## Year2007          -0.352      0.605   -0.58    0.561
## Year2008          -0.696      0.567   -1.23    0.221
## Year2009          -0.715      0.662   -1.08    0.281
## Year2010          -0.955      0.561   -1.70    0.090 .
## Year2011          -1.097      0.554   -1.98    0.049 *
## Year2012          -0.758      0.562   -1.35    0.179
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.714
## Multiple R-squared:  0.212, Adjusted R-squared:  0.145
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## observation 236 is an outlier with |weight| = 0 ( < 0.00039);
## 14 weights are ~= 1. The remaining 241 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0031 0.8850 0.9530 0.8890 0.9850 0.9990
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          3.91e-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 43.423 1          6.590
## LastAuthorFemale  43.757 1          6.615
## Year              1.809 16          1.019

```

Residuals from first and last author



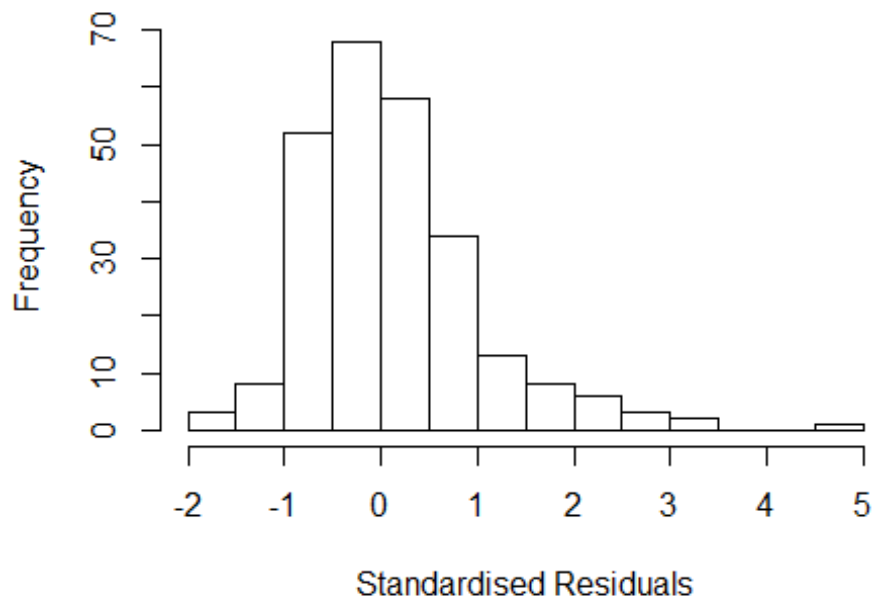
```
## [1] "List of 6 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 41  33846301457 4.222 2002    1208     1    3.245
## 42  60949945511 3.502 2002    1208     1    2.540
## 191 70350028698 3.741 2009    1208     2    3.036
## 193 84870504805 3.306 2009    1208     1    2.617
## 280 84856946247 3.391 2012    1208     1    2.741
## 303 84870415474 5.374 2012    1208     1    4.740
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.549 -0.445 -0.046  0.517  4.740
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.3900     0.5441   2.55   0.011 *
## FirstAuthorFemale1  0.0810     0.5248   0.15   0.878
## LastAuthorFemale1 -0.0653     0.5254  -0.12   0.901
## Year1997         -0.5265     0.6329  -0.83   0.406
## Year1998         -1.0630     0.5645  -1.88   0.061 .
## Year1999         -1.0418     0.5659  -1.84   0.067 .
## Year2000         -0.6000     0.5441  -1.10   0.271
```

```

## Year2001      -0.6851      0.6842      -1.00      0.318
## Year2002      -0.4284      0.6653      -0.64      0.520
## Year2003      -0.3342      0.5980      -0.56      0.577
## Year2004       0.1434      0.5867       0.24      0.807
## Year2005      -0.5914      0.5715      -1.03      0.302
## Year2006      -0.6257      0.5912      -1.06      0.291
## Year2007      -0.3503      0.6032      -0.58      0.562
## Year2008      -0.6938      0.5668      -1.22      0.222
## Year2009      -0.7007      0.6847      -1.02      0.307
## Year2010      -0.8520      0.5583      -1.53      0.128
## Year2011      -1.0391      0.5573      -1.86      0.064 .
## Year2012      -0.7560      0.5702      -1.33      0.186
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.722
## Multiple R-squared:  0.163, Adjusted R-squared:  0.0996
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## observation 236 is an outlier with |weight| = 0 ( < 0.00039);
## 14 weights are ~= 1. The remaining 241 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0063 0.8880 0.9500 0.8870 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      3.91e-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 1.747 1      1.322
## Year      1.747 16      1.018

```

Residuals from first author



```
## [1] "List of 6 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 41  33846301457 4.222 2002    1208     1    3.245
## 42  60949945511 3.502 2002    1208     1    2.540
## 191 70350028698 3.741 2009    1208     2    3.036
## 193 84870504805 3.306 2009    1208     1    2.617
## 280 84856946247 3.391 2012    1208     1    2.741
## 303 84870415474 5.374 2012    1208     1    4.740
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5511 -0.4417 -0.0473  0.5181  4.7456
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.3900     0.5469   2.54   0.012 *
## FirstAuthorFemale1 0.0175     0.1051   0.17   0.868
## Year1997       -0.5265     0.6357  -0.83   0.408
## Year1998       -1.0641     0.5668  -1.88   0.062 .
## Year1999       -1.0426     0.5680  -1.84   0.068 .
## Year2000        -0.6000     0.5469  -1.10   0.274
## Year2001        -0.6874     0.6861  -1.00   0.317
```

```

## Year2002          -0.4426      0.6465    -0.68      0.494
## Year2003          -0.3359      0.5991    -0.56      0.575
## Year2004           0.1437      0.5905     0.24      0.808
## Year2005          -0.5921      0.5740    -1.03      0.303
## Year2006          -0.6289      0.5907    -1.06      0.288
## Year2007          -0.3515      0.6056    -0.58      0.562
## Year2008          -0.6957      0.5678    -1.23      0.222
## Year2009          -0.7179      0.6608    -1.09      0.278
## Year2010          -0.8507      0.5599    -1.52      0.130
## Year2011          -1.0459      0.5550    -1.88      0.061 .
## Year2012          -0.7616      0.5661    -1.35      0.180
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.712
## Multiple R-squared:  0.165, Adjusted R-squared:  0.106
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## observation 236 is an outlier with |weight| = 0 ( < 0.00039);
## 14 weights are ~= 1. The remaining 241 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0022 0.8850 0.9480 0.8840 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          3.91e-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 1.747 1          1.322
## Year            1.747 16          1.018
##
## [1] "List of 6 outliers with residuals above 2.5"
##      ScopusId NLCS Year OneField Fields residuals
## 41 33846301457 4.222 2002 1208 1 3.245
## 42 60949945511 3.502 2002 1208 1 2.540
## 191 70350028698 3.741 2009 1208 2 3.036
## 193 84870504805 3.306 2009 1208 1 2.617

```

```

## 280 84856946247 3.391 2012      1208      1      2.741
## 303 84870415474 5.374 2012      1208      1      4.740
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.549 -0.442 -0.049  0.517  4.744
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3900     0.5470   2.54  0.012 *
## LastAuthorFemale1  0.0138     0.1050   0.13  0.896
## Year1997        -0.5265     0.6357  -0.83  0.408
## Year1998        -1.0623     0.5674  -1.87  0.062 .
## Year1999        -1.0418     0.5688  -1.83  0.068 .
## Year2000        -0.6000     0.5470  -1.10  0.274
## Year2001        -0.6866     0.6884  -1.00  0.320
## Year2002        -0.4409     0.6617  -0.67  0.506
## Year2003        -0.3344     0.6006  -0.56  0.578
## Year2004         0.1449     0.5895   0.25  0.806
## Year2005        -0.5909     0.5743  -1.03  0.305
## Year2006        -0.6273     0.5941  -1.06  0.292
## Year2007        -0.3503     0.6065  -0.58  0.564
## Year2008        -0.6941     0.5696  -1.22  0.224
## Year2009        -0.7153     0.6792  -1.05  0.293
## Year2010        -0.8490     0.5614  -1.51  0.132
## Year2011        -1.0453     0.5602  -1.87  0.063 .
## Year2012        -0.7598     0.5728  -1.33  0.186
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.712
## Multiple R-squared:  0.165, Adjusted R-squared:  0.106
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## observation 236 is an outlier with |weight| = 0 ( < 0.00039);
## 14 weights are ~ 1. The remaining 241 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.002  0.885  0.948  0.884  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      3.91e-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: 256"
## [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]"
##
## 2004 2005 2007 2008 2011
##    1    2    1    1    1
##
## 2004 2005 2007 2008 2011
##    1    2    1    1    0
##
## 2004 2005 2007 2008 2011
##    1    2    1    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: 4"
## [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 2000 2001 2002 2003 2005 2006 2008 2009 2010 2011 2012
##    1    3    1    2    4    7    3    6    4    4    2    9
##
## 1996 2000 2001 2002 2003 2005 2006 2008 2009 2010 2011 2012

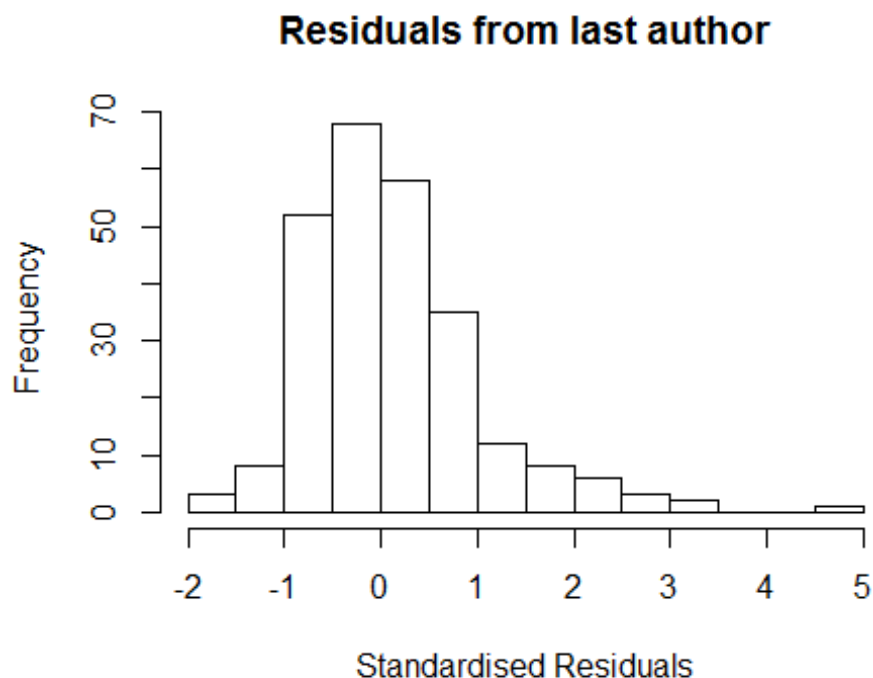
```

```
##      1      2      1      2      3      7      3      6      2      4      1      8
##
## 1996 2000 2001 2002 2003 2005 2006 2008 2009 2010 2011 2012
##      1      2      1      1      2      7      3      6      2      4      1      8
## [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: 1.73205080756888"

## 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 = 0.6
## 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: 1.73205080756888"

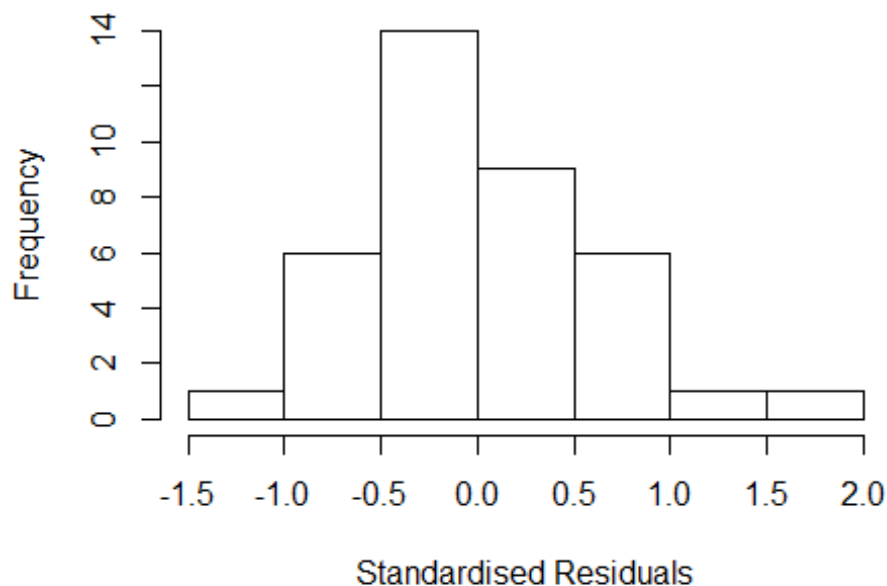
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
```

```
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4, p-value = 0.6
## 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 -4.632e+12  1          NaN
## LastAuthorFemale  1.839e+02  1          13.56
## UniqueAuthors    1.029e+16  2          10071.32
## Year              -2.671e+18 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
## -1.32e+00 -2.01e-01  1.11e-15  1.77e-01  1.69e+00
```

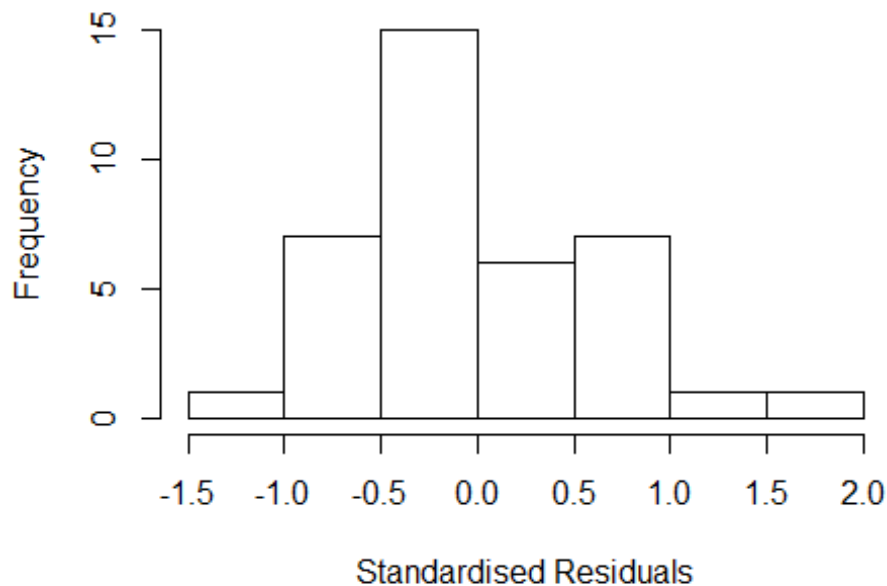
```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.8233     0.3224   11.86 5.0e-11 ***
## FirstAuthorFemale1 -1.1988     0.1017  -11.79 5.6e-11 ***
## LastAuthorFemale1  1.2235     0.2846    4.30 0.00029 ***
## UniqueAuthors2     0.0125     0.3455    0.04 0.97135
## UniqueAuthors3     1.4194     0.2178    6.52 1.5e-06 ***
## Year2000          -2.3403     0.3256   -7.19 3.3e-07 ***
## Year2001          -2.7973     0.3224   -8.68 1.5e-08 ***
## Year2002          -2.4243     0.3224   -7.52 1.6e-07 ***
## Year2003          -1.5446     0.3599   -4.29 0.00030 ***
## Year2005          -2.0011     0.3717   -5.38 2.1e-05 ***
## Year2006          -1.7988     0.1981   -9.08 6.8e-09 ***
## Year2008          -3.2417     0.4005   -8.09 4.9e-08 ***
## Year2009          -3.8202     0.3230  -11.83 5.2e-11 ***
## Year2010          -2.5067     0.4879   -5.14 3.8e-05 ***
## Year2011          -3.0803     0.3224   -9.55 2.8e-09 ***
## Year2012          -3.2777     0.4949   -6.62 1.2e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.58
## Multiple R-squared:  0.633, Adjusted R-squared:  0.382
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 26 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.375  0.905   0.923   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      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 2: First author gender, Last author gender, Year as
factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 4.008e+15 1      6.331e+07

```

```
## LastAuthorFemale 2.893e+02 1 1.701e+01
## Year 1.097e+18 11 6.607e+00
```

Residuals from first and last author



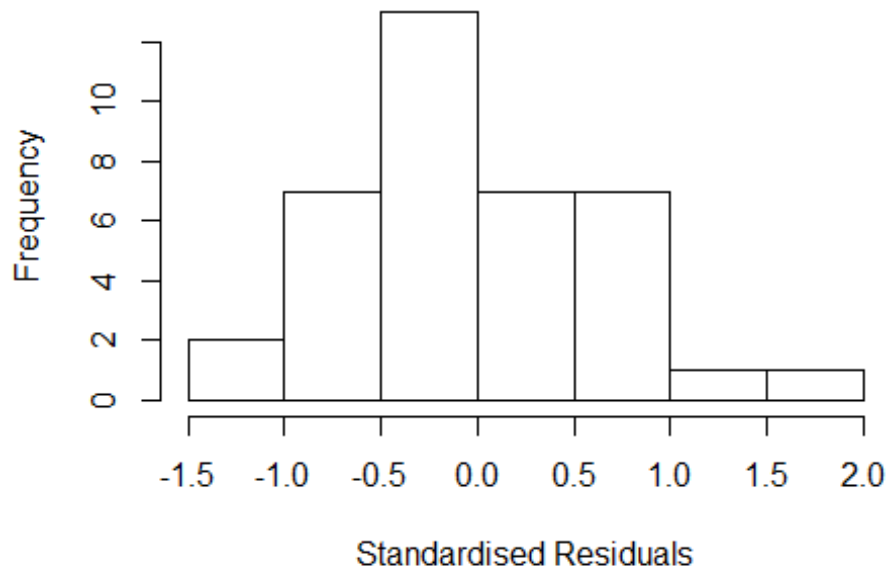
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.283249 -0.267995 -0.000202 0.198107 1.643945
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 3.8340 0.0964 39.79 < 2e-16 ***
## FirstAuthorFemale1 -1.1970 0.0964 -12.42 6.1e-12 ***
## LastAuthorFemale1 1.1200 0.4175 2.68 0.013 *
## Year2000 -2.3510 0.1065 -22.08 < 2e-16 ***
## Year2001 -2.8080 0.0964 -29.14 < 2e-16 ***
## Year2002 -2.4350 0.0964 -25.27 < 2e-16 ***
## Year2003 -1.5045 0.2583 -5.83 5.2e-06 ***
## Year2005 -1.9946 0.2277 -8.76 6.1e-09 ***
## Year2006 -1.8060 0.0542 -33.34 < 2e-16 ***
```

```

## Year2008          -2.7810      0.3658   -7.60   7.7e-08 ***
## Year2009          -3.7220      0.4319   -8.62   8.3e-09 ***
## Year2010          -2.5508      0.4648   -5.49   1.2e-05 ***
## Year2011          -3.0910      0.0964  -32.08   < 2e-16 ***
## Year2012          -3.1380      0.5416   -5.79   5.7e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.679
## Multiple R-squared:  0.498, Adjusted R-squared:  0.225
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 25 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.537  0.893   0.926   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      2.63e-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 2.797e+14 1      1.673e+07
## Year              2.797e+14 11      4.536e+00

```

Residuals from first author



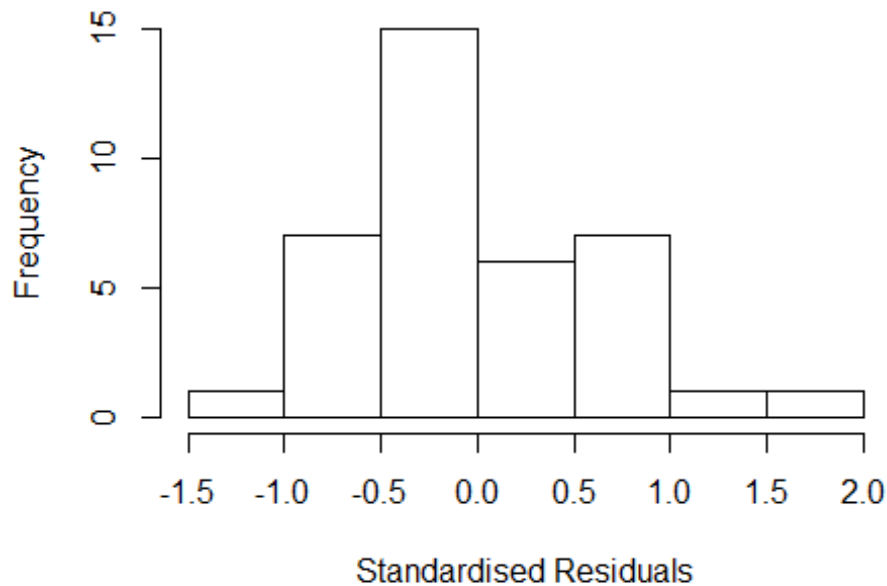
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.26e+00 -4.64e-01 -9.99e-16 4.51e-01 1.65e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 3.064 0.330 9.30 1.4e-09 ***
## FirstAuthorFemale1 -0.427 0.330 -1.29 0.20716
## Year2000 -1.581 0.333 -4.75 7.1e-05 ***
## Year2001 -2.038 0.330 -6.18 1.8e-06 ***
## Year2002 -1.665 0.330 -5.05 3.3e-05 ***
## Year2003 -0.559 0.307 -1.82 0.08028 .
## Year2005 -1.170 0.320 -3.66 0.00118 **
## Year2006 -1.292 0.338 -3.83 0.00077 ***
## Year2008 -1.954 0.422 -4.63 9.8e-05 ***
## Year2009 -2.217 0.354 -6.25 1.5e-06 ***
## Year2010 -1.803 0.538 -3.35 0.00255 **
## Year2011 -2.321 0.330 -7.04 2.2e-07 ***
```

```

## Year2012          -2.028      0.399   -5.08  3.0e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.781
## Multiple R-squared:  0.438, Adjusted R-squared:  0.168
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 28 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.633  0.914  0.946  0.927  0.974  0.999
## Algorithmic parameters:
##           tuning.chi           bb           tuning.psi           refine.tol
##           1.55e+00           5.00e-01           4.69e+00           1.00e-07
##           rel.tol           solve.tol           eps.outlier           eps.x
##           1.00e-07           1.00e-07           2.63e-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 -603 1          NaN
## Year             -603 11          NaN

```


Residuals from last author



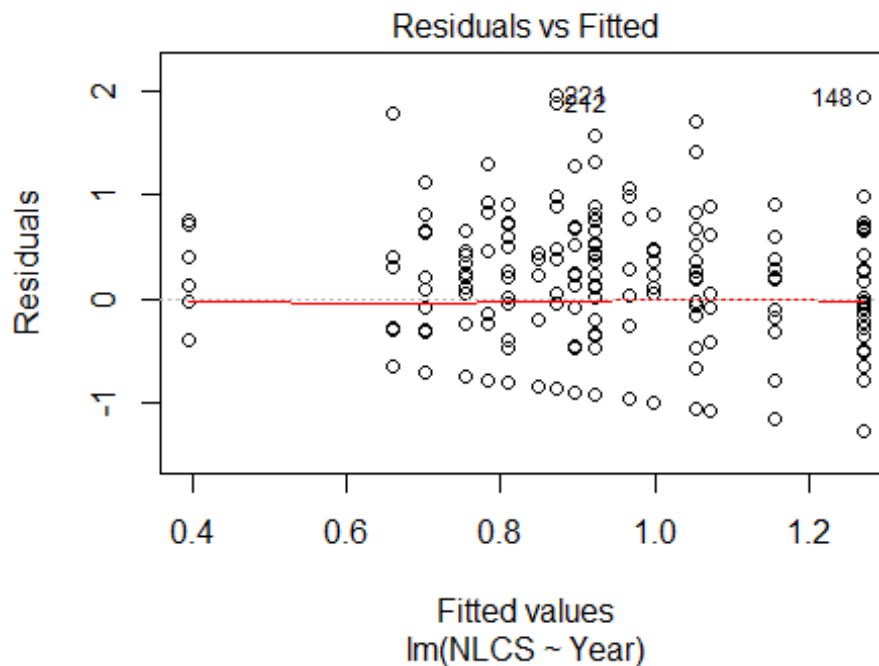
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.254085 -0.580146 -0.000221 0.587157 1.602433
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.64e+00 3.48e-08 7.58e+07 < 2e-16 ***
## LastAuthorFemale1 -7.83e-02 4.00e-01 -2.00e-01 0.84621
## Year2000 -1.15e+00 4.53e-02 -2.55e+01 < 2e-16 ***
## Year2001 -1.61e+00 2.57e-08 -6.26e+07 < 2e-16 ***
## Year2002 -1.24e+00 3.44e-08 -3.59e+07 < 2e-16 ***
## Year2003 -3.07e-01 2.37e-01 -1.30e+00 0.20700
## Year2005 -7.98e-01 2.01e-01 -3.97e+00 0.00054 ***
## Year2006 -9.97e-01 3.54e-01 -2.82e+00 0.00936 **
## Year2008 -1.58e+00 3.30e-01 -4.78e+00 6.5e-05 ***
## Year2009 -1.93e+00 6.44e-01 -2.99e+00 0.00619 **
## Year2010 -1.38e+00 4.30e-01 -3.22e+00 0.00358 **
## Year2011 -1.89e+00 3.70e-08 -5.12e+07 < 2e-16 ***
```

```

## Year2012          -1.90e+00   5.02e-01 -3.78e+00  0.00087 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.822
## Multiple R-squared:  0.404, Adjusted R-squared:  0.118
## 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.
##   0.684  0.924  0.943   0.929   0.955   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           2.63e-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 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
##   10   12   10   14    7   12   12   12   13    6   21   19   15   13   24
## 2011 2012
##   21   32
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    9   11    9   13    5   11   10   10   11    6   16   16   14   12   21
## 2011 2012
##   17   30
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    9   11    9   13    4   11   10   10   11    6   16   16   14   12   20
## 2011 2012

```

```
## 17 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 = 12, df = 16, p-value = 0.7
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.016, df = 1, p-value = 0.9

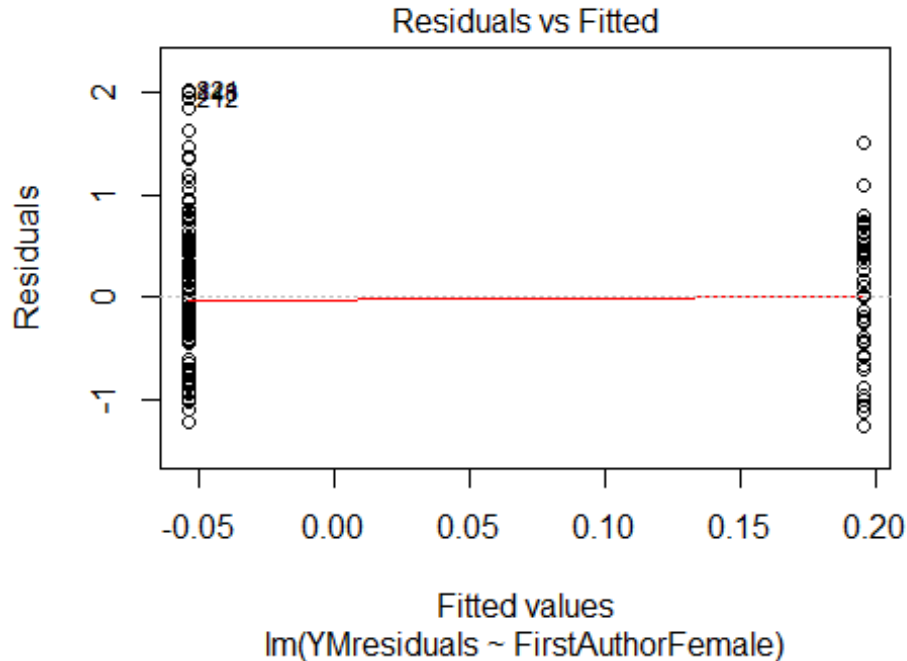
## [1] "Female first author team size 2018 geometric mean: 1.07177346253629"
## [1] "Male first author team size 2018 geometric mean: 1.02189714865412"

## Warning in wilcox.test.default(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 = 0.4
## alternative hypothesis: true location shift is not equal to 0
```

```
##
## [1] "Female last author team size 2018 geometric mean: 1.07177346253629"
## [1] "Male last author team size 2018 geometric mean: 1.02189714865412"

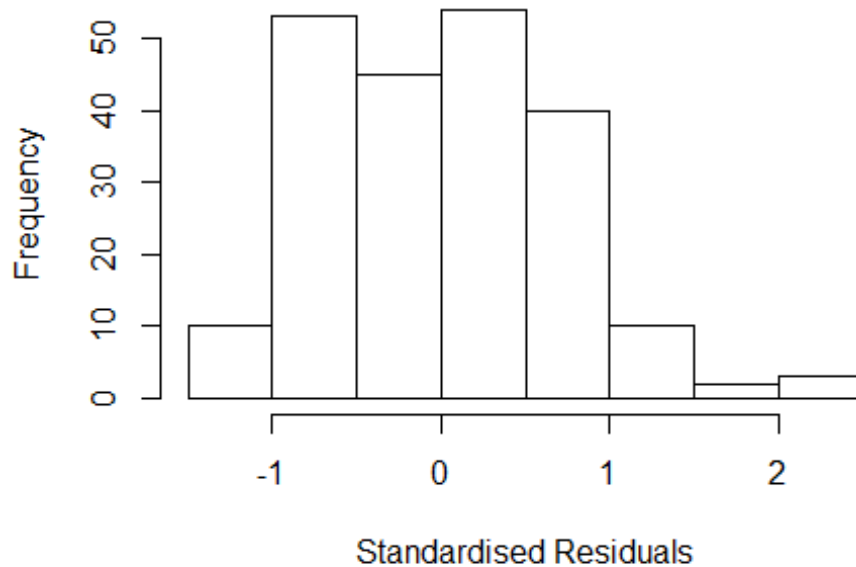
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 170, p-value = 0.4
## 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 | 46.67 | 1 | 6.832 |
| LastAuthorFemale | 35.27 | 1 | 5.939 |
| UniqueAuthors | 671.54 | 4 | 2.256 |
| Year | 190.87 | 16 | 1.178 |

Residuals from first and last author and team size



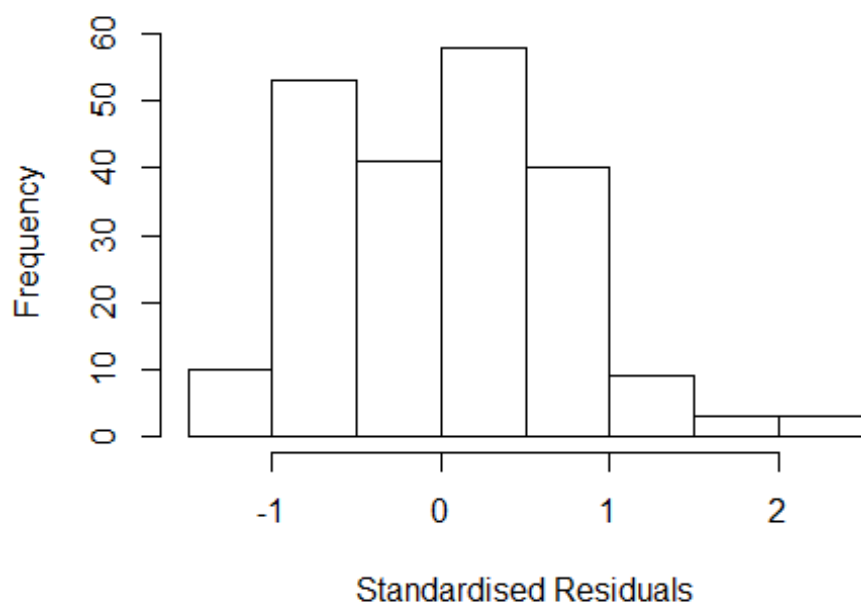
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2273 -0.5949 0.0195 0.5003 2.2401
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.85097 0.32924 2.58 0.01 *
## FirstAuthorFemale1 0.12488 0.27170 0.46 0.65
## LastAuthorFemale1 0.20809 0.26856 0.77 0.44
## UniqueAuthors2 0.34192 0.21297 1.61 0.11
## UniqueAuthors3 -0.30704 0.40342 -0.76 0.45
## UniqueAuthors4 0.00247 0.22128 0.01 0.99
## UniqueAuthors5 -0.19781 0.18889 -1.05 0.30
## Year1997 0.35751 0.37851 0.94 0.35
## Year1998 -0.34384 0.37020 -0.93 0.35
## Year1999 -0.19146 0.37260 -0.51 0.61
```

```

## Year2000      -0.08849    0.41934   -0.21    0.83
## Year2001      0.26574    0.36342    0.73    0.47
## Year2002     -0.52826    0.35915   -1.47    0.14
## Year2003     -0.18087    0.40031   -0.45    0.65
## Year2004      0.15625    0.38641    0.40    0.69
## Year2005      0.11357    0.40918    0.28    0.78
## Year2006     -0.15429    0.37586   -0.41    0.68
## Year2007      0.24383    0.36452    0.67    0.50
## Year2008      0.03825    0.37804    0.10    0.92
## Year2009     -0.24361    0.37115   -0.66    0.51
## Year2010      0.04333    0.38136    0.11    0.91
## Year2011     -0.25603    0.42292   -0.61    0.55
## Year2012     -0.05212    0.37025   -0.14    0.89
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.72
## Multiple R-squared:  0.132, Adjusted R-squared:  0.0334
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 201 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.312  0.891  0.939  0.917  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      4.61e-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 3.985 1      1.996
## LastAuthorFemale  4.437 1      2.106
## Year              1.757 16      1.018

```

Residuals from first and last author

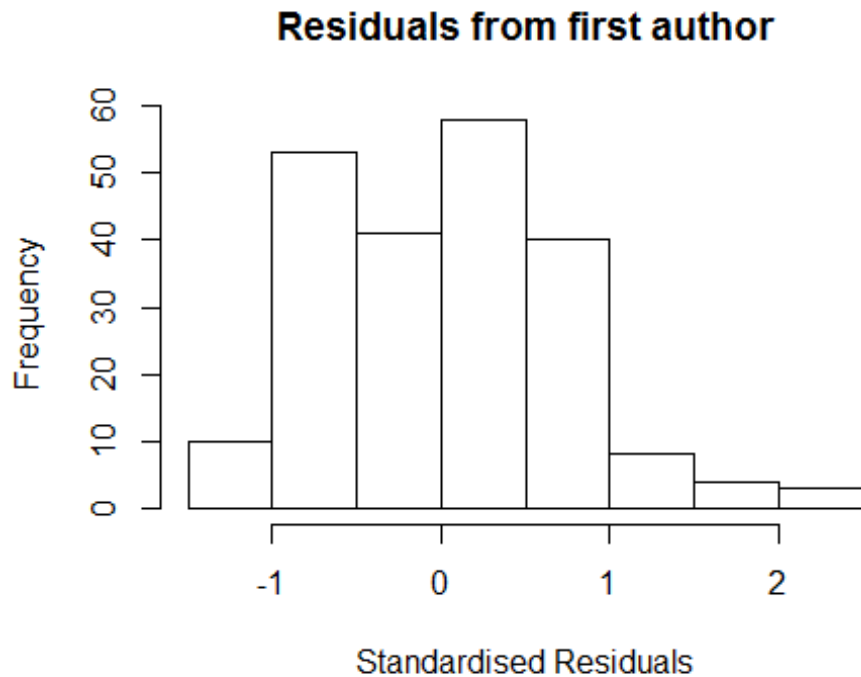


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2642 -0.6177 0.0445 0.5021 2.2173
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.91581 0.31373 2.92 0.0039 **
## FirstAuthorFemale1 0.23438 0.22944 1.02 0.3082
## LastAuthorFemale1 0.11097 0.25617 0.43 0.6654
## Year1997 0.33003 0.37397 0.88 0.3786
## Year1998 -0.41429 0.35322 -1.17 0.2422
## Year1999 -0.25808 0.35786 -0.72 0.4717
## Year2000 -0.15292 0.40735 -0.38 0.7078
## Year2001 0.19939 0.34627 0.58 0.5654
## Year2002 -0.59557 0.34293 -1.74 0.0840 .
## Year2003 -0.24936 0.38509 -0.65 0.5180
## Year2004 0.09168 0.37150 0.25 0.8053
## Year2005 -0.00709 0.37945 -0.02 0.9851
```

```

## Year2006      -0.17779    0.35855   -0.50    0.6205
## Year2007      0.19747    0.34582    0.57    0.5686
## Year2008     -0.02687    0.36467   -0.07    0.9413
## Year2009     -0.26733    0.36246   -0.74    0.4617
## Year2010      0.00304    0.38055    0.01    0.9936
## Year2011     -0.29816    0.41326   -0.72    0.4715
## Year2012     -0.13243    0.34961   -0.38    0.7053
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.708
## Multiple R-squared:  0.119, Adjusted R-squared:  0.0393
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 200 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.307  0.892  0.937   0.912  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      4.61e-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 1.215 1      1.102
## Year              1.215 16      1.006

```

```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2474 -0.6371 0.0453 0.5077 2.1979
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9353 0.2986 3.13 0.002 **
## FirstAuthorFemale1 0.3237 0.1267 2.55 0.011 *
## Year1997 0.3121 0.3617 0.86 0.389
## Year1998 -0.4231 0.3540 -1.20 0.233
## Year1999 -0.2743 0.3482 -0.79 0.432
## Year2000 -0.1732 0.3940 -0.44 0.661
## Year2001 0.1823 0.3347 0.54 0.586
## Year2002 -0.6107 0.3327 -1.84 0.068 .
## Year2003 -0.2623 0.3787 -0.69 0.489
## Year2004 0.0714 0.3568 0.20 0.842
## Year2005 -0.0162 0.3762 -0.04 0.966
## Year2006 -0.2007 0.3407 -0.59 0.556
```

```

## Year2007          0.1884      0.3430      0.55      0.583
## Year2008          -0.0458     0.3517     -0.13     0.896
## Year2009          -0.2997     0.3338     -0.90     0.370
## Year2010          -0.0236     0.3576     -0.07     0.947
## Year2011          -0.2982     0.4160     -0.72     0.474
## Year2012          -0.1417     0.3446     -0.41     0.681
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.732
## Multiple R-squared:  0.116, Adjusted R-squared:  0.0401
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 199 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.347  0.896  0.936  0.917  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      4.61e-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 1.291 1      1.136
## Year      1.291 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
## -1.2658 -0.6268  0.0424  0.4931  2.2082

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.88651    0.31076   2.85  0.0048 **
## LastAuthorFemale1 0.31928    0.13908   2.30  0.0227 *
## Year1997        0.36179    0.37029   0.98  0.3297
## Year1998       -0.38016    0.35127  -1.08  0.2805
## Year1999       -0.22627    0.35515  -0.64  0.5248
## Year2000       -0.12374    0.40507  -0.31  0.7603
## Year2001        0.23324    0.34389   0.68  0.4984
## Year2002       -0.56143    0.33975  -1.65  0.1000
## Year2003       -0.21336    0.38478  -0.55  0.5799
## Year2004        0.12311    0.36912   0.33  0.7391
## Year2005        0.03535    0.37806   0.09  0.9256
## Year2006       -0.12632    0.35014  -0.36  0.7187
## Year2007        0.23502    0.34245   0.69  0.4933
## Year2008        0.00252    0.36223   0.01  0.9945
## Year2009       -0.19076    0.34721  -0.55  0.5833
## Year2010        0.05998    0.36445   0.16  0.8694
## Year2011       -0.25968    0.41029  -0.63  0.5275
## Year2012       -0.09265    0.34337  -0.27  0.7876
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.712
## Multiple R-squared:  0.115, Adjusted R-squared:  0.0396
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 201 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.315  0.891  0.940  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      4.61e-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: 217"
## [1] ""

```

```

## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1212"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    1    1    4    1    6    7    8    5   12   13   12    8    8   17
## 2012
##    16
##
## 1996 1997 1998 1999 2000 2001 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    0    0    1    3    1    5    6    6    5   10   11   10    6    8   16
## 2012
##    15
##
## 1996 1997 1998 1999 2000 2001 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    0    0    1    3    1    5    6    6    5   10   11    9    6    8   16
## 2012
##    15
## [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: 1.20302503608212"

## Warning in wilcox.test.default(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: 1.12246204830937"
## [1] "Male last author team size 2018 geometric mean: 1.16012938616016"

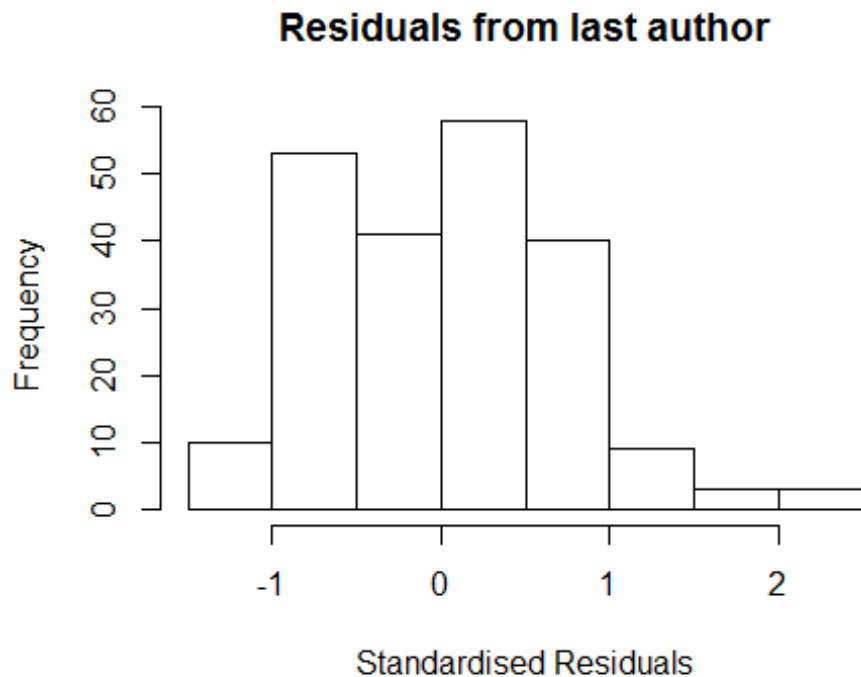
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 42, p-value = 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 1          NaN
## Year              NaN 13         NaN

## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 53 51249141774 3.218 2007    1211      3      2.515
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## 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 -7.02e-01  4.44e-16  4.84e-01  2.52e+00
##
```

```

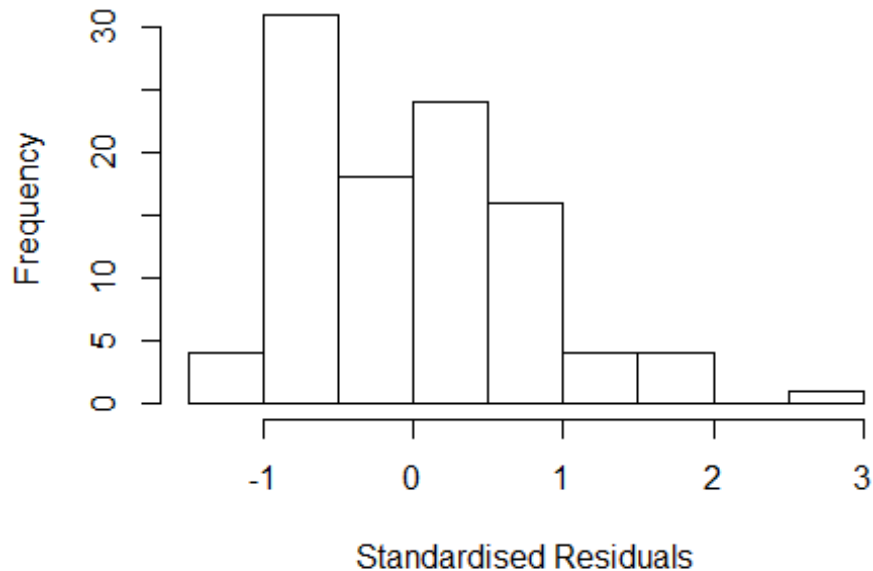
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.442      0.000    Inf < 2e-16 ***
## FirstAuthorFemale1 1.198      0.320     3.75 0.00033 ***
## LastAuthorFemale1 -0.872      0.000   -Inf < 2e-16 ***
## UniqueAuthors2    0.497      0.340     1.46 0.14694
## Year1999          -0.758      0.150    -5.04 2.6e-06 ***
## Year2000          -1.240      0.000   -Inf < 2e-16 ***
## Year2001          -1.469      0.292    -5.04 2.6e-06 ***
## Year2003          -1.725      0.444    -3.89 0.00020 ***
## Year2004          -0.707      0.294    -2.40 0.01857 *
## Year2005          -1.719      0.335    -5.14 1.7e-06 ***
## Year2006          -1.867      0.258    -7.23 1.9e-10 ***
## Year2007          -1.739      0.376    -4.63 1.3e-05 ***
## Year2008          -1.633      0.301    -5.43 5.3e-07 ***
## Year2009          -1.549      0.161    -9.62 3.0e-15 ***
## Year2010          -1.375      0.244    -5.64 2.2e-07 ***
## Year2011          -1.679      0.186    -9.05 4.2e-14 ***
## Year2012          -1.737      0.218    -7.99 6.1e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.801
## Multiple R-squared:  0.231, Adjusted R-squared:  0.0863
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 92 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.304  0.919  0.948  0.921  0.982  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      9.80e-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"

```

```
## 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 13         NaN

## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 53 51249141774 3.218 2007    1211      3      2.53
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.13319 -0.68785  0.00457  0.48620  2.53015
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)       2.442      0.000    Inf < 2e-16 ***
```

```

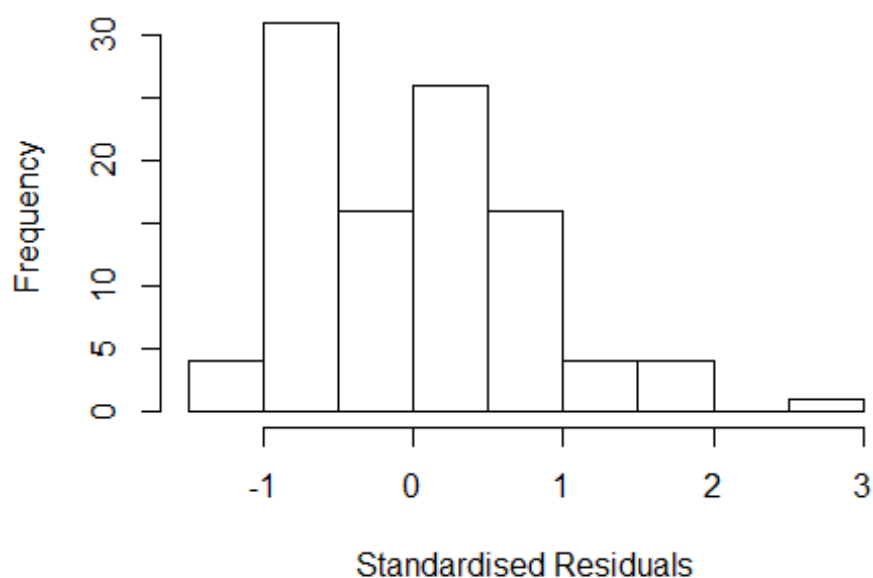
## FirstAuthorFemale1    1.629    0.234    6.96 6.3e-10 ***
## LastAuthorFemale1    -1.266    0.282   -4.49 2.2e-05 ***
## Year1999              -0.758    0.150   -5.04 2.5e-06 ***
## Year2000              -1.240    0.000   -Inf < 2e-16 ***
## Year2001              -1.476    0.295   -5.00 3.0e-06 ***
## Year2003              -1.728    0.436   -3.96 0.00015 ***
## Year2004              -0.705    0.295   -2.39 0.01898 *
## Year2005              -1.734    0.326   -5.32 8.1e-07 ***
## Year2006              -1.870    0.259   -7.21 2.0e-10 ***
## Year2007              -1.754    0.371   -4.72 8.9e-06 ***
## Year2008              -1.644    0.292   -5.64 2.1e-07 ***
## Year2009              -1.554    0.164   -9.45 5.8e-15 ***
## Year2010              -1.309    0.234   -5.59 2.6e-07 ***
## Year2011              -1.682    0.187   -8.97 5.7e-14 ***
## Year2012              -1.737    0.217   -8.00 5.4e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.807
## Multiple R-squared:  0.228, Adjusted R-squared:  0.0935
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~ = 1. The remaining 94 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.304  0.921  0.951  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           9.80e-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"

## 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 13            NaN

## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 53 51249141774 3.218 2007    1211      3      2.53
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q    Median        3Q        Max
## -1.26e+00 -6.44e-01  4.55e-15  4.98e-01  2.57e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.442      0.000    Inf < 2e-16 ***
## FirstAuthorFemale1  0.436      0.315    1.38  0.17096
## Year1999         -0.758      0.150   -5.04  2.5e-06 ***
## Year2000         -1.240      0.000   -Inf < 2e-16 ***
## Year2001         -1.489      0.305   -4.88  4.8e-06 ***
## Year2003         -1.739      0.429   -4.05  0.00011 ***
## Year2004         -0.709      0.294   -2.42  0.01774 *
## Year2005         -1.763      0.316   -5.58  2.7e-07 ***
```

```

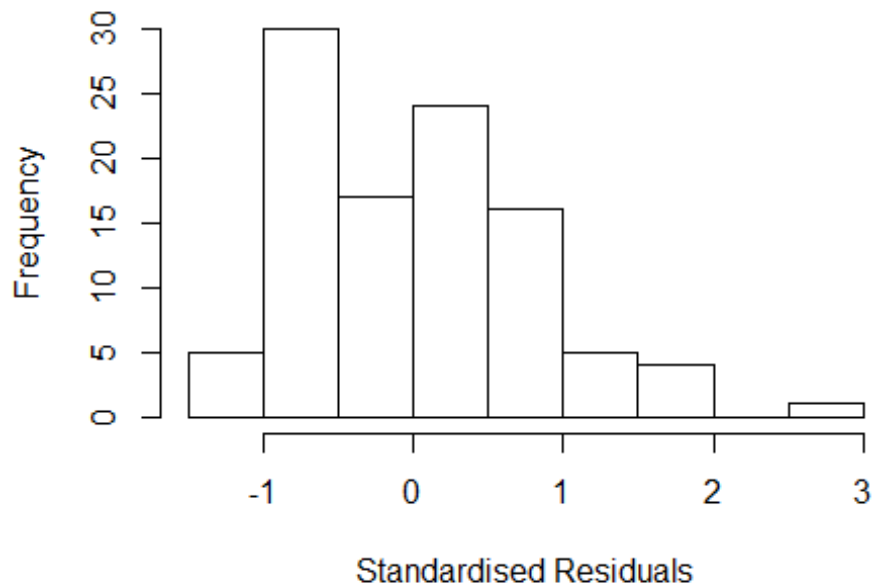
## Year2006          -1.877      0.264   -7.10  3.2e-10 ***
## Year2007          -1.798      0.367   -4.90  4.5e-06 ***
## Year2008          -1.668      0.300   -5.55  3.0e-07 ***
## Year2009          -1.564      0.172   -9.09  3.0e-14 ***
## Year2010          -1.185      0.260   -4.55  1.7e-05 ***
## Year2011          -1.687      0.192   -8.78  1.3e-13 ***
## Year2012          -1.738      0.218   -7.96  5.9e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.791
## Multiple R-squared:  0.209, Adjusted R-squared:  0.0813
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~ = 1. The remaining 94 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.268  0.915  0.948  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      9.80e-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 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 13          NaN

## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 53 51249141774 3.218 2007    1211      3      2.53
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q    Median        3Q       Max
## -1.32e+00 -7.05e-01  8.88e-16  4.98e-01  2.50e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.442      0.000    Inf < 2e-16 ***
## LastAuthorFemale1  0.324      0.295    1.10  0.27522
## Year1999         -0.758      0.150   -5.04  2.5e-06 ***
## Year2000         -1.240      0.000  -Inf < 2e-16 ***
## Year2001         -1.468      0.290   -5.06  2.3e-06 ***
## Year2003         -1.721      0.438   -3.93  0.00017 ***
## Year2004         -0.702      0.295   -2.38  0.01942 *
## Year2005         -1.719      0.331   -5.20  1.3e-06 ***
```

```

## Year2006          -1.866      0.256   -7.28  1.4e-10 ***
## Year2007          -1.728      0.361   -4.78  7.0e-06 ***
## Year2008          -1.631      0.291   -5.61  2.4e-07 ***
## Year2009          -1.548      0.160   -9.66  2.0e-15 ***
## Year2010          -1.124      0.294   -3.82  0.00025 ***
## Year2011          -1.678      0.187   -8.98  5.0e-14 ***
## Year2012          -1.736      0.215   -8.05  3.8e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.821
## Multiple R-squared:  0.19,   Adjusted R-squared:  0.0593
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~ = 1. The remaining 93 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.332  0.923  0.951  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      9.80e-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: 102"
## [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]"
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    2    5    2    6   14   15    4    6    6   15    4    7   29   18
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    4    2    6    9   11    4    4    5   11    4    6   21   16
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012

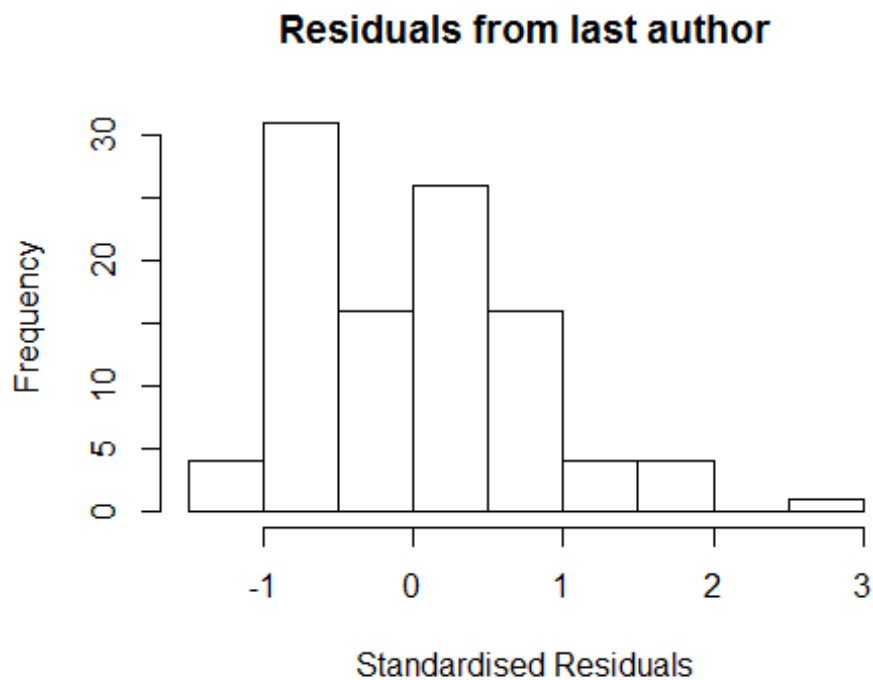
```

```
##      1      1      4      2      6      9     11      4      4      5     11      4      6     21     16
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.18920711500272"
## [1] "Male first author team size 2018 geometric mean: 1.16993081275869"

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

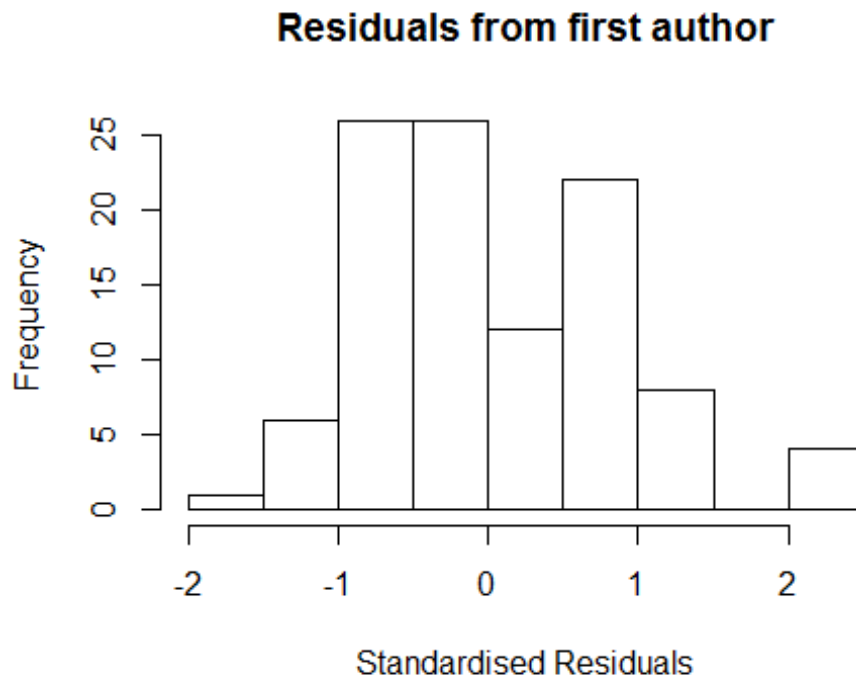
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 28, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.14720269043988"
## [1] "Male last author 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 = 27, 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"
## [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 1.325e+14  1      1.151e+07
## Year              1.325e+14 14      3.194e+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
## -1.5579 -0.6684 -0.0724  0.6126  2.2625
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -0.0692     0.1765   -0.39  0.69582
```

```

## FirstAuthorFemale1  0.0692      0.1765      0.39  0.69582
## Year1999            0.0692      0.1765      0.39  0.69582
## Year2000            0.6577      0.3476      1.89  0.06170 .
## Year2001            2.2005      0.6423      3.43  0.00093 ***
## Year2002            1.1445      0.6207      1.84  0.06854 .
## Year2003            1.3609      0.3946      3.45  0.00086 ***
## Year2004            1.4624      0.2754      5.31  8.0e-07 ***
## Year2005            0.5724      0.2248      2.55  0.01262 *
## Year2006            1.3444      0.3955      3.40  0.00101 **
## Year2007            1.1599      0.3780      3.07  0.00285 **
## Year2008            0.6686      0.2038      3.28  0.00148 **
## Year2009            1.6271      1.1659      1.40  0.16630
## Year2010            0.3153      0.1845      1.71  0.09092 .
## Year2011            0.8087      0.2088      3.87  0.00021 ***
## Year2012            0.9603      0.2203      4.36  3.5e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.947
## Multiple R-squared:  0.188, Adjusted R-squared:  0.0507
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 94 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.547  0.921  0.954  0.929  0.982  0.999
## Algorithmic parameters:
##           tuning.chi           bb           tuning.psi           refine.tol
##           1.55e+00           5.00e-01           4.69e+00           1.00e-07
##           rel.tol           solve.tol           eps.outlier           eps.x
##           1.00e-07           1.00e-07           9.52e-04           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01           5.00e-01
##   nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500          50          2          1          1000          200
##   trace.lev      mts      compute.rd
##           0          1000          0
##           psi           subsampling           cov
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.801e+15  1          4.244e+07
## Year            1.801e+15 14          3.506e+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
## -1.5579 -0.6684 -0.0724  0.6126  2.2625
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -0.0692    0.1765   -0.39  0.69582
## LastAuthorFemale1  0.0692    0.1765    0.39  0.69582
## Year1999        0.0692    0.1765    0.39  0.69582
## Year2000        0.6577    0.3476    1.89  0.06170 .
## Year2001        2.2005    0.6423    3.43  0.00093 ***
## Year2002        1.1445    0.6207    1.84  0.06854 .
## Year2003        1.3609    0.3946    3.45  0.00086 ***
## Year2004        1.4624    0.2754    5.31  8.0e-07 ***
## Year2005        0.5724    0.2248    2.55  0.01262 *
## Year2006        1.3444    0.3955    3.40  0.00101 **
## Year2007        1.1599    0.3780    3.07  0.00285 **
## Year2008        0.6686    0.2038    3.28  0.00148 **
## Year2009        1.6271    1.1659    1.40  0.16630
## Year2010        0.3153    0.1845    1.71  0.09092 .
## Year2011        0.8087    0.2088    3.87  0.00021 ***
## Year2012        0.9603    0.2203    4.36  3.5e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.947
## Multiple R-squared:  0.188, Adjusted R-squared:  0.0507
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 94 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.547  0.921  0.954  0.929  0.982  0.999
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          9.52e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500            50          2            1            1000      200
##      trace.lev      mts      compute.rd
##      0              1000          0
```



```

##          psi          subsampling          cov
##          "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##          "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 105"
## [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
##    7    6   13   17   13    9    9    8   10    7   16   14   18   19   20
## 2011 2012
##   43   69
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    8    9   10    2    7    4    6    7   10    9   12   16   16
## 2011 2012
##   33   54
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    0    7    8    9    1    7    3    5    7    9    9   12   14   14
## 2011 2012
##   29   46
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.64245821947267"
## [1] "Male first author team size 2018 geometric mean: 4.08471483158629"

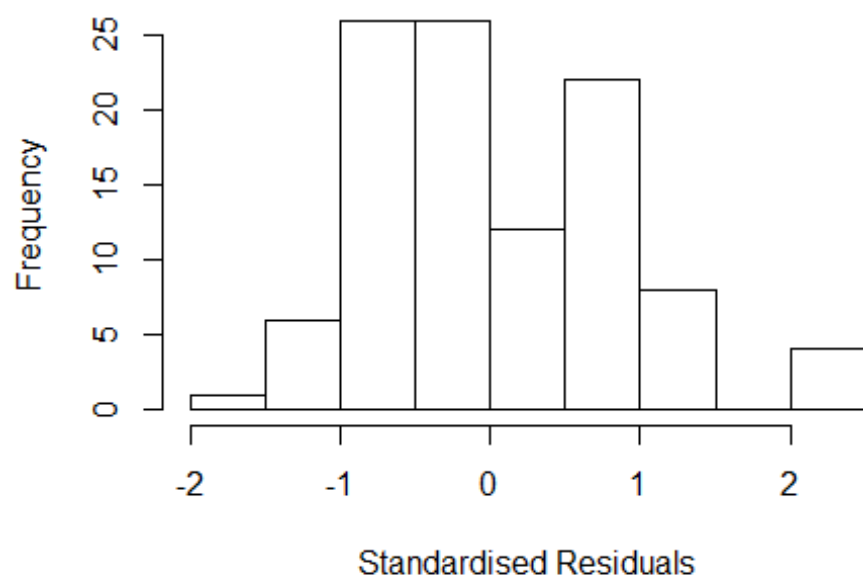
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 240, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.19364008301972"
## [1] "Male last author team size 2018 geometric mean: 4.72307558695382"

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

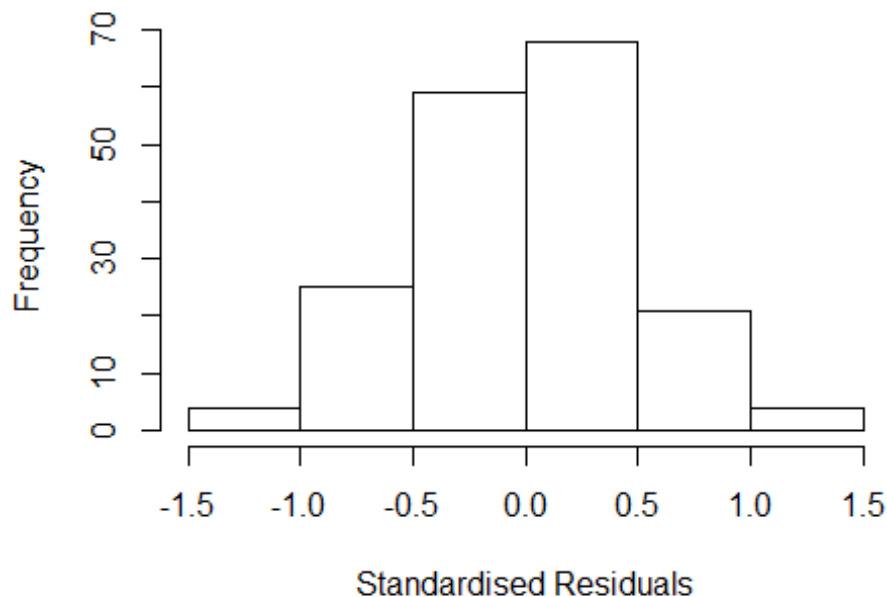
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 100, 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 1.392e+14  1      1.180e+07
## LastAuthorFemale  1.910e+14  1      1.382e+07
## UniqueAuthors    1.622e+16  4      1.062e+02
## Year              1.027e+17 15      3.690e+00
```

Residuals from first and last author and team size



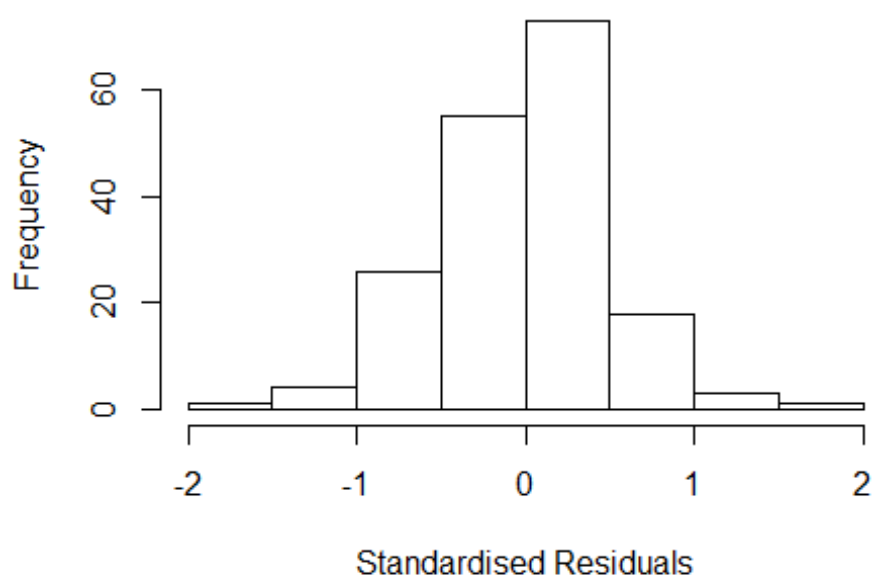
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3531 -0.3098 0.0234 0.2789 1.3922
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.57194 0.19649 2.91 0.0041 **
## FirstAuthorFemale1 0.17249 0.10299 1.67 0.0959 .
## LastAuthorFemale1 0.01656 0.12065 0.14 0.8910
## UniqueAuthors2 0.36808 0.20335 1.81 0.0722 .
## UniqueAuthors3 0.16616 0.20020 0.83 0.4078
## UniqueAuthors4 0.21504 0.18973 1.13 0.2587
## UniqueAuthors5 0.30698 0.19062 1.61 0.1093
## Year1998 0.36897 0.29349 1.26 0.2105
## Year1999 0.05422 0.29530 0.18 0.8545
## Year2000 0.39643 0.48280 0.82 0.4128
```

```

## Year2001      0.41190      0.33821      1.22      0.2251
## Year2002      0.11234      0.38070      0.30      0.7683
## Year2003      0.18694      0.31575      0.59      0.5546
## Year2004      0.52535      0.39792      1.32      0.1887
## Year2005      0.10066      0.30747      0.33      0.7438
## Year2006      0.40699      0.31875      1.28      0.2035
## Year2007     -0.00519      0.61127     -0.01      0.9932
## Year2008      0.20831      0.25784      0.81      0.4203
## Year2009      0.51128      0.30133      1.70      0.0917 .
## Year2010      0.20074      0.29852      0.67      0.5023
## Year2011      0.44814      0.27778      1.61      0.1087
## Year2012      0.28513      0.27779      1.03      0.3062
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.439
## Multiple R-squared:  0.159, Adjusted R-squared:  0.0479
## Convergence in 32 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 161 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.294  0.850  0.950  0.888  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      5.52e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -7.471e+14 1      NaN
## LastAuthorFemale -8.488e+14 1      NaN
## Year -3.446e+15 15      NaN

```

Residuals from first and last author



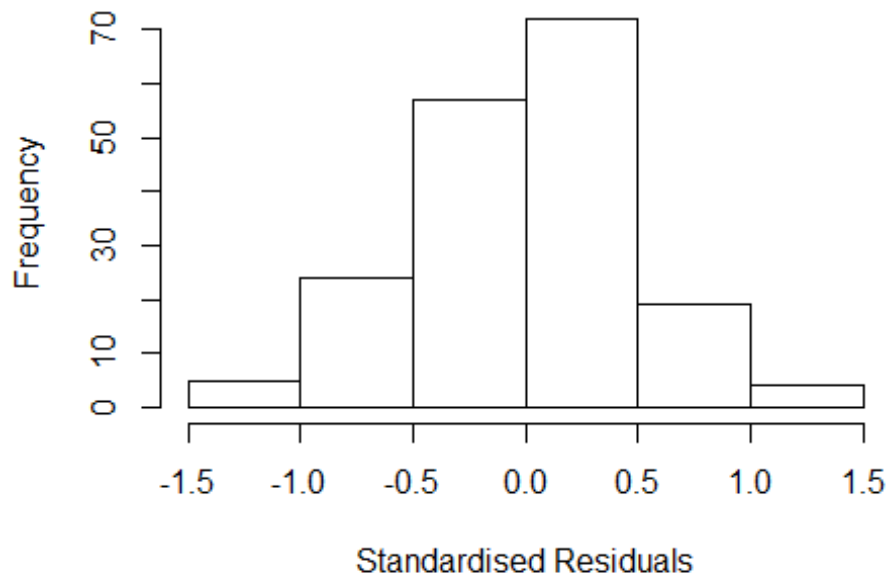
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5313 -0.3297 0.0142 0.2739 1.5183
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.460 0.168 2.73 0.00705 **
## FirstAuthorFemale1 0.233 0.097 2.41 0.01724 *
## LastAuthorFemale1 0.068 0.103 0.66 0.51173
## Year1998 0.652 0.162 4.02 8.7e-05 ***
## Year1999 0.391 0.168 2.33 0.02102 *
## Year2000 0.629 0.401 1.57 0.11847
## Year2001 0.690 0.168 4.10 6.5e-05 ***
## Year2002 0.466 0.331 1.41 0.16134
## Year2003 0.596 0.179 3.32 0.00110 **
## Year2004 0.828 0.329 2.52 0.01286 *
## Year2005 0.450 0.165 2.72 0.00722 **
## Year2006 0.732 0.194 3.77 0.00023 ***
```

```

## Year2007          0.405      0.529      0.77  0.44509
## Year2008          0.515      0.146      3.52  0.00055 ***
## Year2009          0.807      0.187      4.32  2.7e-05 ***
## Year2010          0.548      0.183      2.99  0.00322 **
## Year2011          0.770      0.141      5.45  1.9e-07 ***
## Year2012          0.619      0.142      4.36  2.3e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.43
## Multiple R-squared:  0.124, Adjusted R-squared:  0.0323
## Convergence in 29 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 162 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.179  0.840   0.946   0.879   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      5.52e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 8.085e+11 1      8.992e+05
## Year              8.085e+11 15      2.494e+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
## -1.472 -0.325 0.043 0.274 1.497
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5317 0.0951 5.59 9.3e-08 ***
## FirstAuthorFemale1 0.2293 0.0951 2.41 0.0170 *
## Year1998 0.5945 0.1247 4.77 4.1e-06 ***
## Year1999 0.3391 0.1337 2.54 0.0121 *
## Year2000 0.5816 0.3844 1.51 0.1322
## Year2001 0.6183 0.0951 6.50 9.2e-10 ***
## Year2002 0.3983 0.3104 1.28 0.2012
## Year2003 0.5243 0.1134 4.62 7.6e-06 ***
## Year2004 0.7555 0.3000 2.52 0.0128 *
## Year2005 0.3896 0.1312 2.97 0.0034 **
## Year2006 0.6804 0.1599 4.26 3.5e-05 ***
## Year2007 0.3550 0.5747 0.62 0.5376
```

```

## Year2008          0.4668      0.1077      4.33  2.6e-05 ***
## Year2009          0.7370      0.1270      5.80  3.3e-08 ***
## Year2010          0.4833      0.1471      3.28   0.0012 **
## Year2011          0.7108      0.0894      7.95  2.8e-13 ***
## Year2012          0.5688      0.1058      5.37  2.6e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.423
## Multiple R-squared:  0.12,   Adjusted R-squared:  0.0339
## Convergence in 30 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 167 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.185  0.839  0.948  0.879  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      5.52e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.068e+12  1      1.033e+06
## Year              1.068e+12 15      2.517e+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
## -1.3908 -0.2904  0.0387  0.2948  1.7428
##

```



```

## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.7228    0.0979    7.39 7.3e-12 ***
## LastAuthorFemale1 0.0382    0.0979    0.39 0.69719
## Year1998        0.5091    0.1378    3.70 0.00030 ***
## Year1999        0.1982    0.1490    1.33 0.18519
## Year2000        0.3744    0.3231    1.16 0.24822
## Year2001        0.4272    0.0979    4.36 2.2e-05 ***
## Year2002        0.3313    0.4063    0.82 0.41598
## Year2003        0.3331    0.1157    2.88 0.00451 **
## Year2004        0.5661    0.2973    1.90 0.05864 .
## Year2005        0.2642    0.1412    1.87 0.06317 .
## Year2006        0.5311    0.1392    3.81 0.00019 ***
## Year2007        0.3482    0.5022    0.69 0.48911
## Year2008        0.3563    0.1159    3.07 0.00249 **
## Year2009        0.6035    0.1247    4.84 3.0e-06 ***
## Year2010        0.3514    0.1495    2.35 0.01998 *
## Year2011        0.6298    0.1265    4.98 1.6e-06 ***
## Year2012        0.4470    0.1102    4.06 7.7e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.445
## Multiple R-squared:  0.0663, Adjusted R-squared:  -0.0248
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 163 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0916 0.8470 0.9400 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      5.52e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 181"
## [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]"
##
## 2008 2011 2012
##    1    2    1
##
## 2008 2011 2012
##    1    1    1
##
## 2008 2011 2012
##    1    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: 3"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1302"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    3    3    2    2    1    2    2    3    2    2    3    6    7    5
## 2011 2012
##    8    12
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    1    2    1    0    0    2    2    1    2    2    4    4    4
## 2011 2012
##    5    11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    1    2    1    0    0    2    1    1    2    2    4    4    4
## 2011 2012
##    5    6
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 5.87515877741196"
## [1] "Male first author team size 2018 geometric mean: 5.24148278841779"

```

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

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 10, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6.29970393478685"
## [1] "Male last author team size 2018 geometric mean: 5.02973371873174"

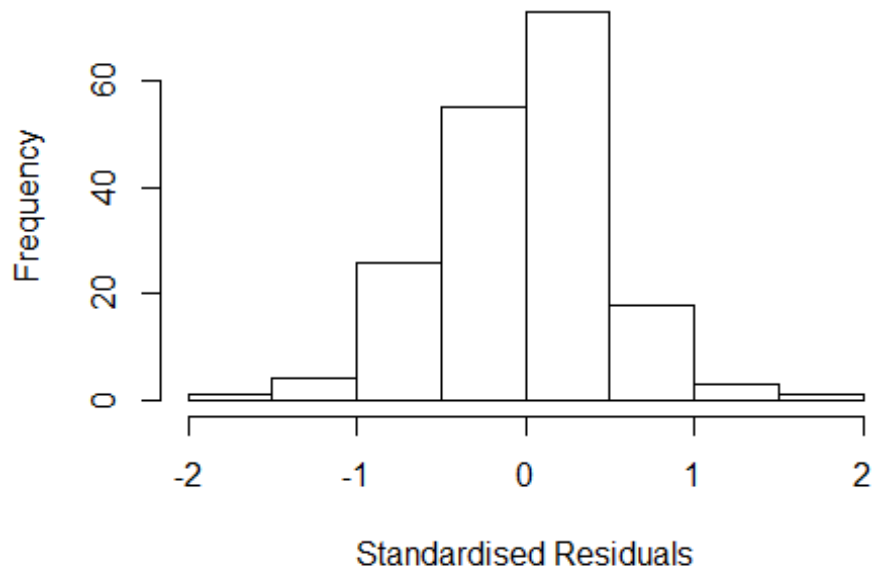
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 12, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
## [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 14            NaN

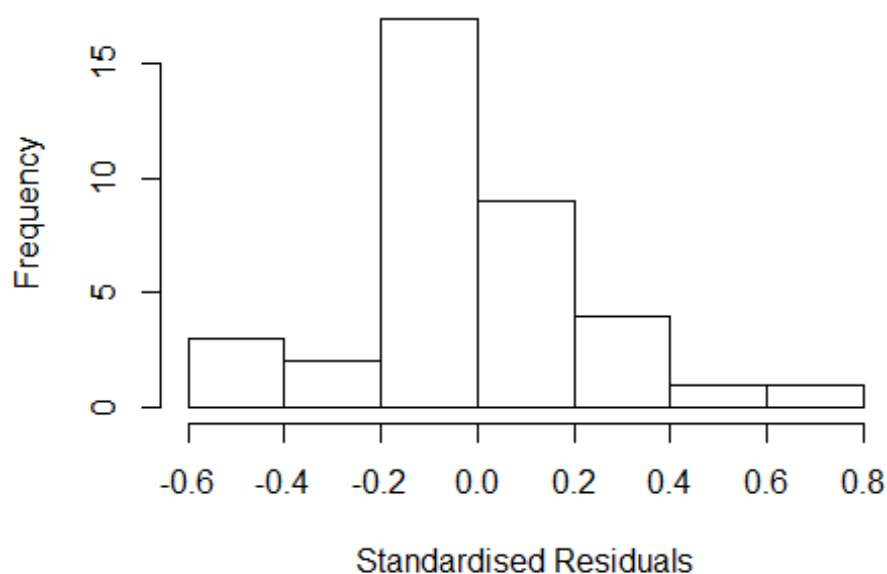
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q    Median        3Q       Max
## -5.26e-01 -9.92e-02 -4.44e-16  1.08e-01  6.43e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.8770     0.0000      Inf < 2e-16 ***
## FirstAuthorFemale1 -0.2135     0.1130    -1.89  0.07348 .
## LastAuthorFemale1 -0.0843     0.1353    -0.62  0.54038
## Year1997         0.4755     0.1130     4.21  0.00043 ***
## Year1998         0.2005     0.1130     1.77  0.09129 .
## Year1999         0.5120     0.6103     0.84  0.41142
## Year2000         0.7105     0.1130     6.29  3.9e-06 ***
```

```

## Year2003          0.6657      0.1335      4.99 7.1e-05 ***
## Year2004          0.5613      0.1353      4.15 0.00050 ***
## Year2005          0.4980      0.0000      Inf < 2e-16 ***
## Year2006         -0.3073      0.0811     -3.79 0.00115 **
## Year2007          0.6586      0.2123      3.10 0.00562 **
## Year2008          0.7401      0.3066      2.41 0.02550 *
## Year2009          0.6295      0.1782      3.53 0.00209 **
## Year2010          0.1678      0.1974      0.85 0.40536
## Year2011          0.4058      0.1590      2.55 0.01899 *
## Year2012          0.7639      0.1419      5.38 2.9e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.317
## Multiple R-squared:  0.582, Adjusted R-squared:  0.247
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 28 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.660  0.905   0.976   0.933   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      2.70e-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"
## 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

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4.95e-01 -1.25e-01 -3.33e-16  1.40e-01  6.94e-01
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    8.77e-01   0.00e+00      Inf < 2e-16 ***
## FirstAuthorFemale1 -1.90e-01   1.05e-01  -1.81e+00  0.08466 .
## Year1997         4.52e-01   1.05e-01   4.30e+00  0.00032 ***
## Year1998         1.77e-01   1.05e-01   1.69e+00  0.10655
## Year1999         5.12e-01   6.44e-01   8.00e-01  0.43546
## Year2000         6.87e-01   1.05e-01   6.54e+00  1.8e-06 ***
## Year2003         6.54e-01   1.24e-01   5.29e+00  3.0e-05 ***
## Year2004         4.77e-01   1.11e-08  4.28e+07  < 2e-16 ***
```

```

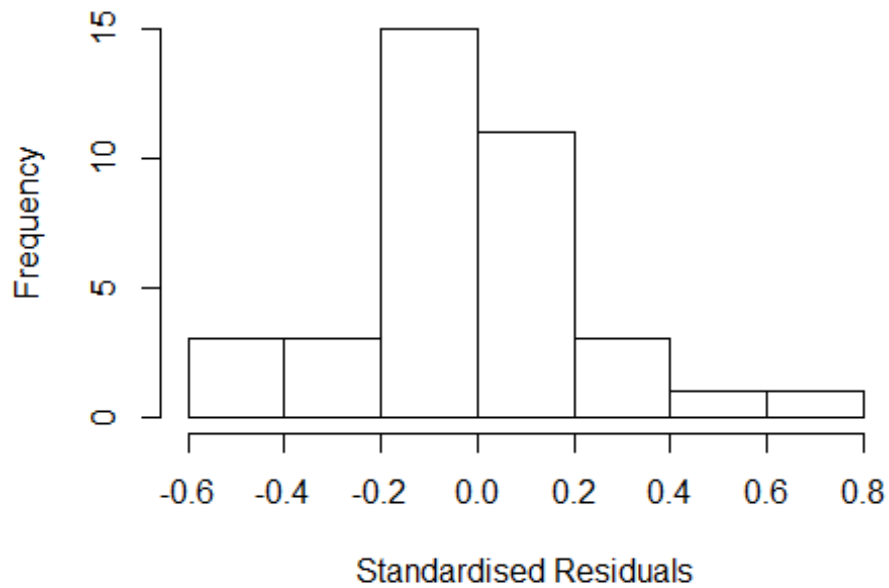
## Year2005          4.98e-01  0.00e+00      Inf < 2e-16 ***
## Year2006        -3.19e-01  8.50e-02 -3.75e+00  0.00117 **
## Year2007          5.93e-01  1.47e-01  4.03e+00  0.00060 ***
## Year2008          6.89e-01  3.20e-01  2.15e+00  0.04323 *
## Year2009          5.71e-01  1.22e-01  4.69e+00  0.00013 ***
## Year2010          1.37e-01  1.78e-01  7.70e-01  0.44885
## Year2011          3.24e-01  7.36e-02  4.41e+00  0.00025 ***
## Year2012          6.97e-01  9.14e-02  7.63e+00  1.8e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.308
## Multiple R-squared:  0.576, Adjusted R-squared:  0.272
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~ = 1. The remaining 30 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.591  0.910  0.978  0.933  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      2.70e-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

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

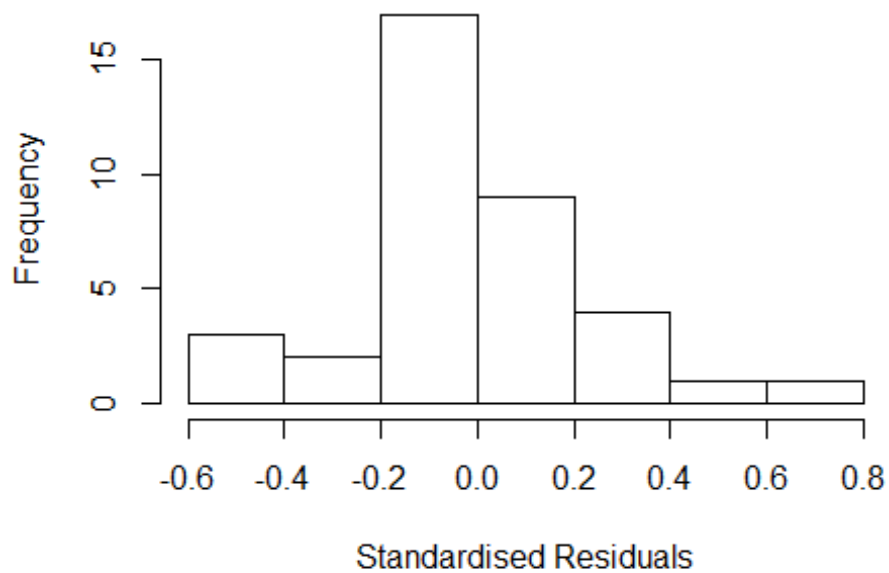
```

Residuals from first author



| | | | |
|---------------------|------|----|---------------------|
| ## | GVIF | Df | $GVIF^{(1/(2*Df))}$ |
| ## LastAuthorFemale | NaN | 1 | NaN |
| ## Year | NaN | 14 | NaN |

Residuals from last author

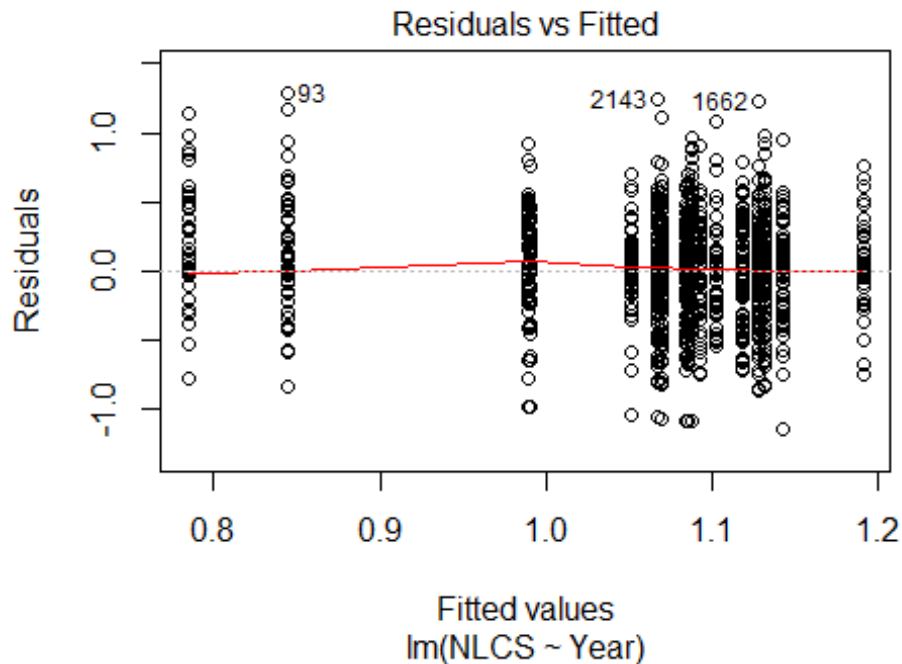



```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -5.30e-01 -6.61e-02 2.22e-16 1.44e-01 8.20e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 8.77e-01 9.23e-09 9.50e+07 < 2e-16 ***
## LastAuthorFemale1 -7.36e-03 1.30e-01 -6.00e-02 0.9553
## Year1997 2.62e-01 0.00e+00 Inf < 2e-16 ***
## Year1998 -1.30e-02 1.20e-08 -1.08e+06 < 2e-16 ***
## Year1999 5.12e-01 6.49e-01 7.90e-01 0.4390
## Year2000 4.97e-01 1.54e-08 3.22e+07 < 2e-16 ***
## Year2003 5.59e-01 3.98e-02 1.40e+01 3.9e-12 ***
## Year2004 4.84e-01 1.30e-01 3.73e+00 0.0012 **
## Year2005 4.98e-01 1.37e-08 3.64e+07 < 2e-16 ***
## Year2006 -4.14e-01 1.43e-01 -2.90e+00 0.0086 **
## Year2007 4.07e-01 1.24e-01 3.28e+00 0.0036 **
## Year2008 5.63e-01 5.31e-01 1.06e+00 0.3008
## Year2009 4.96e-01 1.77e-01 2.80e+00 0.0108 *
## Year2010 3.33e-02 1.35e-01 2.50e-01 0.8077
## Year2011 2.17e-01 1.11e-01 1.96e+00 0.0639 .
## Year2012 6.14e-01 1.14e-01 5.39e+00 2.4e-05 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.307
## Multiple R-squared: 0.538, Adjusted R-squared: 0.208
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 28 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.456 0.905 0.966 0.919 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 2.70e-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 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
##   85   59   74   54   65   74   86   74   73   84  103  115  125  131  141
## 2011 2012
##  150  146
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   61   40   54   43   41   44   61   56   52   63   76   83  105  108  117
## 2011 2012
##  118  110
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   48   36   43   36   35   33   49   46   46   53   69   72   83   96  101
## 2011 2012
##  101   95
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 76, df = 16, p-value = 1e-09

```



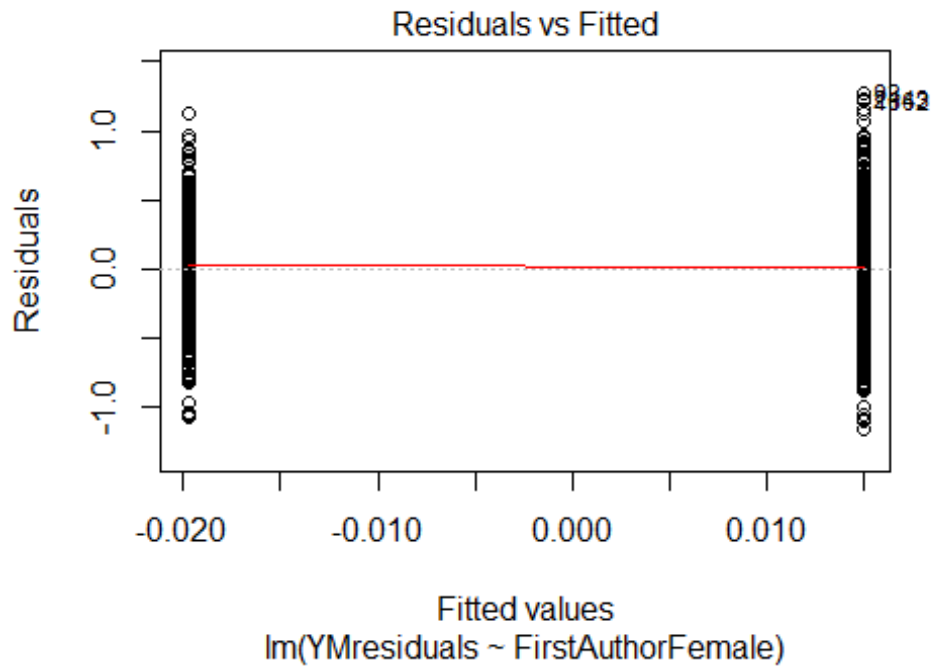
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 4.4, df = 1, p-value = 0.04

## [1] "Female first author team size 2018 geometric mean: 4.61444180067921"
## [1] "Male first author team size 2018 geometric mean: 3.22037430561774"

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

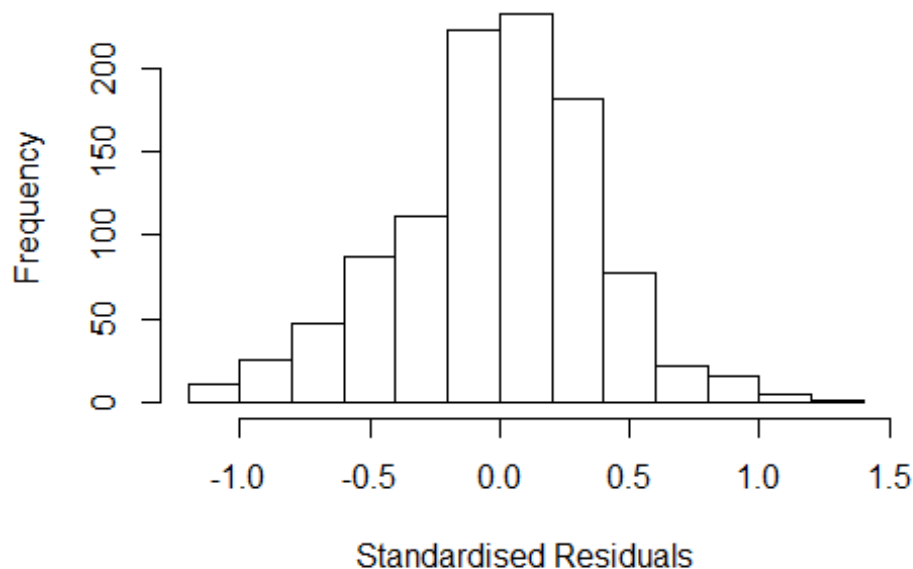
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 230, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.15716666002114"
## [1] "Male last author team size 2018 geometric mean: 3.27114686803443"

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



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 200, p-value = 0.01
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.170  1      1.082
## LastAuthorFemale  1.168  1      1.081
## UniqueAuthors    2.038  4      1.093
## Year              2.298 16      1.026
```

Residuals from first and last author and team size



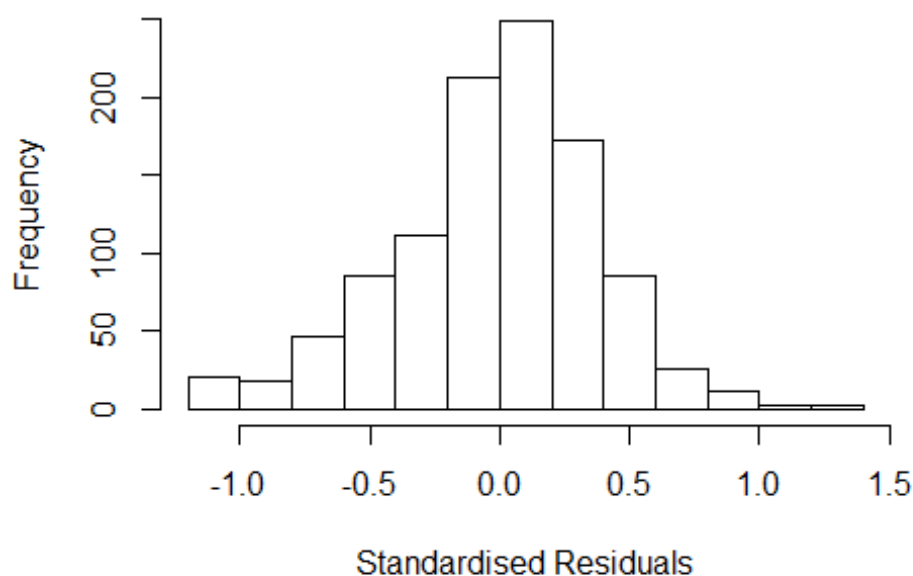
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1661 -0.2323 0.0149 0.2348 1.2846
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.63603 0.17397 3.66 0.00027 ***
## FirstAuthorFemale1 -0.04761 0.02493 -1.91 0.05648 .
## LastAuthorFemale1 -0.00655 0.02964 -0.22 0.82507
## UniqueAuthors2 0.20942 0.16729 1.25 0.21092
## UniqueAuthors3 0.22916 0.16544 1.39 0.16632
## UniqueAuthors4 0.25857 0.16451 1.57 0.11632
## UniqueAuthors5 0.30498 0.16304 1.87 0.06169 .
## Year1997 0.29655 0.18783 1.58 0.11470
## Year1998 -0.03609 0.12559 -0.29 0.77390
## Year1999 0.18356 0.10464 1.75 0.07968 .
```

```

## Year2000      0.22511      0.13266      1.70  0.09003 .
## Year2001      0.34628      0.10164      3.41  0.00068 ***
## Year2002      0.25581      0.09923      2.58  0.01008 *
## Year2003      0.22757      0.09411      2.42  0.01578 *
## Year2004      0.22830      0.09640      2.37  0.01805 *
## Year2005      0.25635      0.09938      2.58  0.01003 *
## Year2006      0.28516      0.09366      3.04  0.00239 **
## Year2007      0.25665      0.09429      2.72  0.00660 **
## Year2008      0.23959      0.09044      2.65  0.00820 **
## Year2009      0.25702      0.09063      2.84  0.00466 **
## Year2010      0.21122      0.09161      2.31  0.02133 *
## Year2011      0.17829      0.09091      1.96  0.05014 .
## Year2012      0.28731      0.09357      3.07  0.00219 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.35
## Multiple R-squared:  0.0743, Adjusted R-squared:  0.0543
## Convergence in 30 IRWLS iterations
##
## Robustness weights:
## 91 weights are ~= 1. The remaining 951 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.149  0.853  0.948  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          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 1.124 1          1.060
## LastAuthorFemale  1.131 1          1.063
## Year              1.203 16          1.006

```

Residuals from first and last author



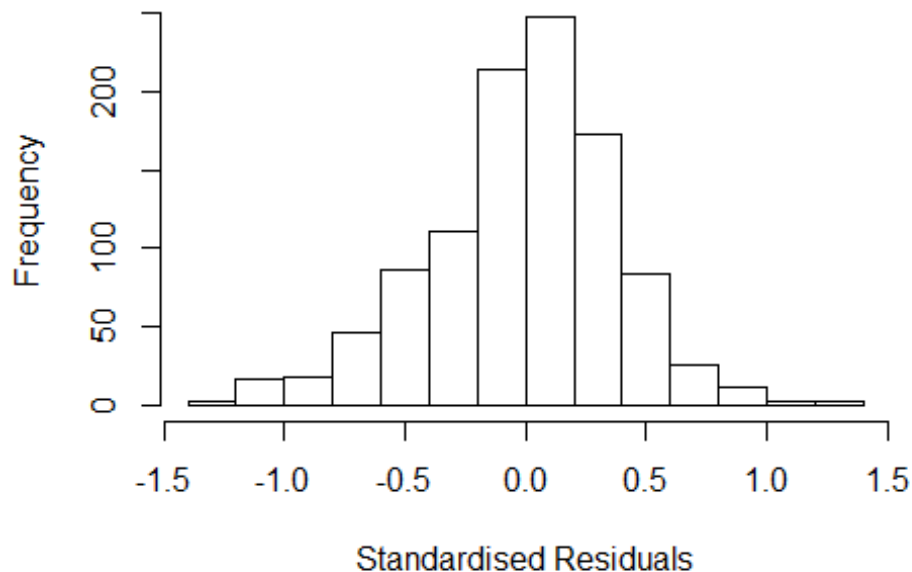
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.197 -0.239  0.015  0.240  1.275
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.8546    0.0857   9.97 < 2e-16 ***
## FirstAuthorFemale1 -0.0394    0.0246  -1.60  0.10917
## LastAuthorFemale1 -0.0041    0.0295  -0.14  0.88941
## Year1997         0.3427    0.1824   1.88  0.06050 .
## Year1998        -0.0382    0.1240  -0.31  0.75800
## Year1999         0.1913    0.1023   1.87  0.06189 .
## Year2000         0.2352    0.1322   1.78  0.07560 .
## Year2001         0.3562    0.1001   3.56  0.00039 ***
## Year2002         0.2656    0.1004   2.64  0.00831 **
## Year2003         0.2365    0.0943   2.51  0.01231 *
## Year2004         0.2600    0.0974   2.67  0.00772 **
## Year2005         0.2787    0.0995   2.80  0.00519 **
```

```

## Year2006          0.3153      0.0936      3.37  0.00078 ***
## Year2007          0.2835      0.0943      3.01  0.00270 **
## Year2008          0.2700      0.0900      3.00  0.00277 **
## Year2009          0.2900      0.0903      3.21  0.00136 **
## Year2010          0.2474      0.0910      2.72  0.00668 **
## Year2011          0.2198      0.0905      2.43  0.01540 *
## Year2012          0.3222      0.0930      3.46  0.00056 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.357
## Multiple R-squared:  0.0563, Adjusted R-squared:  0.0397
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 90 weights are ~= 1. The remaining 952 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.176  0.859  0.949  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      9.60e-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 1.095 1          1.046
## Year              1.095 16          1.003

```


Residuals from first author



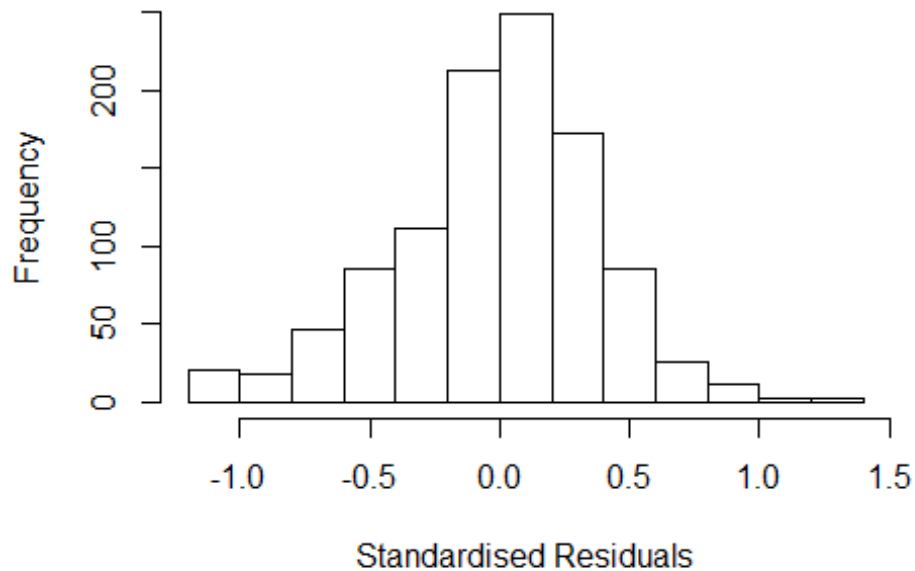
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2038 -0.2402  0.0149  0.2391  1.2762
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.8538    0.0859   9.94 < 2e-16 ***
## FirstAuthorFemale1 -0.0398    0.0243  -1.64  0.10178
## Year1997        0.3499    0.1847   1.89  0.05840 .
## Year1998       -0.0372    0.1254  -0.30  0.76647
## Year1999        0.1921    0.1029   1.87  0.06216 .
## Year2000        0.2397    0.1330   1.80  0.07182 .
## Year2001        0.3561    0.1006   3.54  0.00042 ***
## Year2002        0.2663    0.1012   2.63  0.00861 **
## Year2003        0.2371    0.0948   2.50  0.01252 *
## Year2004        0.2590    0.0980   2.64  0.00834 **
## Year2005        0.2794    0.1000   2.79  0.00529 **
## Year2006        0.3158    0.0940   3.36  0.00081 ***
```

```

## Year2007          0.2842      0.0946      3.00  0.00274 **
## Year2008          0.2706      0.0905      2.99  0.00286 **
## Year2009          0.2904      0.0907      3.20  0.00140 **
## Year2010          0.2477      0.0917      2.70  0.00705 **
## Year2011          0.2196      0.0913      2.41  0.01631 *
## Year2012          0.3222      0.0937      3.44  0.00061 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.351
## Multiple R-squared:  0.0569, Adjusted R-squared:  0.0413
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 90 weights are ~= 1. The remaining 952 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.159  0.856  0.948  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      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 1.113 1          1.055
## Year              1.113 16          1.003

```

Residuals from last author



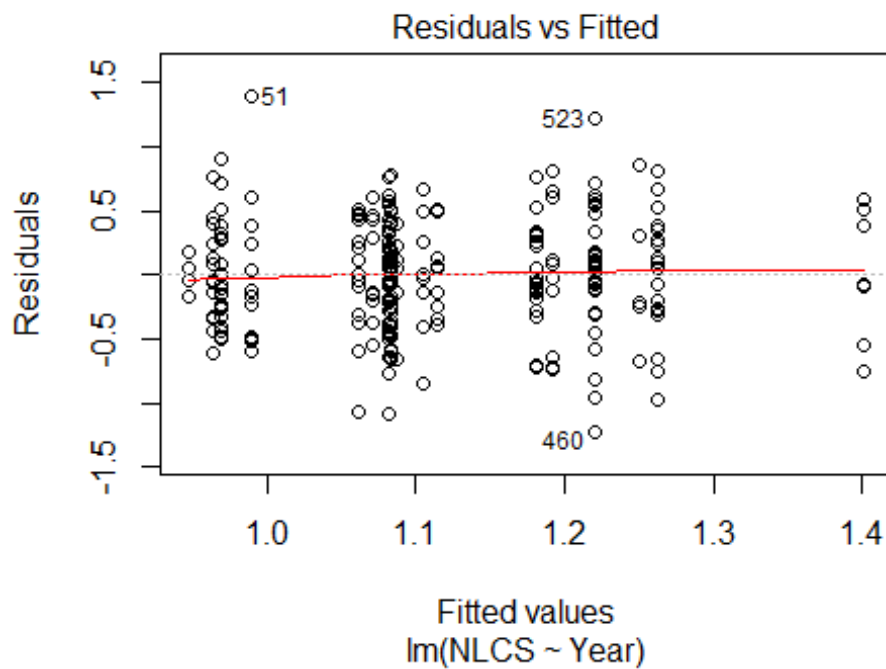
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1909 -0.2461 0.0208 0.2402 1.2930
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8370 0.0865 9.67 < 2e-16 ***
## LastAuthorFemale1 -0.0106 0.0293 -0.36 0.71665
## Year1997 0.3539 0.1878 1.88 0.05984 .
## Year1998 -0.0332 0.1264 -0.26 0.79274
## Year1999 0.1912 0.1040 1.84 0.06633 .
## Year2000 0.2430 0.1321 1.84 0.06620 .
## Year2001 0.3588 0.1016 3.53 0.00043 ***
## Year2002 0.2721 0.1024 2.66 0.00799 **
## Year2003 0.2334 0.0960 2.43 0.01525 *
## Year2004 0.2650 0.0991 2.67 0.00763 **
## Year2005 0.2791 0.1016 2.75 0.00613 **
## Year2006 0.3201 0.0954 3.36 0.00082 ***
```

```

## Year2007          0.2826      0.0960      2.94  0.00333 **
## Year2008          0.2688      0.0918      2.93  0.00348 **
## Year2009          0.2917      0.0921      3.17  0.00159 **
## Year2010          0.2505      0.0928      2.70  0.00704 **
## Year2011          0.2206      0.0924      2.39  0.01707 *
## Year2012          0.3241      0.0948      3.42  0.00065 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.352
## Multiple R-squared:  0.0544, Adjusted R-squared:  0.0387
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 92 weights are ~= 1. The remaining 950 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.148  0.856  0.947  0.888  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+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 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
##   15   10   19   14   16   13   14   25   18   20   38   31   36   36   51
## 2011 2012
##   48   55
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    8    7   12    5    5    6   10   10   10   14   18   18   22   23   36
## 2011 2012

```

```
## 37 42
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 7 7 9 5 5 6 7 5 7 12 15 16 19 15 30
## 2011 2012
## 26 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 = 19, df = 16, p-value = 0.3
```



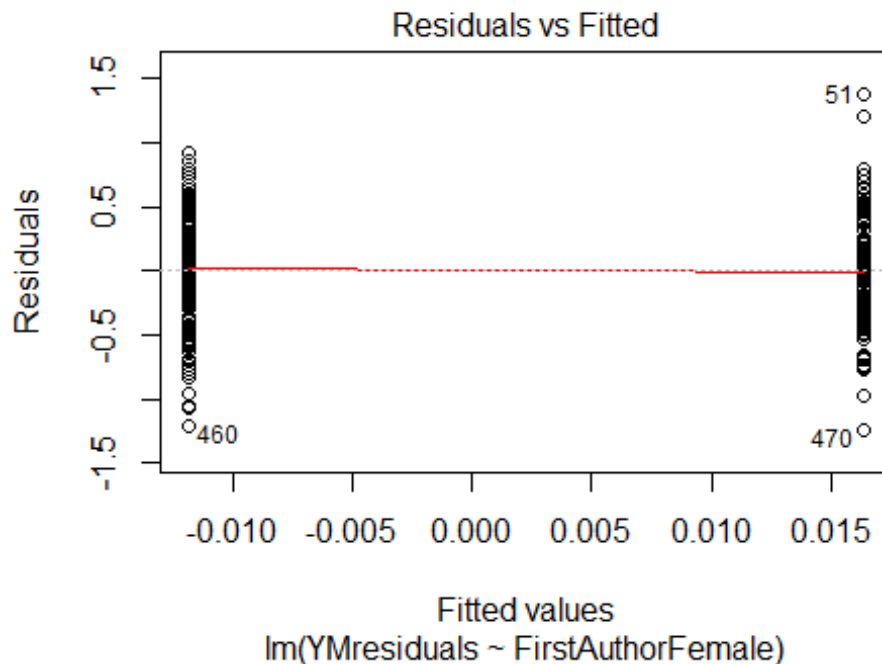
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.18, df = 1, p-value = 0.7

## [1] "Female first author team size 2018 geometric mean: 3.77762959804593"
## [1] "Male first author team size 2018 geometric mean: 4.10714661365223"

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

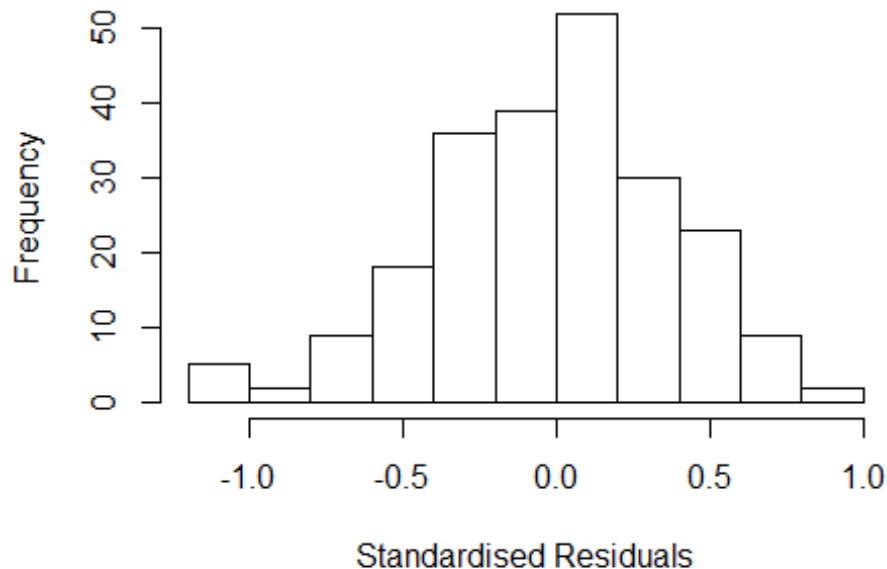
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 19, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.33788708479079"
## [1] "Male last author team size 2018 geometric mean: 3.69885428168899"

## Warning in wilcox.test.default(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 = 0.5
## 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 1.228 1      1.108
## LastAuthorFemale  2.015 1      1.420
## UniqueAuthors    4.654 4      1.212
## Year              7.285 16     1.064
```

Residuals from first and last author and team size



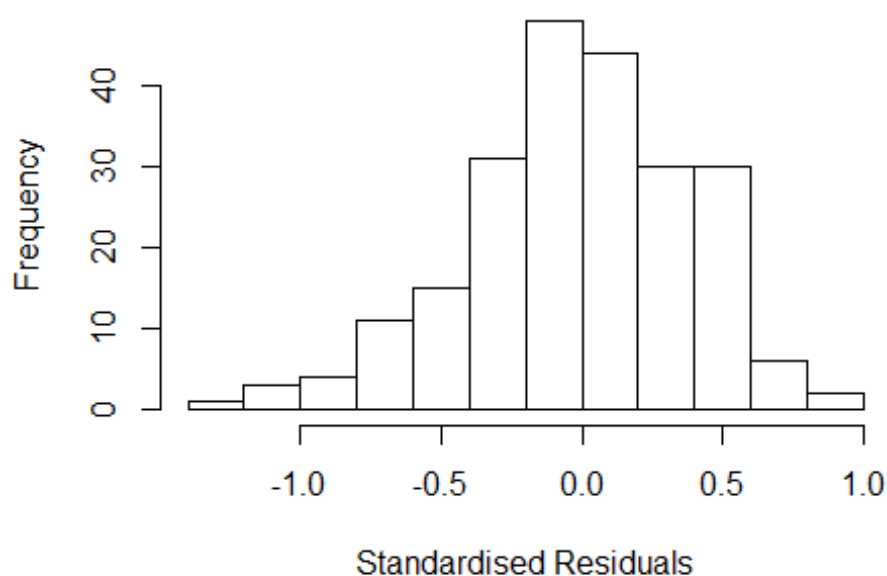
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.153 -0.268 0.017 0.259 0.962
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.23276 0.18958 6.50 6.1e-10 ***
## FirstAuthorFemale1 -0.02454 0.05653 -0.43 0.665
## LastAuthorFemale1 -0.03763 0.07474 -0.50 0.615
## UniqueAuthors2 -0.03411 0.14212 -0.24 0.811
## UniqueAuthors3 -0.03938 0.12959 -0.30 0.762
## UniqueAuthors4 0.12691 0.14071 0.90 0.368
## UniqueAuthors5 0.08182 0.12841 0.64 0.525
## Year1997 0.22918 0.31612 0.72 0.469
## Year1998 -0.35961 0.20837 -1.73 0.086 .
## Year1999 -0.29268 0.21039 -1.39 0.166
```

```

## Year2000      -0.02435    0.32386   -0.08    0.940
## Year2001      -0.13973    0.23666   -0.59    0.556
## Year2002      -0.37626    0.25140   -1.50    0.136
## Year2003      -0.04103    0.23952   -0.17    0.864
## Year2004      -0.09286    0.22141   -0.42    0.675
## Year2005      -0.31304    0.22361   -1.40    0.163
## Year2006      -0.13264    0.22145   -0.60    0.550
## Year2007       0.07694    0.24329    0.32    0.752
## Year2008      -0.00949    0.21393   -0.04    0.965
## Year2009      -0.14881    0.21539   -0.69    0.490
## Year2010      -0.02082    0.20345   -0.10    0.919
## Year2011      -0.16161    0.20775   -0.78    0.438
## Year2012      -0.18839    0.20213   -0.93    0.352
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.398
## Multiple R-squared:  0.121, Adjusted R-squared:  0.025
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 206 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.382  0.880  0.953  0.910  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      4.44e-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 1.203 1      1.097
## LastAuthorFemale  1.657 1      1.287
## Year              1.920 16      1.021

```


Residuals from first and last author



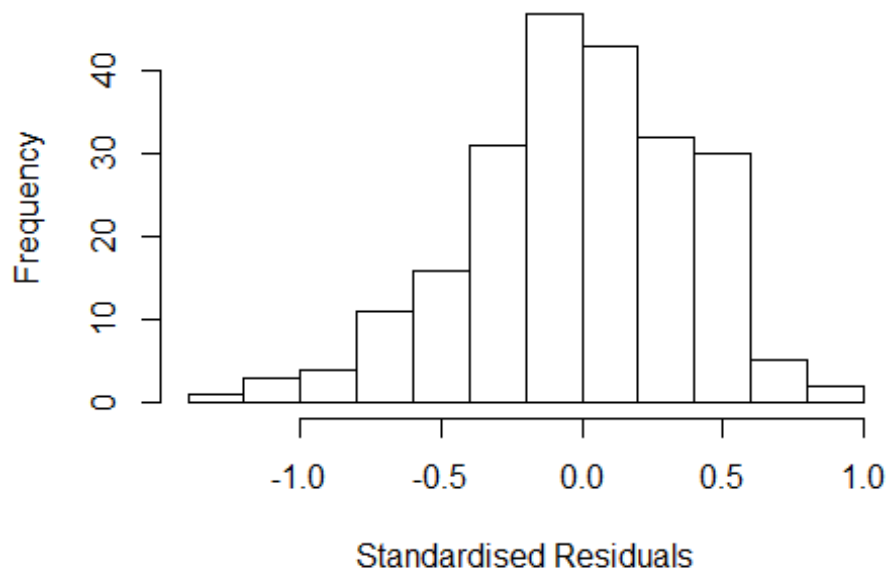
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2163 -0.2685 -0.0121  0.2482  0.9060
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.22499    0.18775   6.52 5.2e-10 ***
## FirstAuthorFemale1 -0.01797    0.05714  -0.31  0.753
## LastAuthorFemale1 -0.02193    0.06906  -0.32  0.751
## Year1997         0.20510    0.28628   0.72  0.475
## Year1998        -0.38629    0.20988  -1.84  0.067 .
## Year1999        -0.26732    0.19366  -1.38  0.169
## Year2000        -0.01004    0.33996  -0.03  0.976
## Year2001        -0.09797    0.22684  -0.43  0.666
## Year2002        -0.30544    0.23759  -1.29  0.200
## Year2003        -0.07723    0.22902  -0.34  0.736
## Year2004        -0.09194    0.21803  -0.42  0.674
## Year2005        -0.27826    0.21880  -1.27  0.205
```

```

## Year2006      -0.11170    0.20990   -0.53    0.595
## Year2007      0.11155    0.22762    0.49    0.625
## Year2008      0.02977    0.19753    0.15    0.880
## Year2009     -0.13953    0.21055   -0.66    0.508
## Year2010      0.00929    0.19226    0.05    0.961
## Year2011     -0.12387    0.19385   -0.64    0.524
## Year2012     -0.13863    0.19440   -0.71    0.477
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.412
## Multiple R-squared:  0.0967, Adjusted R-squared:  0.0178
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 208 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.364  0.889  0.957  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      4.44e-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 1.162 1      1.078
## Year              1.162 16      1.005

```

Residuals from first author



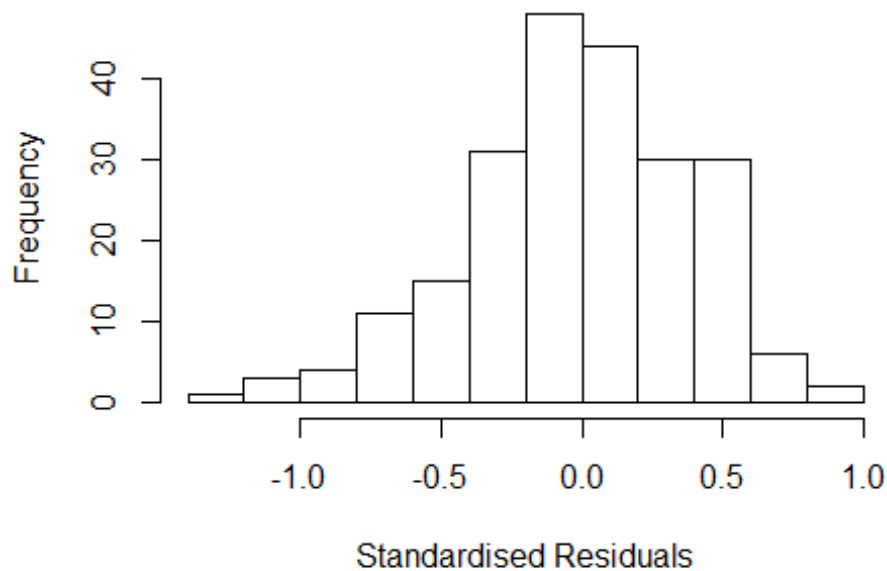
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.20881 -0.26890 -0.00541 0.24985 0.88486
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.21202 0.18031 6.72 1.7e-10 ***
## FirstAuthorFemale1 -0.02210 0.05628 -0.39 0.69
## Year1997 0.21406 0.28791 0.74 0.46
## Year1998 -0.38232 0.21029 -1.82 0.07 .
## Year1999 -0.25186 0.18425 -1.37 0.17
## Year2000 0.00213 0.32665 0.01 0.99
## Year2001 -0.09154 0.22318 -0.41 0.68
## Year2002 -0.29585 0.23378 -1.27 0.21
## Year2003 -0.06515 0.22503 -0.29 0.77
## Year2004 -0.09349 0.21973 -0.43 0.67
## Year2005 -0.26885 0.21381 -1.26 0.21
## Year2006 -0.10055 0.20540 -0.49 0.62
```

```

## Year2007      0.12088    0.22379    0.54    0.59
## Year2008      0.04114    0.19162    0.21    0.83
## Year2009     -0.12819    0.20635   -0.62    0.54
## Year2010      0.01889    0.19019    0.10    0.92
## Year2011     -0.11761    0.19379   -0.61    0.54
## Year2012     -0.12958    0.19157   -0.68    0.50
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.413
## Multiple R-squared:  0.096, Adjusted R-squared:  0.0218
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 209 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.371  0.887   0.958   0.915   0.986   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      4.44e-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 1.621 1      1.273
## Year      1.621 16      1.015

```

Residuals from last author



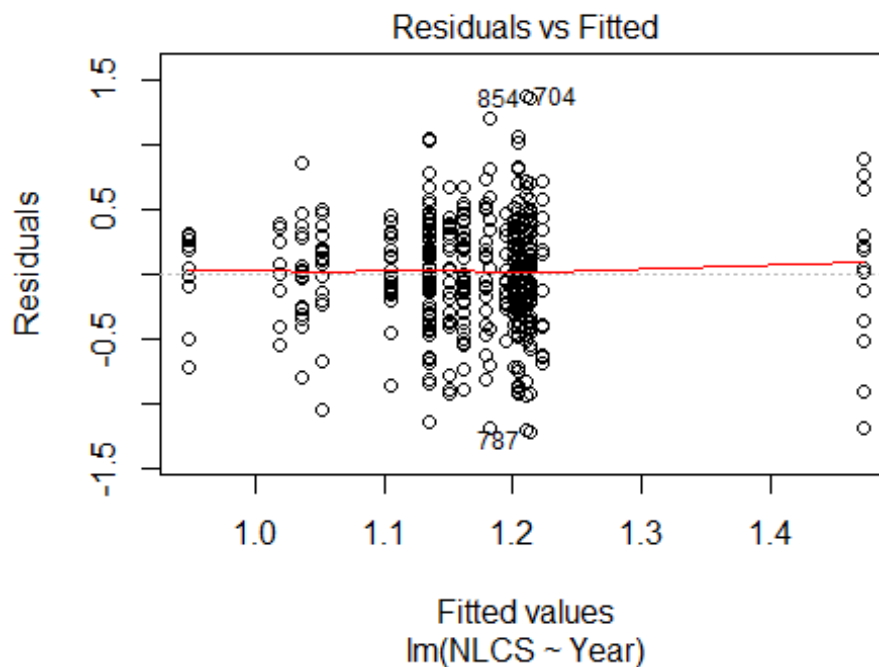
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.22734 -0.26266 -0.00581 0.25773 0.91889
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.21957 0.18554 6.57 3.9e-10 ***
## LastAuthorFemale1 -0.02645 0.06816 -0.39 0.698
## Year1997 0.20864 0.28410 0.73 0.464
## Year1998 -0.38568 0.20961 -1.84 0.067 .
## Year1999 -0.27271 0.19282 -1.41 0.159
## Year2000 -0.01301 0.34415 -0.04 0.970
## Year2001 -0.10146 0.22632 -0.45 0.654
## Year2002 -0.30914 0.23790 -1.30 0.195
## Year2003 -0.08491 0.22925 -0.37 0.711
## Year2004 -0.08831 0.21822 -0.40 0.686
## Year2005 -0.28319 0.21873 -1.29 0.197
## Year2006 -0.11157 0.20924 -0.53 0.594
```

```

## Year2007      0.11147      0.22744      0.49      0.625
## Year2008      0.02570      0.19749      0.13      0.897
## Year2009     -0.14071      0.21079     -0.67      0.505
## Year2010      0.00777      0.19194      0.04      0.968
## Year2011     -0.12476      0.19314     -0.65      0.519
## Year2012     -0.13542      0.19335     -0.70      0.484
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.411
## Multiple R-squared:  0.0969, Adjusted R-squared:  0.0227
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 209 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.351  0.889  0.957  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      4.44e-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: 225"
## [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
##   31   26   24   39   40   32   42   23   29   39   52   55   62   59   89
## 2011 2012
##  104   91
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   13   11   13   17   16    8   17   11   15   21   25   30   35   38   61
## 2011 2012

```

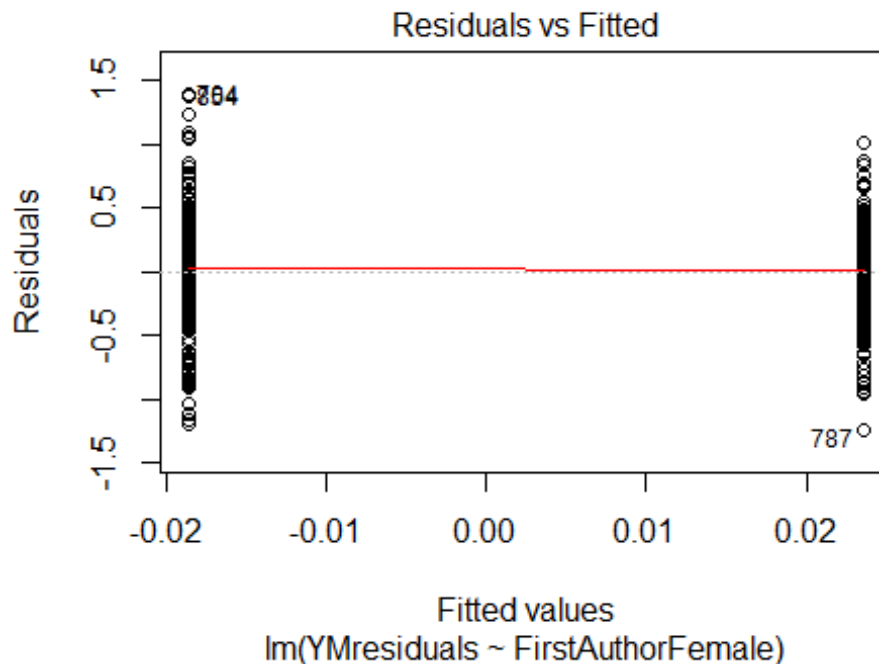
```
## 76 57
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 10 9 10 12 15 8 15 7 12 15 23 25 24 31 49
## 2011 2012
## 60 46
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 36, df = 16, p-value = 0.003
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 8.7, df = 1, p-value = 0.003
## [1] "Female first author team size 2018 geometric mean: 4.2450857048404"
## [1] "Male first author team size 2018 geometric mean: 3.02220266793812"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

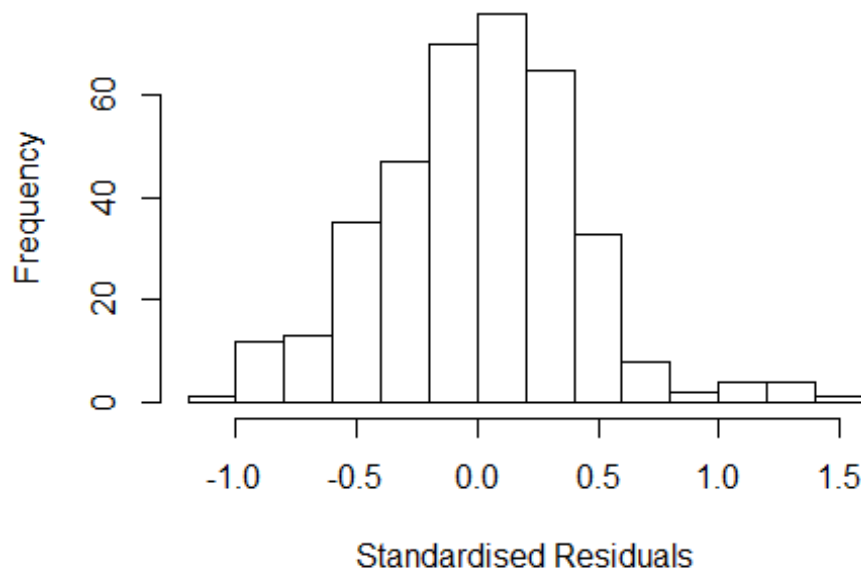
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 150, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.10987472271945"
## [1] "Male last author team size 2018 geometric mean: 3.53007118495345"

## Warning in wilcox.test.default(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 = 0.6
## 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 1.425  1      1.194
## LastAuthorFemale  1.296  1      1.138
## UniqueAuthors    2.899  4      1.142
## Year              3.729 16      1.042
```


Residuals from first and last author and team size



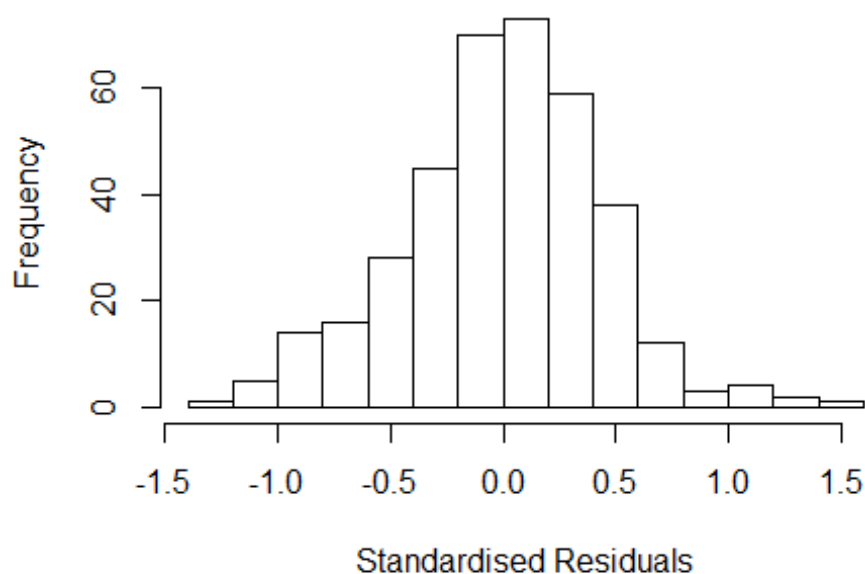
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1833 -0.2531 0.0206 0.2648 1.5251
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.88471 0.23800 3.72 0.00023 ***
## FirstAuthorFemale1 -0.02216 0.04611 -0.48 0.63116
## LastAuthorFemale1 0.00798 0.05386 0.15 0.88235
## UniqueAuthors2 0.56741 0.16317 3.48 0.00057 ***
## UniqueAuthors3 0.71686 0.16851 4.25 2.7e-05 ***
## UniqueAuthors4 0.68166 0.16595 4.11 5.0e-05 ***
## UniqueAuthors5 0.78401 0.15967 4.91 1.4e-06 ***
## Year1997 -0.29662 0.32567 -0.91 0.36303
## Year1998 -0.34877 0.25212 -1.38 0.16746
## Year1999 -0.63244 0.20481 -3.09 0.00218 **
```

```

## Year2000      -0.52440    0.21475   -2.44  0.01511 *
## Year2001      -0.49234    0.21349   -2.31  0.02169 *
## Year2002      -0.31972    0.22420   -1.43  0.15475
## Year2003      -0.70840    0.23593   -3.00  0.00287 **
## Year2004      -0.37044    0.19718   -1.88  0.06113 .
## Year2005      -0.42362    0.20262   -2.09  0.03728 *
## Year2006      -0.41389    0.20385   -2.03  0.04307 *
## Year2007      -0.27322    0.18774   -1.46  0.14648
## Year2008      -0.42787    0.19256   -2.22  0.02693 *
## Year2009      -0.41991    0.19201   -2.19  0.02941 *
## Year2010      -0.34213    0.18611   -1.84  0.06687 .
## Year2011      -0.39479    0.18690   -2.11  0.03537 *
## Year2012      -0.42628    0.19449   -2.19  0.02906 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.387
## Multiple R-squared:  0.185, Adjusted R-squared:  0.134
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 32 weights are ~= 1. The remaining 339 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0856 0.8720 0.9490 0.8980 0.9850 0.9990
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          2.70e-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 1.366 1          1.169
## LastAuthorFemale 1.237 1          1.112
## Year      1.651 16          1.016

```

Residuals from first and last author



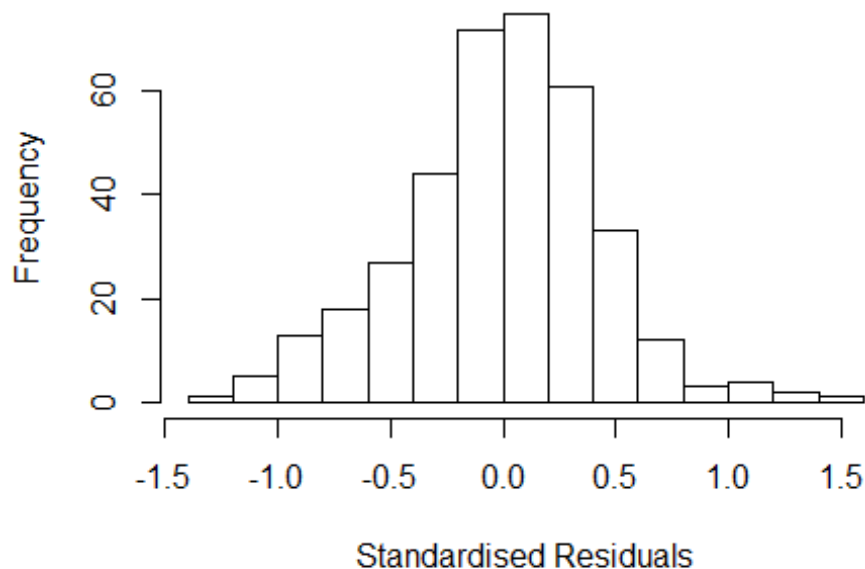
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2813 -0.2542 0.0207 0.2729 1.4770
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4873 0.1746 8.52 4.9e-16 ***
## FirstAuthorFemale1 0.0267 0.0481 0.56 0.5787
## LastAuthorFemale1 0.0372 0.0545 0.68 0.4959
## Year1997 -0.3801 0.6041 -0.63 0.5297
## Year1998 -0.3137 0.2678 -1.17 0.2422
## Year1999 -0.5811 0.2021 -2.88 0.0043 **
## Year2000 -0.5093 0.2034 -2.50 0.0127 *
## Year2001 -0.4807 0.2138 -2.25 0.0251 *
## Year2002 -0.3524 0.2178 -1.62 0.1065
## Year2003 -0.6739 0.2250 -3.00 0.0029 **
## Year2004 -0.3219 0.1912 -1.68 0.0932 .
## Year2005 -0.3419 0.1878 -1.82 0.0695 .
```

```

## Year2006          -0.3238      0.2009   -1.61    0.1079
## Year2007          -0.2703      0.1853   -1.46    0.1454
## Year2008          -0.4151      0.1944   -2.14    0.0334 *
## Year2009          -0.3793      0.1892   -2.00    0.0457 *
## Year2010          -0.2698      0.1815   -1.49    0.1379
## Year2011          -0.3512      0.1851   -1.90    0.0586 .
## Year2012          -0.3668      0.1920   -1.91    0.0569 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.4
## Multiple R-squared:  0.0553, Adjusted R-squared:  0.007
## Convergence in 32 IRWLS iterations
##
## Robustness weights:
## 32 weights are ~= 1. The remaining 339 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.144  0.877   0.953   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      2.70e-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 1.372 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
## -1.2546 -0.2502 0.0174 0.2754 1.4704
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4871 0.1749 8.50 5.4e-16 ***
## FirstAuthorFemale1 0.0279 0.0484 0.58 0.5639
## Year1997 -0.3799 0.6095 -0.62 0.5334
## Year1998 -0.3141 0.2687 -1.17 0.2432
## Year1999 -0.5670 0.2000 -2.83 0.0049 **
## Year2000 -0.4992 0.2027 -2.46 0.0143 *
## Year2001 -0.4757 0.2139 -2.22 0.0268 *
## Year2002 -0.3465 0.2176 -1.59 0.1122
## Year2003 -0.6516 0.2224 -2.93 0.0036 **
## Year2004 -0.3190 0.1911 -1.67 0.0959 .
## Year2005 -0.3344 0.1871 -1.79 0.0748 .
## Year2006 -0.3177 0.2008 -1.58 0.1145
```

```

## Year2007          -0.2643      0.1855   -1.43    0.1549
## Year2008          -0.4141      0.1950   -2.12    0.0344 *
## Year2009          -0.3725      0.1899   -1.96    0.0507 .
## Year2010          -0.2604      0.1814   -1.44    0.1521
## Year2011          -0.3427      0.1843   -1.86    0.0638 .
## Year2012          -0.3609      0.1924   -1.88    0.0616 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.399
## Multiple R-squared:  0.054, Adjusted R-squared:  0.00841
## Convergence in 32 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 341 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.145  0.871  0.951  0.893  0.986  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      2.70e-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 1.238 1      1.112
## Year      1.238 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
## -1.2663 -0.2509  0.0177  0.2720  1.4669

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.4935    0.1717   8.70 <2e-16 ***
## LastAuthorFemale1  0.0383    0.0548   0.70  0.4845
## Year1997         -0.3671    0.6045  -0.61  0.5441
## Year1998         -0.3057    0.2609  -1.17  0.2422
## Year1999         -0.5701    0.1988  -2.87  0.0044 **
## Year2000         -0.4994    0.2017  -2.48  0.0138 *
## Year2001         -0.4716    0.2107  -2.24  0.0258 *
## Year2002         -0.3504    0.2166  -1.62  0.1066
## Year2003         -0.6647    0.2227  -2.99  0.0030 **
## Year2004         -0.3150    0.1887  -1.67  0.0960 .
## Year2005         -0.3361    0.1860  -1.81  0.0716 .
## Year2006         -0.3226    0.1990  -1.62  0.1059
## Year2007         -0.2669    0.1829  -1.46  0.1454
## Year2008         -0.4092    0.1918  -2.13  0.0336 *
## Year2009         -0.3754    0.1874  -2.00  0.0459 *
## Year2010         -0.2655    0.1794  -1.48  0.1397
## Year2011         -0.3418    0.1819  -1.88  0.0611 .
## Year2012         -0.3606    0.1894  -1.90  0.0577 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.0539, Adjusted R-squared:  0.00839
## Convergence in 31 IRWLS iterations
##
## Robustness weights:
## 33 weights are ~= 1. The remaining 338 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.154  0.876  0.952  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      2.70e-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: 371"
## [1] ""

```

```

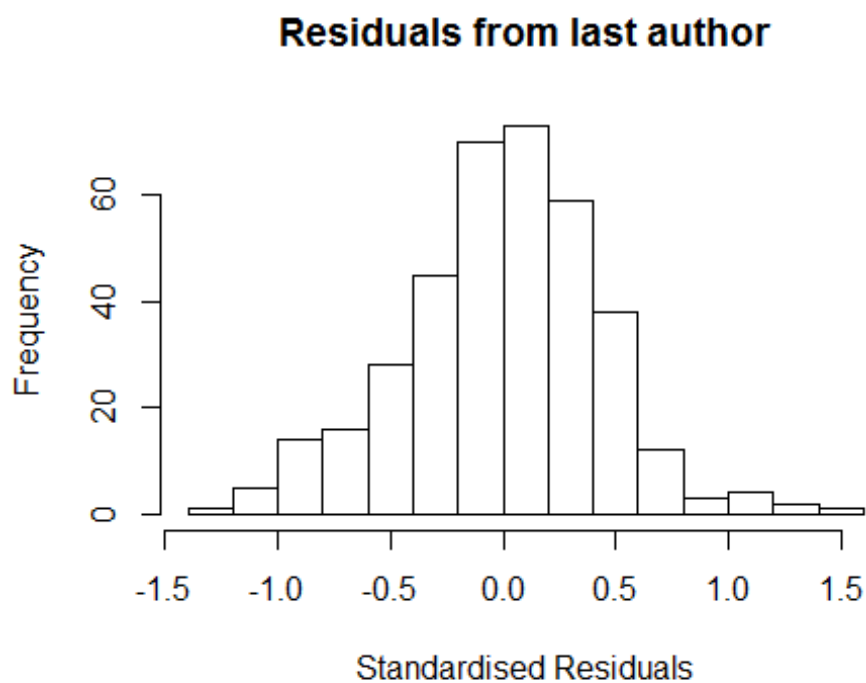
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1306"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 13 9 8 11 10 17 15 15 15 20 35 23 31 36 40
## 2011 2012
## 39 26
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1 3 1 8 2 6 9 9 5 13 17 14 15 21 23
## 2011 2012
## 26 19
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1 3 1 7 2 6 8 7 5 9 14 10 13 16 19
## 2011 2012
## 26 17
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 7.70980096180214"
## [1] "Male first author team size 2018 geometric mean: 6.09334624882746"

## Warning in wilcox.test.default(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 = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 8.17992564255437"
## [1] "Male last author team size 2018 geometric mean: 6.57589968005202"

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

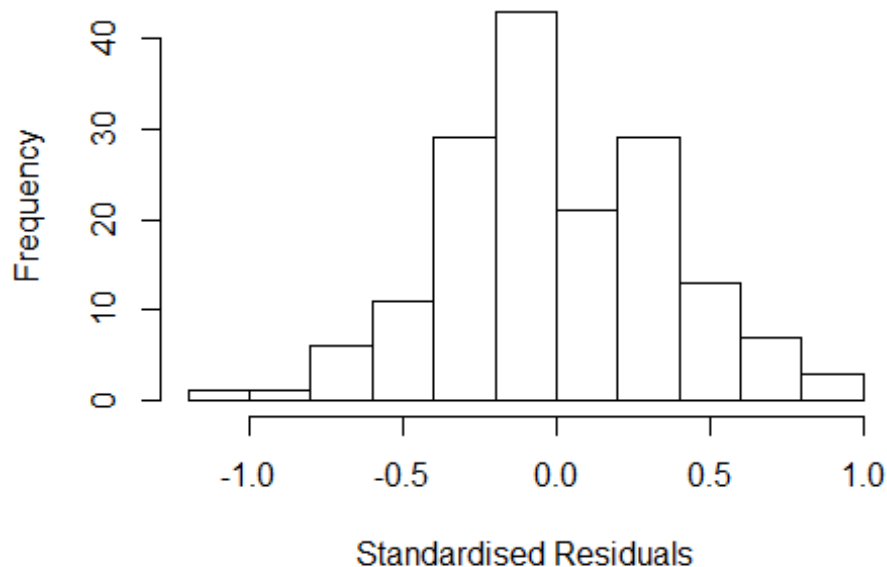
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 45, p-value = 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 | 4.961e+13 | 1 | 7.043e+06 |
| ## LastAuthorFemale | 2.606e+00 | 1 | 1.614e+00 |
| ## UniqueAuthors | 5.687e+14 | 4 | 6.988e+01 |
| ## Year | 2.281e+15 | 16 | 3.020e+00 |

Residuals from first and last author and team size



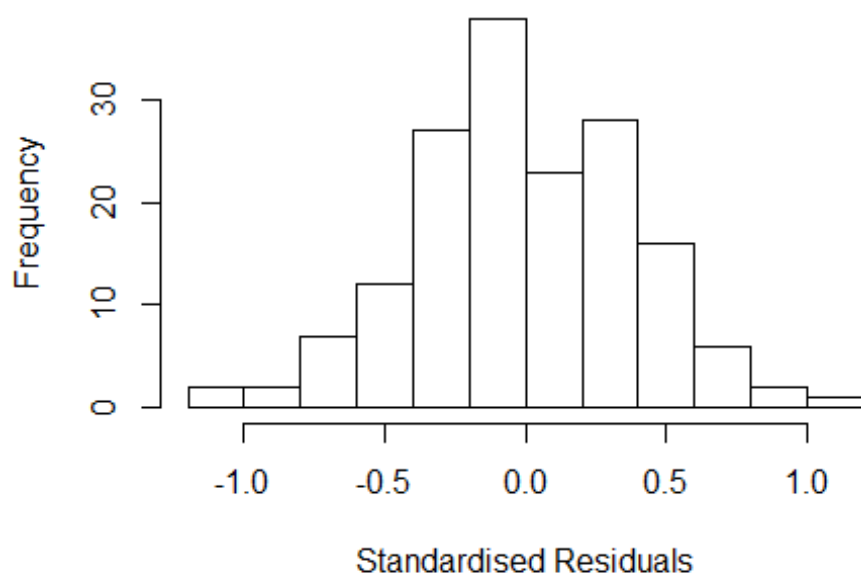
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0245 -0.2418 -0.0325 0.2596 0.9544
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7463 0.1711 4.36 2.5e-05 ***
## FirstAuthorFemale1 0.0341 0.0692 0.49 0.62343
## LastAuthorFemale1 0.1403 0.0750 1.87 0.06350 .
## UniqueAuthors2 0.3747 0.2959 1.27 0.20760
## UniqueAuthors3 0.5067 0.1711 2.96 0.00360 **
## UniqueAuthors4 0.4795 0.1656 2.90 0.00438 **
## UniqueAuthors5 0.5435 0.1379 3.94 0.00013 ***
## Year1997 0.1408 0.1981 0.71 0.47861
## Year1998 -0.2799 0.1371 -2.04 0.04313 *
## Year1999 0.0748 0.1727 0.43 0.66545
```

```

## Year2000          -0.2011      0.2524   -0.80   0.42703
## Year2001          -0.4040      0.1543   -2.62   0.00980 **
## Year2002          -0.3462      0.2186   -1.58   0.11537
## Year2003          -0.0965      0.2295   -0.42   0.67486
## Year2004          -0.1795      0.1743   -1.03   0.30491
## Year2005          -0.2285      0.1661   -1.38   0.17112
## Year2006          -0.2714      0.1477   -1.84   0.06818 .
## Year2007          -0.2350      0.1777   -1.32   0.18821
## Year2008          -0.3797      0.1440   -2.64   0.00933 **
## Year2009          -0.2148      0.1836   -1.17   0.24403
## Year2010          -0.2987      0.1699   -1.76   0.08082 .
## Year2011          -0.2395      0.1303   -1.84   0.06807 .
## Year2012          -0.1042      0.1324   -0.79   0.43282
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.389
## Multiple R-squared:  0.216, Adjusted R-squared:  0.0935
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 146 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.467  0.886  0.956  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           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 4.504e+15  1           6.711e+07
## LastAuthorFemale  2.443e+00  1           1.563e+00
## Year              1.052e+16 16           3.167e+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
## -1.0735 -0.2346 -0.0118  0.2521  1.0218
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.25e+00  4.86e-08  2.58e+07 < 2e-16 ***
## FirstAuthorFemale1  9.75e-02  7.27e-02  1.34e+00  0.1818
## LastAuthorFemale1  1.42e-01  7.11e-02  2.00e+00  0.0477 *
## Year1997          9.41e-02  1.46e-01  6.50e-01  0.5199
## Year1998         -3.07e-01  7.27e-02 -4.22e+00  4.3e-05 ***
## Year1999          5.82e-02  1.61e-01  3.60e-01  0.7182
## Year2000         -2.49e-01  2.81e-01 -8.90e-01  0.3756
## Year2001         -4.84e-01  1.88e-01 -2.58e+00  0.0109 *
## Year2002         -3.65e-01  1.91e-01 -1.92e+00  0.0574 .
## Year2003         -1.79e-01  1.43e-01 -1.26e+00  0.2108
## Year2004         -2.08e-01  1.57e-01 -1.33e+00  0.1872
## Year2005         -2.88e-01  1.08e-01 -2.67e+00  0.0085 **
```

```

## Year2006      -3.31e-01  1.27e-01 -2.61e+00  0.0100 *
## Year2007      -2.80e-01  1.38e-01 -2.03e+00  0.0444 *
## Year2008      -4.00e-01  9.04e-02 -4.43e+00  1.9e-05 ***
## Year2009      -2.32e-01  1.42e-01 -1.63e+00  0.1050
## Year2010      -3.64e-01  1.23e-01 -2.96e+00  0.0036 **
## Year2011      -2.72e-01  9.36e-02 -2.91e+00  0.0042 **
## Year2012      -1.96e-01  8.75e-02 -2.24e+00  0.0264 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.388
## Multiple R-squared:  0.125, Adjusted R-squared:  0.016
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 145 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.425  0.862  0.959  0.905  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      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 2.606e+11 1      5.105e+05
## Year              2.606e+11 16      2.274e+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

```

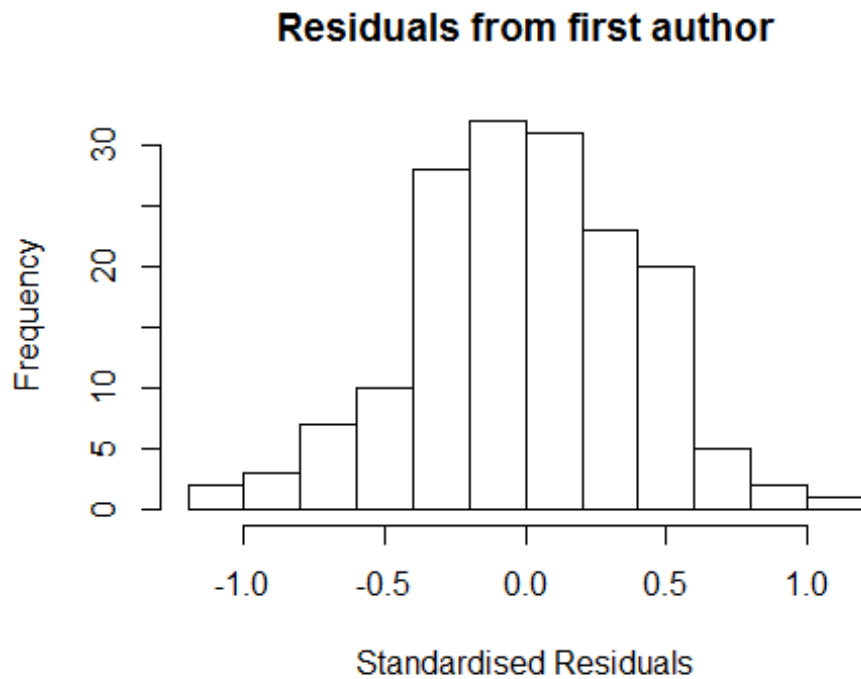
```

## -1.04756 -0.25341 0.00093 0.25717 1.14724
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.25e+00   1.03e-07  1.22e+07 < 2e-16 ***
## FirstAuthorFemale1 1.26e-01   7.35e-02  1.71e+00 0.08911 .
## Year1997         6.59e-02   1.46e-01  4.50e-01 0.65309
## Year1998        -3.35e-01   7.35e-02 -4.56e+00 1.1e-05 ***
## Year1999         1.02e-01   1.65e-01  6.20e-01 0.53786
## Year2000        -1.78e-01   2.06e-01 -8.70e-01 0.38843
## Year2001        -5.07e-01   1.87e-01 -2.71e+00 0.00756 **
## Year2002        -3.18e-01   1.98e-01 -1.60e+00 0.11093
## Year2003        -2.06e-01   1.46e-01 -1.42e+00 0.15885
## Year2004        -1.63e-01   1.47e-01 -1.10e+00 0.27123
## Year2005        -2.47e-01   1.15e-01 -2.14e+00 0.03416 *
## Year2006        -3.23e-01   1.31e-01 -2.47e+00 0.01459 *
## Year2007        -2.67e-01   1.41e-01 -1.89e+00 0.06032 .
## Year2008        -3.41e-01   9.96e-02 -3.43e+00 0.00079 ***
## Year2009        -2.05e-01   1.47e-01 -1.40e+00 0.16468
## Year2010        -3.25e-01   1.19e-01 -2.73e+00 0.00718 **
## Year2011        -2.55e-01   9.42e-02 -2.71e+00 0.00753 **
## Year2012        -1.64e-01   8.53e-02 -1.92e+00 0.05676 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.39
## Multiple R-squared:  0.105, Adjusted R-squared:  0.000358
## 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.
##   0.368 0.866  0.956  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          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"

```

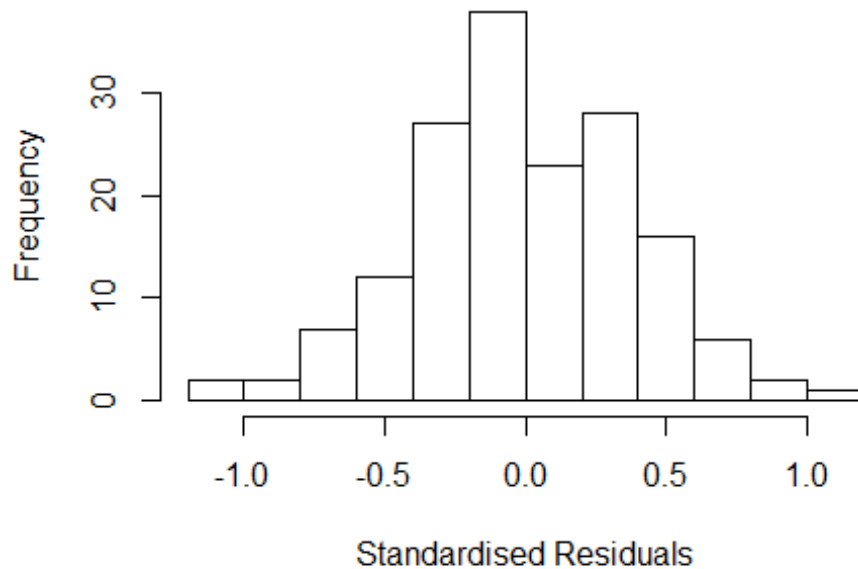
```
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

## 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
## -1.164813 -0.254521 -0.000111 0.263213 0.959293
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2530 0.0000 Inf < 2e-16 ***
## LastAuthorFemale1 0.1620 0.0688 2.36 0.0198 *
## Year1997 0.1916 0.1267 1.51 0.1326
## Year1998 -0.2090 0.0000 -Inf < 2e-16 ***
## Year1999 0.1200 0.1377 0.87 0.3848
## Year2000 -0.2595 0.2922 -0.89 0.3759
## Year2001 -0.4050 0.1678 -2.41 0.0170 *
## Year2002 -0.3276 0.1897 -1.73 0.0863 .
## Year2003 -0.0882 0.1026 -0.86 0.3913
## Year2004 -0.1421 0.1619 -0.88 0.3814
## Year2005 -0.2614 0.0982 -2.66 0.0086 **
## Year2006 -0.2708 0.1103 -2.45 0.0153 *
```

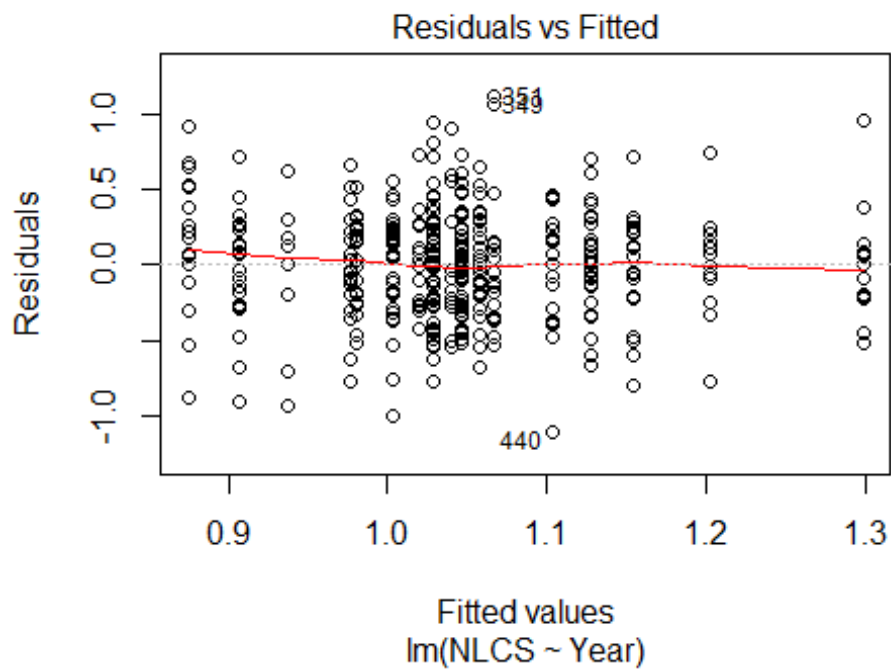


```

## Year2007          -0.2198      0.1451   -1.51    0.1320
## Year2008          -0.3566      0.0821   -4.35    2.6e-05 ***
## Year2009          -0.1865      0.1298   -1.44    0.1531
## Year2010          -0.3103      0.1298   -2.39    0.0181 *
## Year2011          -0.2238      0.0882   -2.54    0.0122 *
## Year2012          -0.1487      0.0804   -1.85    0.0663 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.386
## Multiple R-squared:  0.115, Adjusted R-squared:  0.0115
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 143 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.342  0.857   0.950   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      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 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
##   23   14   24   25   31   32   36   31   31   31   31   41   43   36   48
## 2011 2012
##   62   55
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   13    9   13   16   19   13   21   22   19   21   21   29   33   24   36
## 2011 2012

```

```
## 52 43
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 11 8 12 15 13 12 17 16 15 18 17 21 27 20 29
## 2011 2012
## 40 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 = 26, df = 16, p-value = 0.06
```



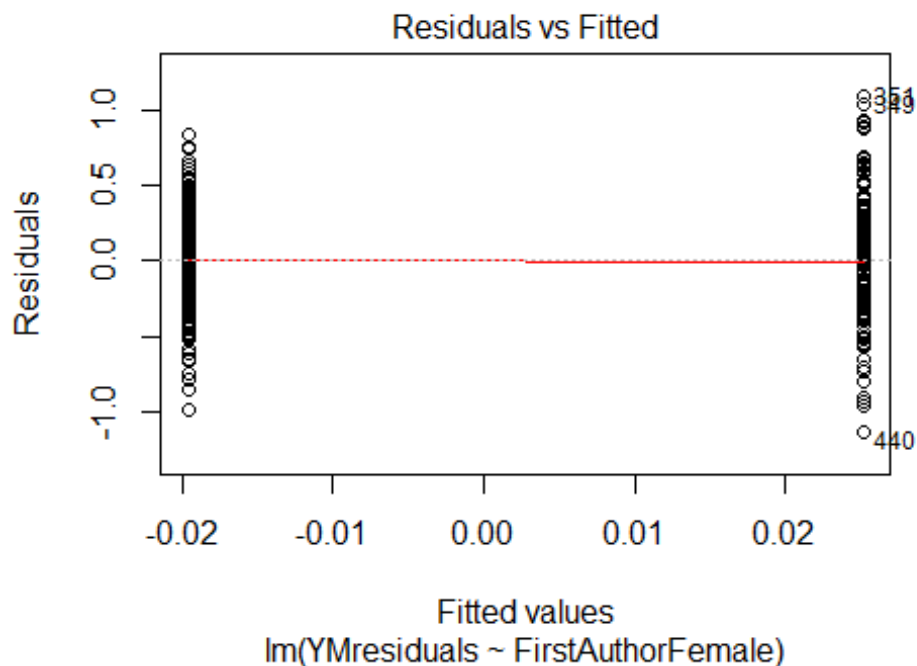
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.5, df = 1, p-value = 0.06

## [1] "Female first author team size 2018 geometric mean: 5.53286520949602"
## [1] "Male first author team size 2018 geometric mean: 4.51536609126718"

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

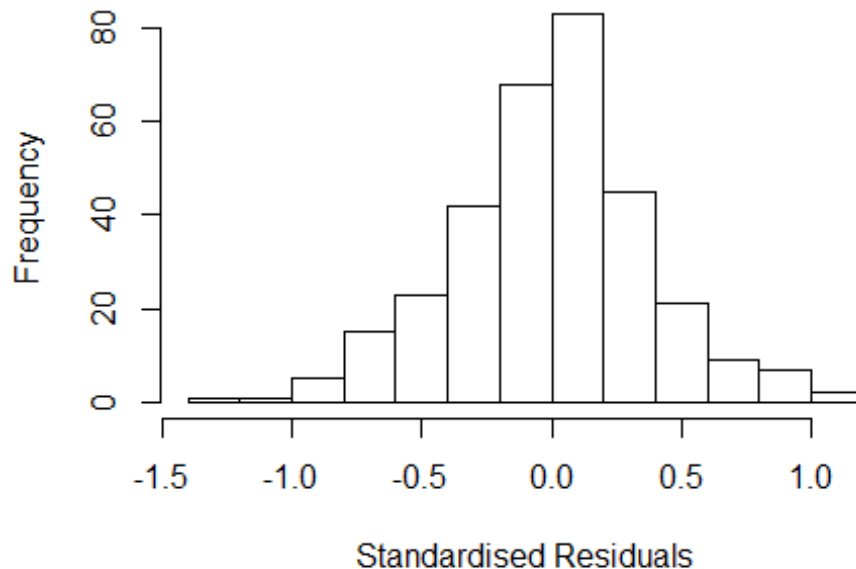
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 150, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.08460581366762"
## [1] "Male last author team size 2018 geometric mean: 4.93497785691118"

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



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 75, 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 1.493  1      1.222
## LastAuthorFemale  1.233  1      1.110
## UniqueAuthors    3.509  4      1.170
## Year              4.852 16      1.051
```

Residuals from first and last author and team size



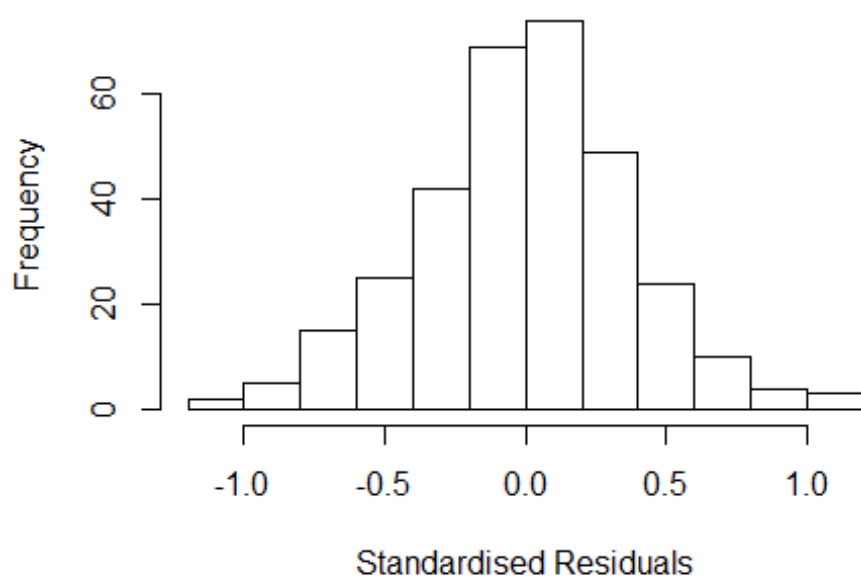
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2206 -0.2465 0.0173 0.2189 1.1525
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2000 0.1259 9.53 < 2e-16 ***
## FirstAuthorFemale1 -0.0635 0.0478 -1.33 0.18540
## LastAuthorFemale1 -0.0264 0.0465 -0.57 0.57088
## UniqueAuthors2 0.2158 0.1231 1.75 0.08065 .
## UniqueAuthors3 0.1420 0.1229 1.16 0.24862
## UniqueAuthors4 0.2204 0.1236 1.78 0.07542 .
## UniqueAuthors5 0.2210 0.1192 1.85 0.06462 .
## Year1997 -0.2867 0.2669 -1.07 0.28348
## Year1998 -0.3235 0.1711 -1.89 0.05970 .
## Year1999 -0.3774 0.1100 -3.43 0.00069 ***
```

```

## Year2000          -0.6055      0.2934    -2.06    0.03989 *
## Year2001          -0.1859      0.1328    -1.40    0.16286
## Year2002          -0.1497      0.1280    -1.17    0.24316
## Year2003          -0.4393      0.1322    -3.32    0.00100 **
## Year2004          -0.3883      0.1442    -2.69    0.00747 **
## Year2005          -0.3851      0.1303    -2.96    0.00337 **
## Year2006          -0.1951      0.1240    -1.57    0.11645
## Year2007          -0.2896      0.1157    -2.50    0.01286 *
## Year2008          -0.1646      0.1091    -1.51    0.13241
## Year2009          -0.2948      0.1069    -2.76    0.00619 **
## Year2010          -0.3342      0.1065    -3.14    0.00186 **
## Year2011          -0.3444      0.1065    -3.23    0.00136 **
## Year2012          -0.2980      0.1102    -2.71    0.00722 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.345
## Multiple R-squared:  0.11,   Adjusted R-squared:  0.0443
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 294 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.186  0.864  0.948   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          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 1.449 1          1.204
## LastAuthorFemale  1.167 1          1.080
## Year              1.636 16          1.016

```

Residuals from first and last author



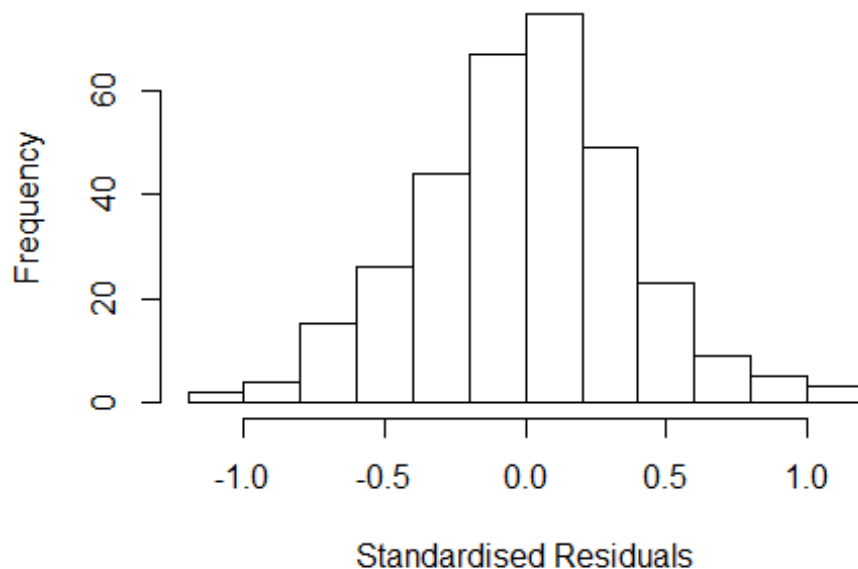
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1925 -0.2412  0.0102  0.2219  1.1700
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3596     0.0962   14.13  <2e-16 ***
## FirstAuthorFemale1 -0.0469     0.0483   -0.97   0.3322
## LastAuthorFemale1 -0.0230     0.0464   -0.50   0.6200
## Year1997          -0.2788     0.2562   -1.09   0.2774
## Year1998          -0.3229     0.1742   -1.85   0.0648 .
## Year1999          -0.3443     0.1121   -3.07   0.0023 **
## Year2000          -0.6063     0.3518   -1.72   0.0859 .
## Year2001          -0.1429     0.1297   -1.10   0.2716
## Year2002          -0.1257     0.1281   -0.98   0.3270
## Year2003          -0.4287     0.1408   -3.04   0.0025 **
## Year2004          -0.3496     0.1401   -2.50   0.0131 *
## Year2005          -0.3469     0.1275   -2.72   0.0069 **
```

```

## Year2006          -0.1671      0.1325    -1.26    0.2081
## Year2007          -0.2592      0.1182    -2.19    0.0291 *
## Year2008          -0.1377      0.1078    -1.28    0.2026
## Year2009          -0.2776      0.1102    -2.52    0.0123 *
## Year2010          -0.3168      0.1126    -2.81    0.0052 **
## Year2011          -0.3088      0.1088    -2.84    0.0049 **
## Year2012          -0.2808      0.1105    -2.54    0.0115 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.346
## Multiple R-squared:  0.0927, Adjusted R-squared:  0.0388
## Convergence in 27 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 294 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.209  0.863   0.953   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      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 1.429 1      1.196
## Year              1.429 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
## -1.19254 -0.23593  0.00539  0.21825  1.17896
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3570     0.0971   13.97  <2e-16 ***
## FirstAuthorFemale1 -0.0498     0.0479   -1.04   0.2992
## Year1997         -0.2850     0.2529   -1.13   0.2606
## Year1998         -0.3252     0.1737   -1.87   0.0621 .
## Year1999         -0.3445     0.1141   -3.02   0.0028 **
## Year2000         -0.6102     0.3577   -1.71   0.0890 .
## Year2001         -0.1483     0.1301   -1.14   0.2551
## Year2002         -0.1301     0.1281   -1.02   0.3105
## Year2003         -0.4316     0.1408   -3.07   0.0024 **
## Year2004         -0.3559     0.1390   -2.56   0.0109 *
## Year2005         -0.3528     0.1281   -2.75   0.0062 **
## Year2006         -0.1644     0.1334   -1.23   0.2188
```

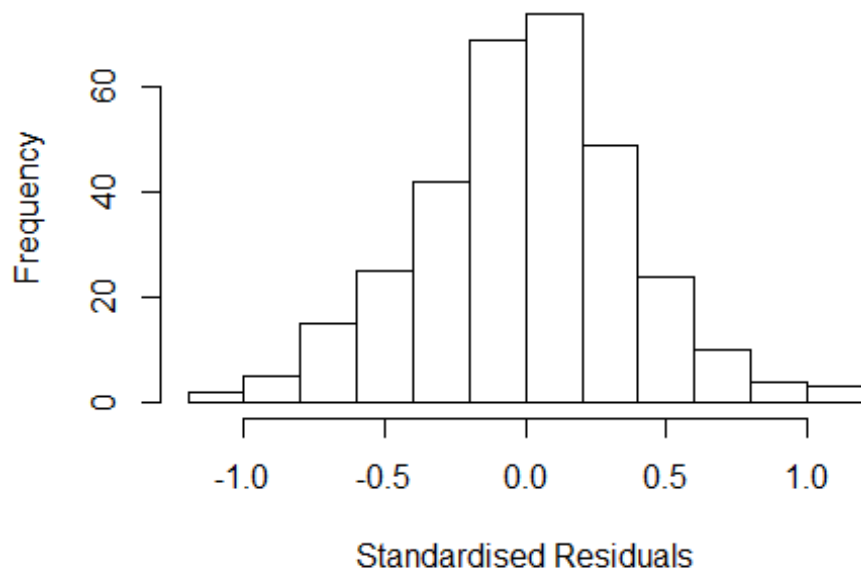


```

## Year2007          -0.2611      0.1194    -2.19    0.0295 *
## Year2008          -0.1399      0.1090    -1.28    0.2005
## Year2009          -0.2810      0.1116    -2.52    0.0123 *
## Year2010          -0.3232      0.1134    -2.85    0.0047 **
## Year2011          -0.3112      0.1105    -2.82    0.0052 **
## Year2012          -0.2846      0.1111    -2.56    0.0109 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.347
## Multiple R-squared:  0.0921, Adjusted R-squared:  0.0414
## Convergence in 27 IRWLS iterations
##
## Robustness weights:
## 27 weights are ~= 1. The remaining 295 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.213  0.865   0.954   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      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 1.162 1          1.078
## Year            1.162 16          1.005

```

Residuals from last author



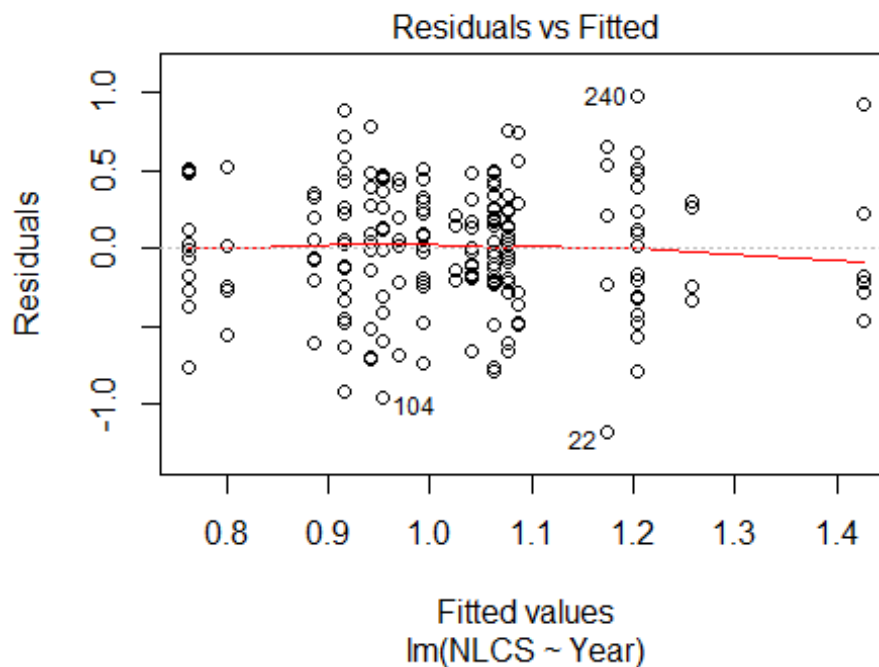
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1657 -0.2262 0.0134 0.2187 1.1868
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3285 0.0905 14.67 <2e-16 ***
## LastAuthorFemale1 -0.0293 0.0465 -0.63 0.5292
## Year1997 -0.2701 0.2600 -1.04 0.2999
## Year1998 -0.3189 0.1712 -1.86 0.0635 .
## Year1999 -0.3388 0.1134 -2.99 0.0030 **
## Year2000 -0.6178 0.3579 -1.73 0.0853 .
## Year2001 -0.1332 0.1267 -1.05 0.2939
## Year2002 -0.1107 0.1299 -0.85 0.3949
## Year2003 -0.4283 0.1368 -3.13 0.0019 **
## Year2004 -0.3353 0.1390 -2.41 0.0165 *
## Year2005 -0.3356 0.1267 -2.65 0.0085 **
## Year2006 -0.1627 0.1337 -1.22 0.2247
```

```

## Year2007          -0.2450      0.1170    -2.09    0.0371 *
## Year2008          -0.1398      0.1080    -1.29    0.1965
## Year2009          -0.2793      0.1101    -2.54    0.0117 *
## Year2010          -0.3125      0.1122    -2.79    0.0057 **
## Year2011          -0.3036      0.1093    -2.78    0.0058 **
## Year2012          -0.2698      0.1093    -2.47    0.0141 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.345
## Multiple R-squared:  0.0921, Adjusted R-squared:  0.0413
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 27 weights are ~= 1. The remaining 295 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.214  0.852  0.953  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      3.11e-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 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
##   20   10   13   10   12   12   14   17   10   10   19   21   25   23   33
## 2011 2012
##   42   27
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    8    5    6    8    4    4   12   12    7    6   11   13   19   17   20
## 2011 2012

```

```
## 29 19
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 7 3 5 8 4 3 10 9 6 6 9 11 17 16 18
## 2011 2012
## 29 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 = 19, df = 16, p-value = 0.3
```



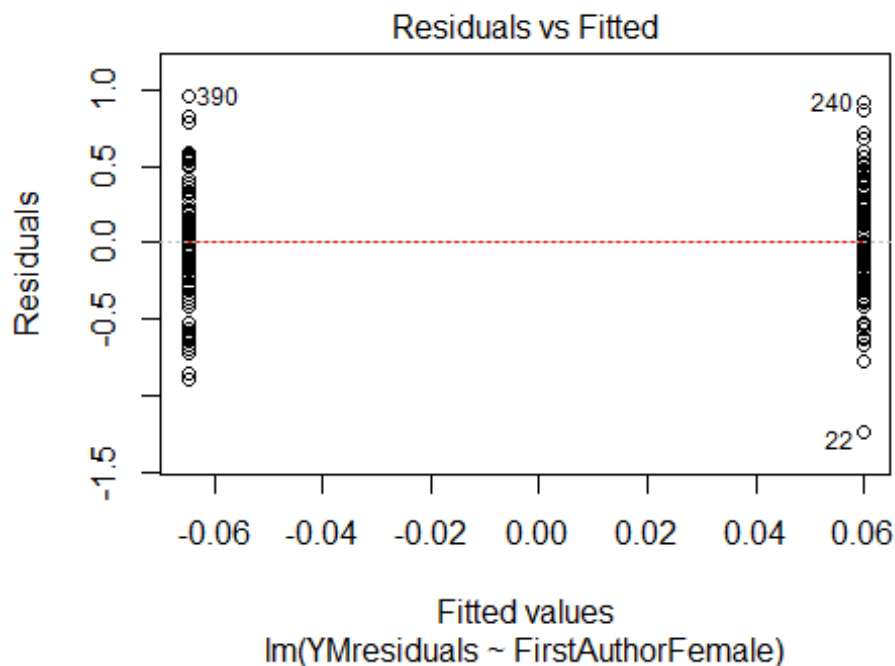
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.5, df = 1, p-value = 0.5

## [1] "Female first author team size 2018 geometric mean: 6.86828545531999"
## [1] "Male first author team size 2018 geometric mean: 3.6889272467625"

## Warning in wilcox.test.default(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 = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 7"
## [1] "Male last author team size 2018 geometric mean: 4.29888610333493"

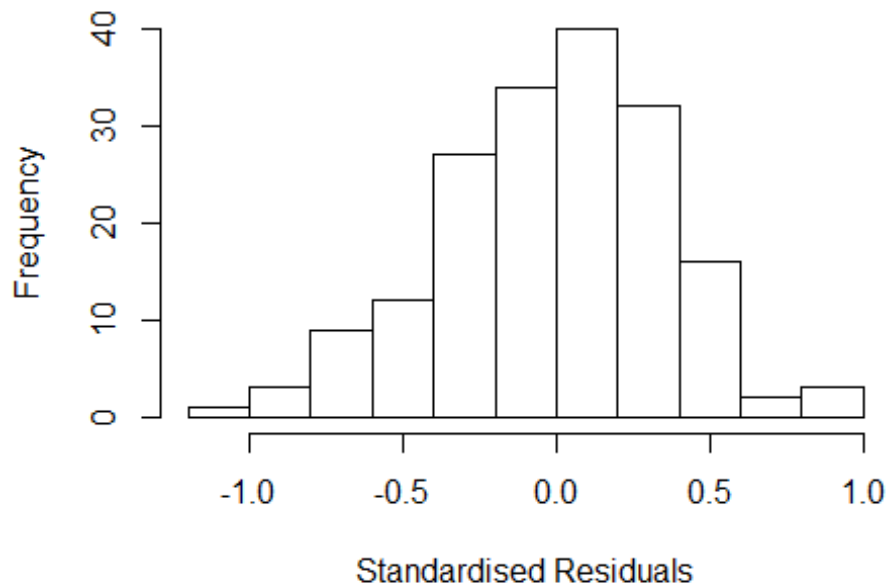
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## 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] "Regression 1: First author gender, last author gender, team size,
## Year as factors"
##
```

| | GVIF | Df | GVIF^(1/(2*Df)) |
|-------------------|--------|----|-----------------|
| FirstAuthorFemale | 1.559 | 1 | 1.249 |
| LastAuthorFemale | 1.781 | 1 | 1.334 |
| UniqueAuthors | 18.982 | 4 | 1.445 |
| Year | 33.981 | 16 | 1.116 |

Residuals from first and last author and team size



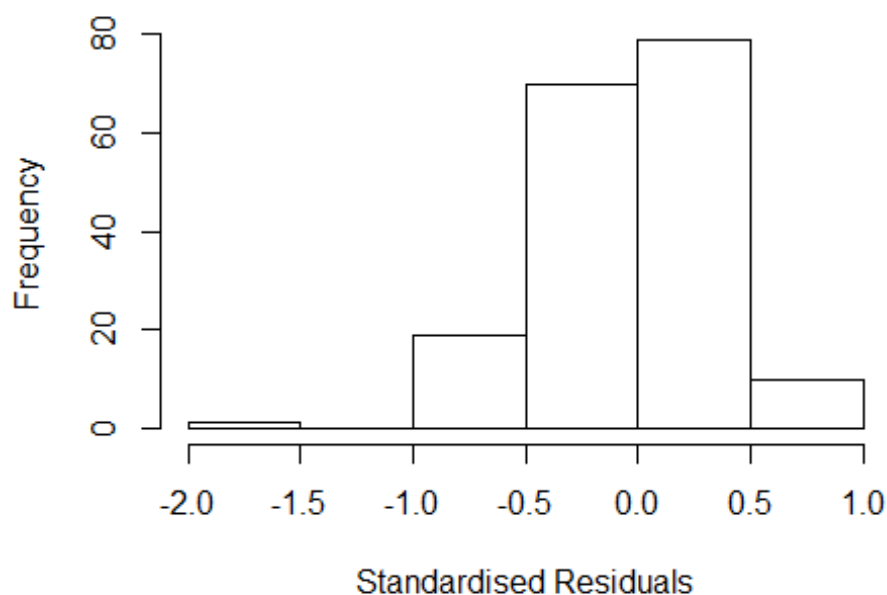
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.039 -0.229  0.024  0.260  0.934
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.45672    0.26616   1.72   0.0882 .
## FirstAuthorFemale1 -0.08393    0.06303  -1.33   0.1850
## LastAuthorFemale1  0.11326    0.07808   1.45   0.1489
## UniqueAuthors2    0.55531    0.23210   2.39   0.0179 *
## UniqueAuthors3    0.47790    0.25626   1.86   0.0641 .
## UniqueAuthors4    0.52858    0.23803   2.22   0.0278 *
## UniqueAuthors5    0.58495    0.21745   2.69   0.0079 **
## Year1997          0.58209    0.51230   1.14   0.2576
## Year1998         -0.16267    0.27265  -0.60   0.5516
## Year1999          0.01721    0.17609   0.10   0.9223
```

```

## Year2000      0.37122      0.19829      1.87      0.0631 .
## Year2001      0.06672      0.15954      0.42      0.6764
## Year2002      0.02503      0.19300      0.13      0.8970
## Year2003     -0.00287      0.19143     -0.01      0.9881
## Year2004      0.39039      0.33395      1.17      0.2442
## Year2005      0.35160      0.21193      1.66      0.0991 .
## Year2006     -0.04581      0.18391     -0.25      0.8036
## Year2007      0.06471      0.15844      0.41      0.6835
## Year2008      0.23365      0.17396      1.34      0.1812
## Year2009      0.05861      0.15382      0.38      0.7037
## Year2010      0.01799      0.15619      0.12      0.9085
## Year2011      0.05851      0.14869      0.39      0.6945
## Year2012     -0.01432      0.21013     -0.07      0.9457
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.374
## Multiple R-squared:  0.185, Adjusted R-squared:  0.0695
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 166 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.421  0.890  0.956  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      5.59e-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 1.644 1      1.282
## LastAuthorFemale  1.615 1      1.271
## Year              2.368 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
## -1.7656 -0.2450 -0.0128 0.2665 0.9804
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9728 0.1240 7.85 5.7e-13 ***
## FirstAuthorFemale1 -0.0832 0.0651 -1.28 0.203
## LastAuthorFemale1 0.1080 0.0768 1.41 0.162
## Year1997 0.7928 0.1334 5.94 1.7e-08 ***
## Year1998 -0.1500 0.2599 -0.58 0.565
## Year1999 0.0209 0.1742 0.12 0.904
## Year2000 0.2799 0.2054 1.36 0.175
## Year2001 0.0930 0.1628 0.57 0.568
## Year2002 0.0539 0.1915 0.28 0.779
## Year2003 -0.1106 0.2102 -0.53 0.600
## Year2004 0.1797 0.2863 0.63 0.531
## Year2005 0.3713 0.2092 1.77 0.078 .
```

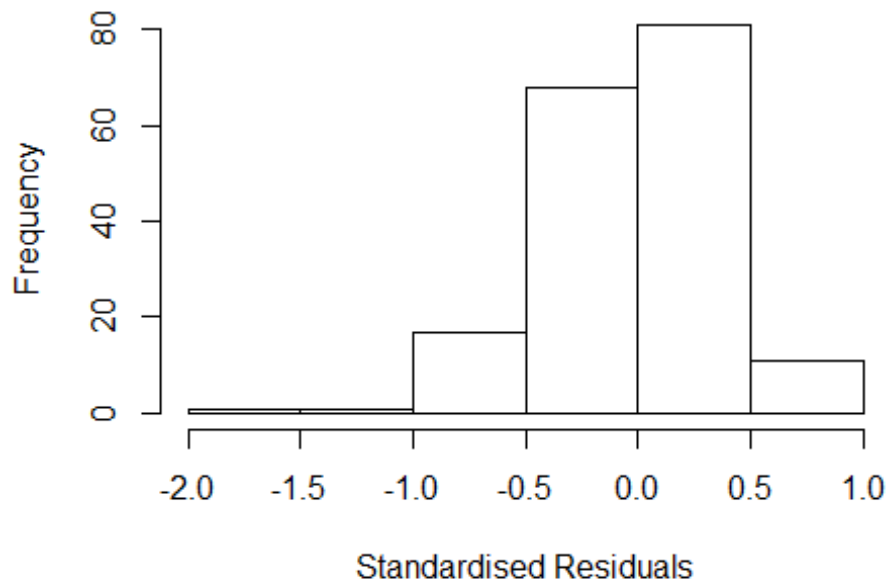


```

## Year2006          -0.0847      0.1670    -0.51     0.613
## Year2007           0.0963      0.1536      0.63     0.532
## Year2008           0.2268      0.1651      1.37     0.172
## Year2009           0.0994      0.1416      0.70     0.484
## Year2010           0.0582      0.1414      0.41     0.681
## Year2011           0.1112      0.1334      0.83     0.406
## Year2012           0.0370      0.1822      0.20     0.839
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.389
## Multiple R-squared:  0.14,    Adjusted R-squared:  0.0433
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 169 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0038 0.8840 0.9560 0.9120 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      5.59e-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 1.561 1      1.249
## Year              1.561 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
## -1.76475 -0.24436  0.00311  0.25025  0.98809
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.9666    0.1233   7.84 5.8e-13 ***
## FirstAuthorFemale1 -0.0706    0.0634  -1.11  0.267
## Year1997          0.7981    0.1330   6.00 1.2e-08 ***
## Year1998         -0.1354    0.2770  -0.49  0.626
## Year1999          0.0508    0.1645   0.31  0.758
## Year2000          0.3089    0.1899   1.63  0.106
## Year2001          0.1324    0.1672   0.79  0.430
## Year2002          0.1112    0.1822   0.61  0.542
## Year2003         -0.0870    0.1997  -0.44  0.664
## Year2004          0.1860    0.2854   0.65  0.515
## Year2005          0.3933    0.2117   1.86  0.065 .
## Year2006         -0.0736    0.1668  -0.44  0.660
```

```

## Year2007          0.1307      0.1488      0.88      0.381
## Year2008          0.2473      0.1654      1.49      0.137
## Year2009          0.1129      0.1406      0.80      0.423
## Year2010          0.0943      0.1384      0.68      0.497
## Year2011          0.1486      0.1309      1.14      0.258
## Year2012          0.0532      0.1867      0.29      0.776
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.39
## Multiple R-squared:  0.128, Adjusted R-squared:  0.0357
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 169 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0048 0.8820 0.9590 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      5.59e-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 1.544 1      1.243
## Year      1.544 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
## -1.764273 -0.229240  0.000701  0.274579  1.017037

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9329    0.1150    8.11 1.2e-13 ***
## LastAuthorFemale1 0.0955    0.0766    1.25  0.214
## Year1997        0.8313    0.1244    6.68 3.7e-10 ***
## Year1998       -0.1665    0.2578   -0.65  0.519
## Year1999        0.0243    0.1831    0.13  0.895
## Year2000        0.3039    0.2136    1.42  0.157
## Year2001        0.1081    0.1568    0.69  0.491
## Year2002        0.0784    0.1911    0.41  0.682
## Year2003       -0.1108    0.1968   -0.56  0.574
## Year2004        0.2198    0.2820    0.78  0.437
## Year2005        0.3956    0.2103    1.88  0.062 .
## Year2006       -0.1071    0.1670   -0.64  0.522
## Year2007        0.1039    0.1518    0.68  0.494
## Year2008        0.2300    0.1625    1.42  0.159
## Year2009        0.0996    0.1379    0.72  0.471
## Year2010        0.0784    0.1352    0.58  0.563
## Year2011        0.1100    0.1314    0.84  0.404
## Year2012        0.0299    0.1769    0.17  0.866
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.391
## Multiple R-squared:  0.13, Adjusted R-squared:  0.0385
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 163 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0053 0.8870 0.9530 0.9080 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      5.59e-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: 179"
## [1] ""

```

```

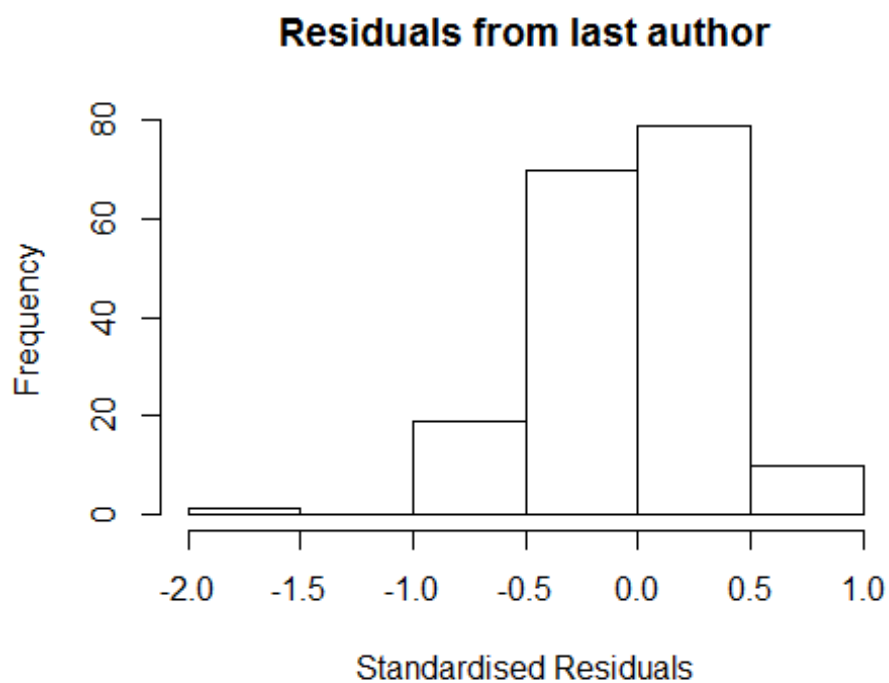
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1309"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    5    2    5    5    5    9    7   10    5    7   15   19   17   13
## 2011 2012
##   14   18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    1    0    3    1    3    4    5    7    2    4   11   11   14   11
## 2011 2012
##   13   12
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    0    0    2    0    3    4    3    4    2    2    8    8   11    9
## 2011 2012
##   10    8
## [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: 4.40923831739395"

## Warning in wilcox.test.default(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 = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.64621617328617"
## [1] "Male last author team size 2018 geometric mean: 4.38432765486578"

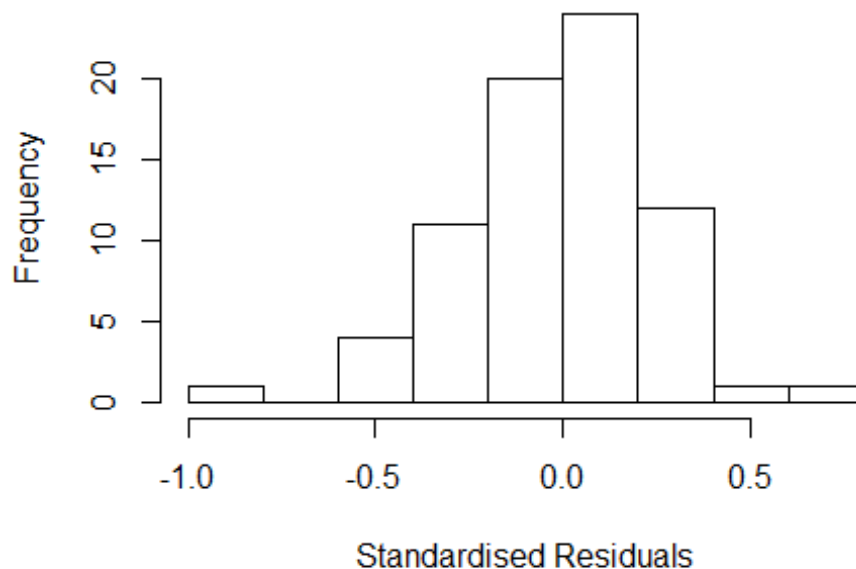
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## 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] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.597 1      1.612
## LastAuthorFemale  2.710 1      1.646
## UniqueAuthors    26.732 4      1.508
## Year              98.709 12     1.211
```

Residuals from first and last author and team size



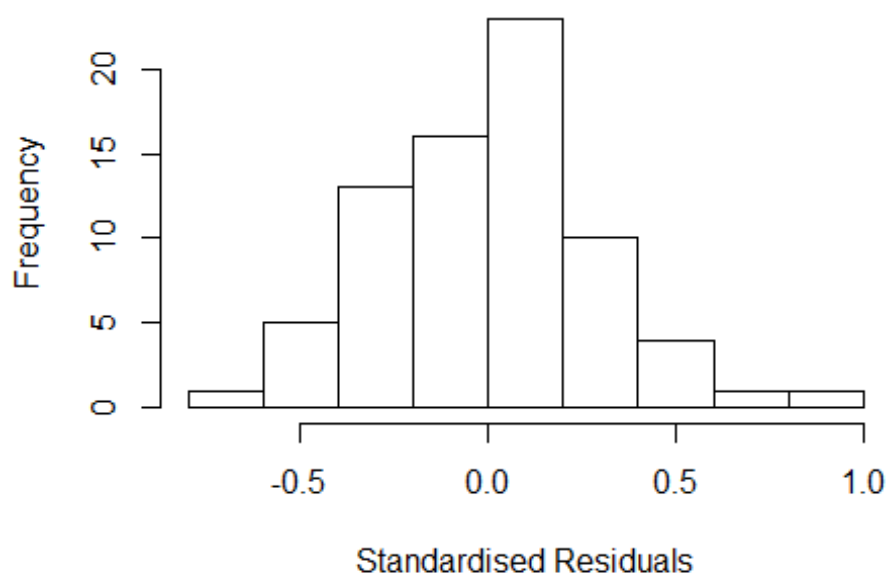
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8361 -0.1868 0.0177 0.1644 0.6215
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.4627 0.2378 1.95 0.0568 .
## FirstAuthorFemale1 -0.0394 0.0793 -0.50 0.6210
## LastAuthorFemale1 0.1631 0.0752 2.17 0.0345 *
## UniqueAuthors2 0.2444 0.1697 1.44 0.1555
## UniqueAuthors3 0.2099 0.1796 1.17 0.2475
## UniqueAuthors4 0.2320 0.1810 1.28 0.2052
## UniqueAuthors5 0.5127 0.1685 3.04 0.0036 **
## Year2001 0.4626 0.1885 2.45 0.0173 *
## Year2002 0.3027 0.3323 0.91 0.3662
## Year2003 0.1929 0.2667 0.72 0.4726
```

```

## Year2004          0.3043      0.2032      1.50      0.1400
## Year2005          0.4308      0.2202      1.96      0.0555 .
## Year2006          0.5267      0.1889      2.79      0.0073 **
## Year2007          0.2815      0.1864      1.51      0.1368
## Year2008          0.1636      0.2408      0.68      0.4997
## Year2009          0.1007      0.2137      0.47      0.6395
## Year2010         -0.0718      0.1929     -0.37      0.7113
## Year2011          0.3001      0.2131      1.41      0.1647
## Year2012          0.0194      0.2106      0.09      0.9268
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.263
## Multiple R-squared:  0.37,    Adjusted R-squared:  0.164
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  5 weights are ~= 1. The remaining 69 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.291  0.901  0.954   0.911   0.984   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      1.35e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.304 1      1.818
## LastAuthorFemale  1.722 1      1.312
## Year              5.376 12      1.073

```


Residuals from first and last author



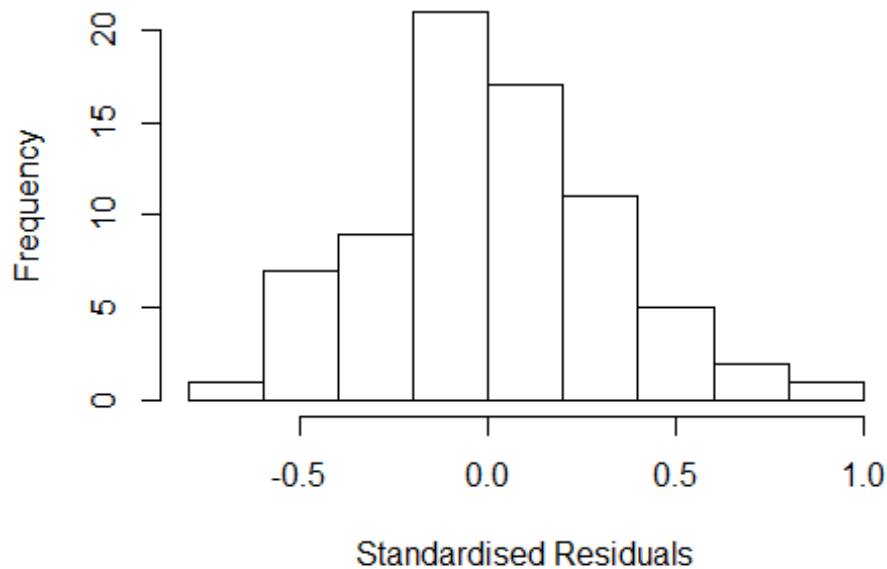
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.6197 -0.1989 0.0185 0.1733 0.8080
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7986 0.1008 7.92 7.4e-11 ***
## FirstAuthorFemale1 0.0458 0.0869 0.53 0.6007
## LastAuthorFemale1 0.1795 0.0815 2.20 0.0315 *
## Year2001 0.4328 0.1473 2.94 0.0047 **
## Year2002 0.2644 0.7022 0.38 0.7079
## Year2003 0.0913 0.1958 0.47 0.6427
## Year2004 0.2640 0.1169 2.26 0.0276 *
## Year2005 0.2196 0.1116 1.97 0.0537 .
## Year2006 0.5013 0.1634 3.07 0.0033 **
## Year2007 0.2014 0.1106 1.82 0.0738 .
## Year2008 0.0448 0.1761 0.25 0.8002
## Year2009 -0.0383 0.1816 -0.21 0.8336
```

```

## Year2010          -0.0027      0.1136   -0.02   0.9811
## Year2011          0.1753      0.1319    1.33   0.1888
## Year2012          0.0918      0.1482    0.62   0.5382
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.295
## Multiple R-squared:  0.2,   Adjusted R-squared:  0.00962
## Convergence in 31 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 66 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.432  0.869  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          1.35e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.137 1          1.771
## Year              3.137 12          1.049

```

Residuals from first author



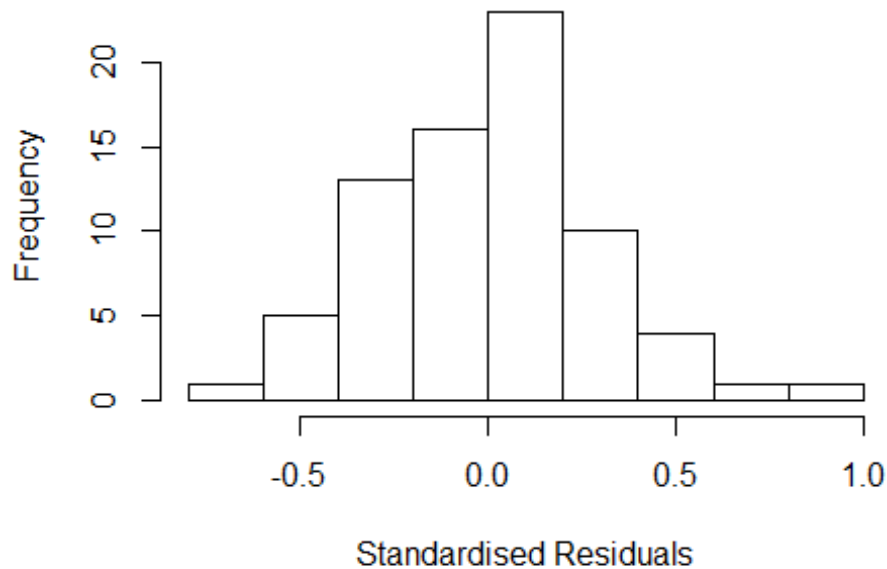
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.6819 -0.1583 -0.0174 0.1960 0.8057
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7922 0.1072 7.39 5.4e-10 ***
## FirstAuthorFemale1 0.0587 0.0960 0.61 0.5434
## Year2001 0.4353 0.1539 2.83 0.0064 **
## Year2002 0.2731 0.6949 0.39 0.6957
## Year2003 0.0981 0.1989 0.49 0.6239
## Year2004 0.3596 0.1275 2.82 0.0065 **
## Year2005 0.2132 0.1174 1.82 0.0744 .
## Year2006 0.5910 0.2431 2.43 0.0180 *
## Year2007 0.2438 0.1168 2.09 0.0412 *
## Year2008 0.1576 0.1569 1.00 0.3192
## Year2009 0.0487 0.1895 0.26 0.7980
## Year2010 0.0621 0.1171 0.53 0.5976
```

```

## Year2011          0.2310      0.1355      1.71      0.0932 .
## Year2012          0.1877      0.1620      1.16      0.2513
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.299
## Multiple R-squared:  0.146, Adjusted R-squared:  -0.0392
## Convergence in 30 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 66 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.447  0.855  0.948  0.904  0.986  0.999
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          1.35e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.814 1          1.347
## Year            1.814 12          1.025

```

Residuals from last author



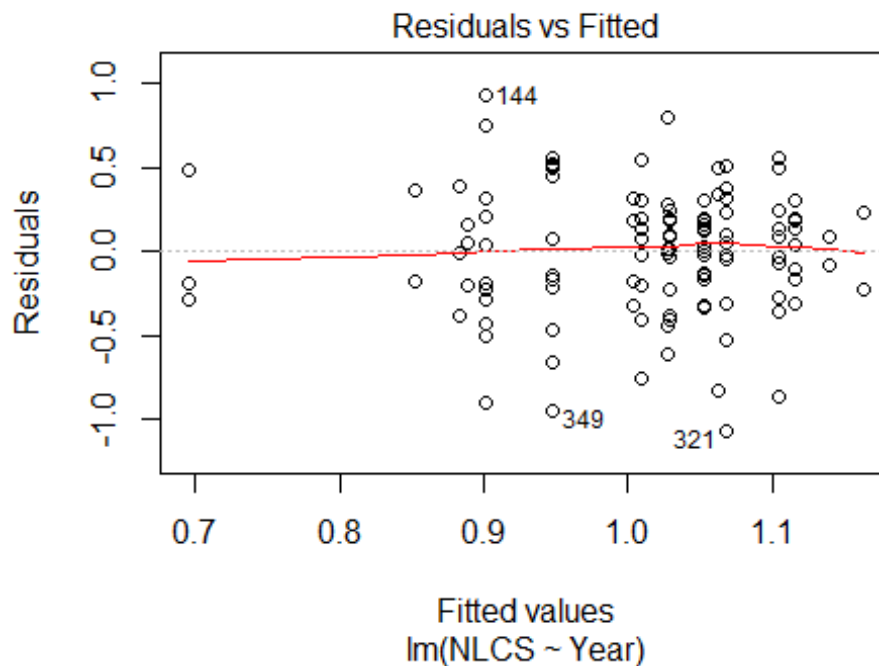
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.6045 -0.1893 0.0148 0.1780 0.8015
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.82150 0.07344 11.19 2.6e-16 ***
## LastAuthorFemale1 0.18247 0.08385 2.18 0.03349 *
## Year2001 0.42384 0.12294 3.45 0.00104 **
## Year2002 0.24801 0.67848 0.37 0.71599
## Year2003 0.06789 0.18443 0.37 0.71410
## Year2004 0.23964 0.09363 2.56 0.01302 *
## Year2005 0.24250 0.08765 2.77 0.00752 **
## Year2006 0.49977 0.13828 3.61 0.00062 ***
## Year2007 0.20082 0.09493 2.12 0.03855 *
## Year2008 0.06004 0.16872 0.36 0.72318
## Year2009 -0.03004 0.18109 -0.17 0.86881
## Year2010 0.00801 0.10093 0.08 0.93702
```

```

## Year2011          0.18301    0.12535    1.46  0.14952
## Year2012          0.07944    0.13243    0.60  0.55083
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.288
## Multiple R-squared:  0.196, Adjusted R-squared:  0.0218
## Convergence in 31 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 70 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.419  0.859  0.960   0.910   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           1.35e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01             5.00e-01
##   nResample      max.it    best.r.s    k.fast.s      k.max maxit.scale
##           500          50        2          1        1000         200
##   trace.lev      mts    compute.rd
##           0          1000         0
##           psi          subsampling          cov
##           "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 74"
## [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
##   18   13   12    9   14   11   15   12   17   15   17   20   20   12   22
## 2011 2012
##   21   20
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    2    3    2    3    3    4    3   12   10   11   10   14    6   11
## 2011 2012
##   13   12
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    0    3    0    2    2    4    2   10    9    9    9    9    6    9

```

```
## 2011 2012
## 11 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 = 0.06
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.8, df = 1, p-value = 0.2

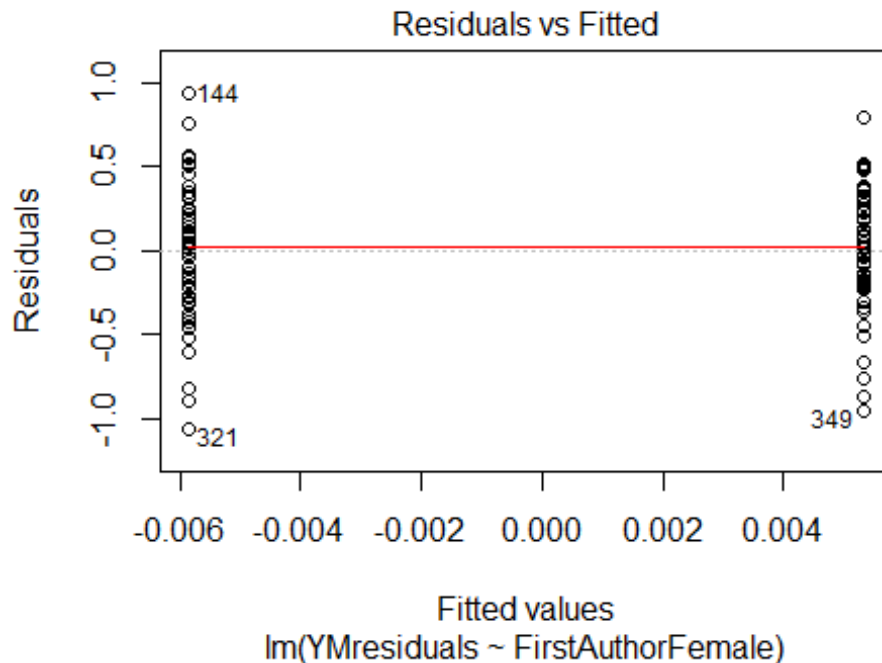
## [1] "Female first author team size 2018 geometric mean: 6.69432950082169"
## [1] "Male first author team size 2018 geometric mean: 5.47722557505166"

## 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 = 4, p-value = 0.8
```

```
## 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: 6.22332977288478"

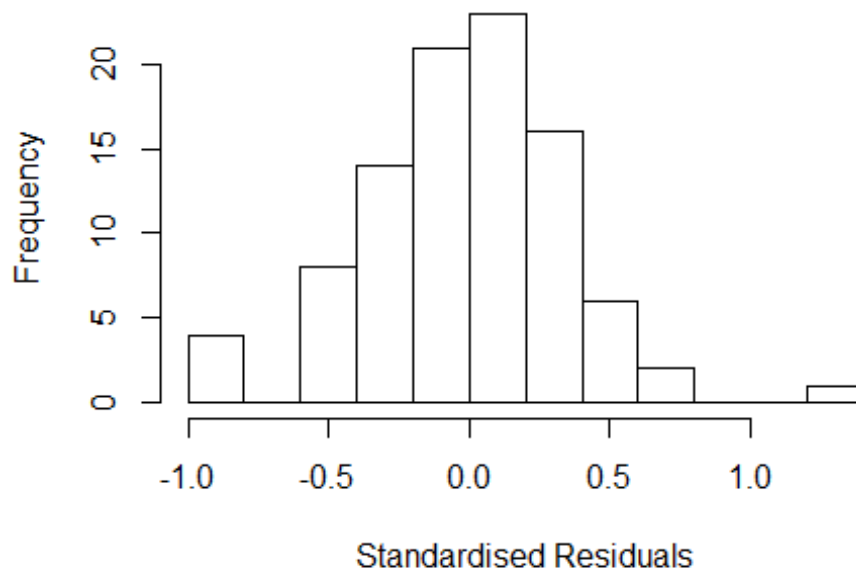
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## 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] "Regression 1: First author gender, last author gender, team size,
## Year as factors"
##
```

| | GVIF | Df | GVIF ^{1/(2*Df)} |
|-------------------|---------|----|--------------------------|
| FirstAuthorFemale | 2.094 | 1 | 1.447 |
| LastAuthorFemale | 4.801 | 1 | 2.191 |
| UniqueAuthors | 321.751 | 4 | 2.058 |
| Year | 703.692 | 13 | 1.287 |

Residuals from first and last author and team size



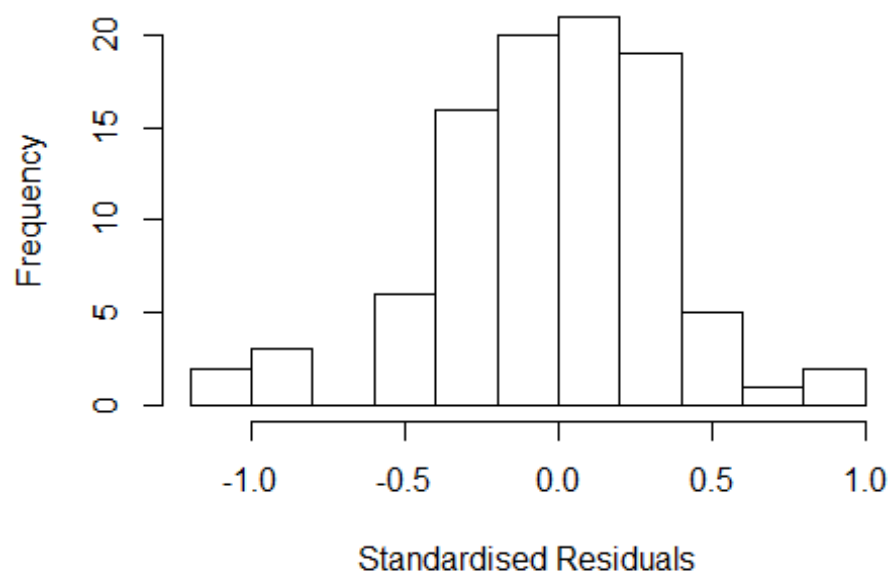
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9480 -0.2272 0.0134 0.2133 1.2068
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.1976 0.4842 0.41 0.68
## FirstAuthorFemale1 0.0520 0.0848 0.61 0.54
## LastAuthorFemale1 0.0651 0.1082 0.60 0.55
## UniqueAuthors2 0.4984 0.4769 1.05 0.30
## UniqueAuthors3 0.6583 0.4337 1.52 0.13
## UniqueAuthors4 0.6338 0.4298 1.47 0.14
## UniqueAuthors5 0.7428 0.4469 1.66 0.10
## Year2000 -0.0286 1.1232 -0.03 0.98
## Year2001 0.1940 0.2845 0.68 0.50
## Year2002 0.0475 0.2303 0.21 0.84
```

```

## Year2003          -0.0494      0.3011    -0.16      0.87
## Year2004           0.1166      0.2647      0.44      0.66
## Year2005           0.1679      0.2219      0.76      0.45
## Year2006           0.1966      0.3087      0.64      0.53
## Year2007           0.2307      0.2067      1.12      0.27
## Year2008           0.2587      0.2042      1.27      0.21
## Year2009           0.0960      0.3060      0.31      0.75
## Year2010           0.0624      0.2191      0.28      0.78
## Year2011           0.1990      0.2285      0.87      0.39
## Year2012           0.1659      0.3066      0.54      0.59
##
## Robust residual standard error: 0.349
## Multiple R-squared:  0.18,   Adjusted R-squared:  -0.0281
## Convergence in 35 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 85 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.209  0.890  0.951  0.899  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      1.05e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.393 1      1.180
## LastAuthorFemale  1.971 1      1.404
## Year              2.317 13      1.033

```

Residuals from first and last author



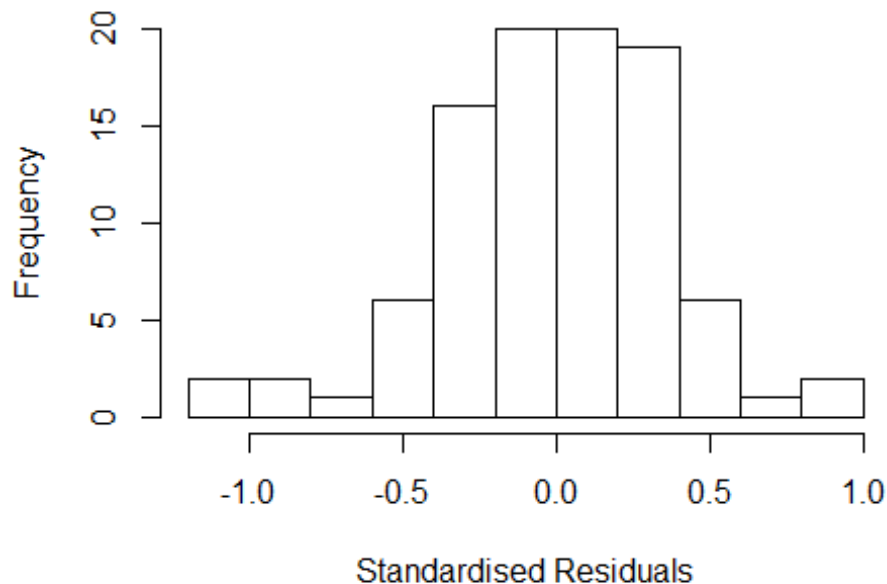
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.122 -0.286 0.017 0.250 0.912
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8595 0.2084 4.12 9.1e-05 ***
## FirstAuthorFemale1 0.0325 0.0835 0.39 0.70
## LastAuthorFemale1 0.0271 0.1075 0.25 0.80
## Year2000 -0.0445 0.9519 -0.05 0.96
## Year2001 -0.0868 0.3801 -0.23 0.82
## Year2002 0.1206 0.2400 0.50 0.62
## Year2003 0.0509 0.2897 0.18 0.86
## Year2004 0.0584 0.2901 0.20 0.84
## Year2005 0.1348 0.2360 0.57 0.57
## Year2006 0.2923 0.2419 1.21 0.23
## Year2007 0.2510 0.2084 1.20 0.23
## Year2008 0.2870 0.2059 1.39 0.17
```

```

## Year2009          0.1029      0.2879      0.36      0.72
## Year2010          0.1173      0.2132      0.55      0.58
## Year2011          0.2211      0.2340      0.94      0.35
## Year2012          0.2302      0.2962      0.78      0.44
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.358
## Multiple R-squared:  0.0731, Adjusted R-squared:  -0.103
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 86 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.306  0.898  0.941  0.897  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      1.05e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.293 1          1.137
## Year              1.293 13          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
## -1.12338 -0.28613 0.00923 0.25283 0.89802
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8563 0.2055 4.17 7.8e-05 ***
## FirstAuthorFemale1 0.0373 0.0809 0.46 0.65
## Year2000 -0.0413 0.9140 -0.05 0.96
## Year2001 -0.0860 0.3740 -0.23 0.82
## Year2002 0.1275 0.2402 0.53 0.60
## Year2003 0.0494 0.2886 0.17 0.86
## Year2004 0.0757 0.2695 0.28 0.78
## Year2005 0.1413 0.2343 0.60 0.55
## Year2006 0.2999 0.2345 1.28 0.20
## Year2007 0.2517 0.2062 1.22 0.23
## Year2008 0.2995 0.1972 1.52 0.13
## Year2009 0.1092 0.2869 0.38 0.70
```

```

## Year2010          0.1215      0.2100      0.58      0.56
## Year2011          0.2273      0.2342      0.97      0.33
## Year2012          0.2298      0.2949      0.78      0.44
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.364
## Multiple R-squared:  0.0707, Adjusted R-squared:  -0.092
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 87 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.320  0.901  0.943   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      1.05e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.82 1          1.349
## Year              1.82 13          1.023
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.109 -0.282  0.024  0.261  0.908
##
## Coefficients:
##      Estimate Std. Error t value Pr(>|t|)

```

```

## (Intercept)          0.8824      0.2059      4.28 5.1e-05 ***
## LastAuthorFemale1    0.0387      0.1019      0.38 0.71
## Year2000             -0.0674      0.9833     -0.07 0.95
## Year2001             -0.0934      0.4069     -0.23 0.82
## Year2002              0.1125      0.2475      0.45 0.65
## Year2003              0.0606      0.2986      0.20 0.84
## Year2004              0.0397      0.2951      0.13 0.89
## Year2005              0.1272      0.2435      0.52 0.60
## Year2006              0.2848      0.2474      1.15 0.25
## Year2007              0.2457      0.2169      1.13 0.26
## Year2008              0.2772      0.2127      1.30 0.20
## Year2009              0.1056      0.2986      0.35 0.72
## Year2010              0.1083      0.2221      0.49 0.63
## Year2011              0.2087      0.2416      0.86 0.39
## Year2012              0.2265      0.2963      0.76 0.45
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.354
## Multiple R-squared:  0.072, Adjusted R-squared:  -0.0904
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 86 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.307  0.896  0.942  0.895  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          1.05e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01          5.00e-01
##   nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500          50          2          1          1000          200
##   trace.lev      mts      compute.rd
##           0          1000          0
##           psi          subsampling          cov
##           "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 95"
## [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
##   33   18    3   34    3   26   40   37   43   62   51   52   55   53   81
## 2011 2012
##   54   80
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   13    9    1   16    1    8   26   25   29   43   30   29   33   35   45
## 2011 2012
##   33   32
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    9    8    0   11    1    5   25   19   23   39   30   25   31   32   39
## 2011 2012
##   29   27
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.0878562434672"
## [1] "Male first author team size 2018 geometric mean: 4.95063192251826"

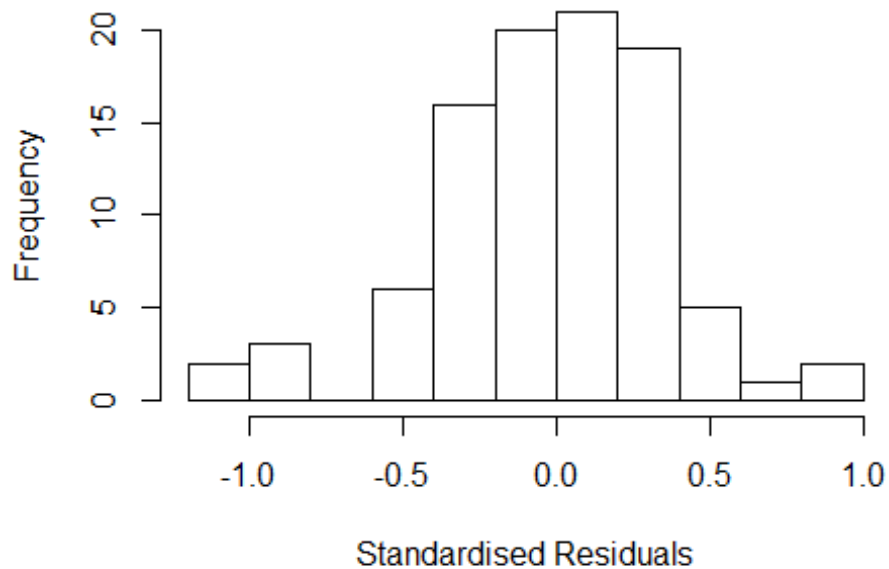
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 170, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.92125491178664"
## [1] "Male last author team size 2018 geometric mean: 4.91443659221379"

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

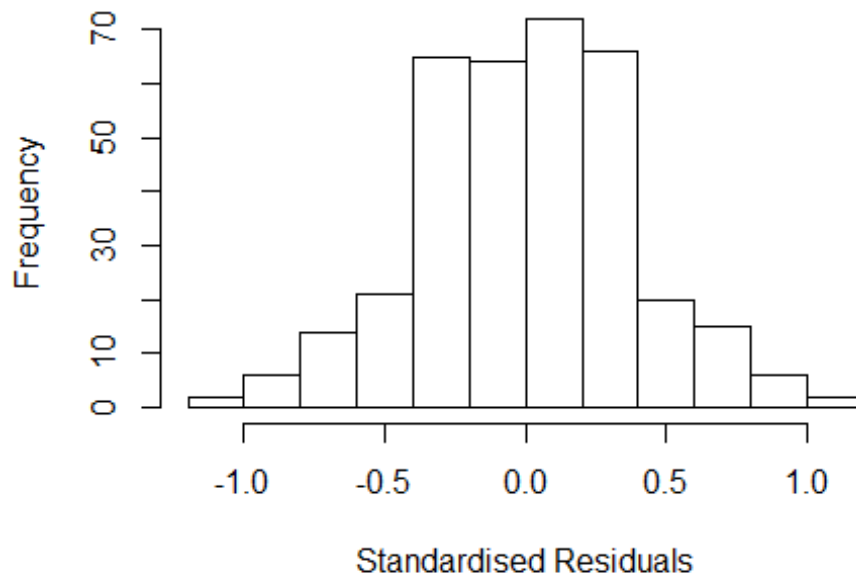
```


Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 160, 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 3.023 1      1.739
## LastAuthorFemale  1.779 1      1.334
## UniqueAuthors    28.273 4      1.519
## Year              42.835 15     1.133
```

Residuals from first and last author and team size



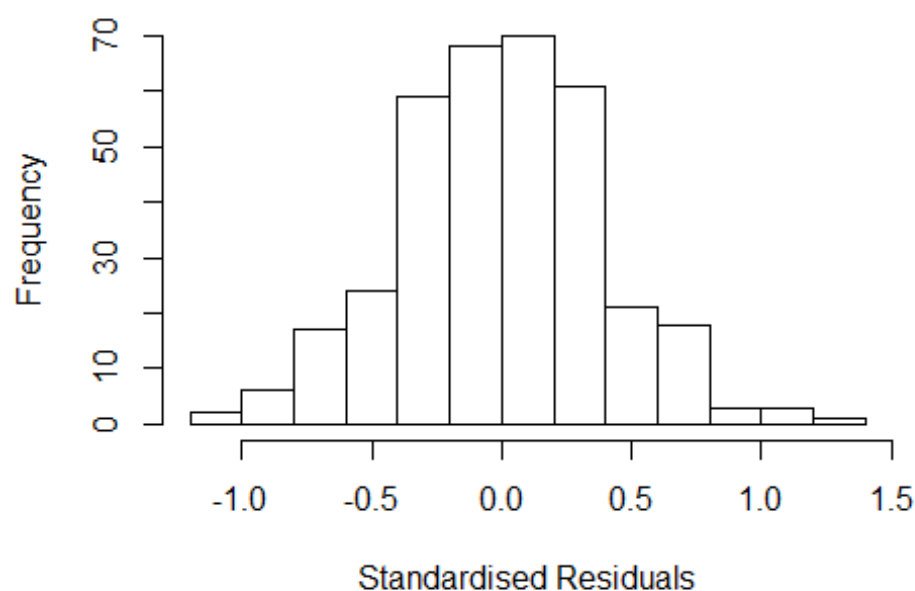
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.06761 -0.26042 0.00853 0.24687 1.12251
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.60036 0.25122 2.39 0.01742 *
## FirstAuthorFemale1 -0.09341 0.04644 -2.01 0.04507 *
## LastAuthorFemale1 0.00326 0.05178 0.06 0.94985
## UniqueAuthors2 0.38952 0.10173 3.83 0.00015 ***
## UniqueAuthors3 0.39281 0.10149 3.87 0.00013 ***
## UniqueAuthors4 0.41997 0.10431 4.03 7.0e-05 ***
## UniqueAuthors5 0.46725 0.09214 5.07 6.6e-07 ***
## Year1997 0.08031 0.28778 0.28 0.78038
## Year1999 0.32077 0.24027 1.34 0.18278
## Year2000 -0.19493 0.21678 -0.90 0.36921
```

```

## Year2001      0.17154      0.24515      0.70  0.48458
## Year2002     -0.08827      0.27889     -0.32  0.75182
## Year2003      0.06514      0.23839      0.27  0.78485
## Year2004     -0.22583      0.25495     -0.89  0.37637
## Year2005      0.03077      0.23135      0.13  0.89429
## Year2006      0.18460      0.23510      0.79  0.43290
## Year2007      0.07821      0.23792      0.33  0.74257
## Year2008      0.14766      0.23768      0.62  0.53487
## Year2009      0.13625      0.23743      0.57  0.56647
## Year2010      0.10321      0.23829      0.43  0.66521
## Year2011      0.10654      0.23935      0.45  0.65653
## Year2012      0.25193      0.23866      1.06  0.29190
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.374
## Multiple R-squared:  0.22,   Adjusted R-squared:  0.171
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 330 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.348  0.894  0.954  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      2.83e-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 2.735 1 1.654
## LastAuthorFemale 1.609 1 1.269
## Year 3.978 15 1.047

```

Residuals from first and last author



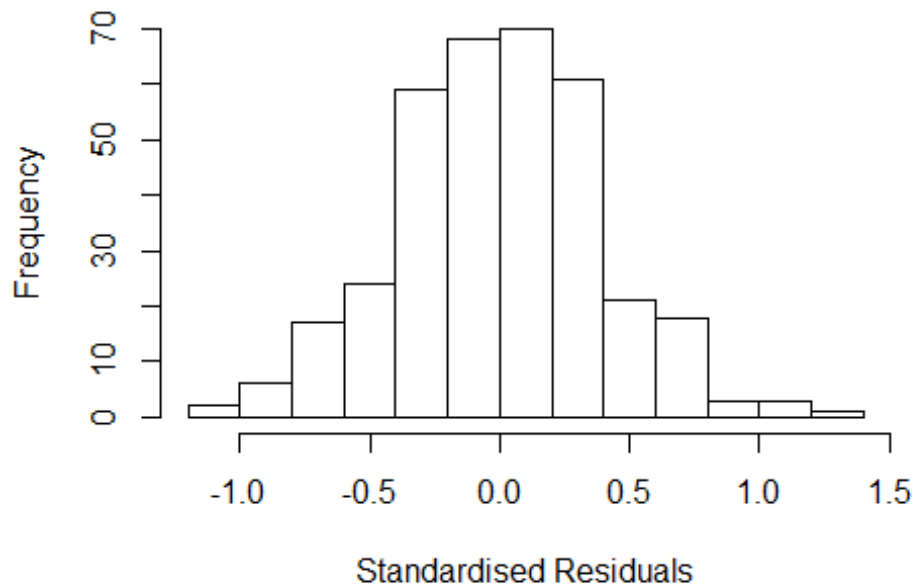
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.094 -0.254 0.010 0.284 1.207
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.99895 0.23014 4.34 1.9e-05 ***
## FirstAuthorFemale1 -0.06503 0.04646 -1.40 0.163
## LastAuthorFemale1 -0.00957 0.05268 -0.18 0.856
## Year1997 0.09553 0.28592 0.33 0.739
## Year1999 0.32551 0.23637 1.38 0.169
## Year2000 -0.20192 0.22506 -0.90 0.370
## Year2001 0.18716 0.24759 0.76 0.450
## Year2002 -0.20654 0.28663 -0.72 0.472
## Year2003 0.02029 0.24481 0.08 0.934
## Year2004 -0.42547 0.24784 -1.72 0.087 .
## Year2005 0.02804 0.23133 0.12 0.904
## Year2006 0.18125 0.23587 0.77 0.443
```

```

## Year2007          0.07201    0.23560    0.31    0.760
## Year2008          0.13893    0.23592    0.59    0.556
## Year2009          0.13388    0.23610    0.57    0.571
## Year2010          0.09978    0.23581    0.42    0.672
## Year2011          0.09268    0.23790    0.39    0.697
## Year2012          0.25627    0.23716    1.08    0.281
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.381
## Multiple R-squared:  0.16,   Adjusted R-squared:  0.117
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 334 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.294  0.878  0.949  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      2.83e-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 2.645 1          1.626
## Year              2.645 15          1.033

```

Residuals from first author



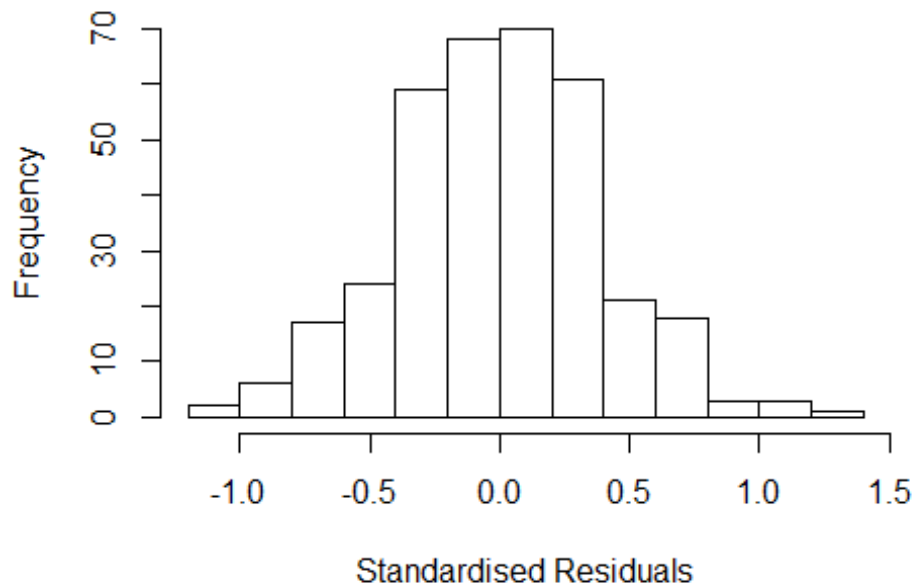
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.09285 -0.25194 0.00576 0.27787 1.20550
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9945 0.2242 4.44 1.2e-05 ***
## FirstAuthorFemale1 -0.0651 0.0457 -1.43 0.15
## Year1997 0.0984 0.2831 0.35 0.73
## Year1999 0.3290 0.2316 1.42 0.16
## Year2000 -0.1974 0.2196 -0.90 0.37
## Year2001 0.1916 0.2425 0.79 0.43
## Year2002 -0.2009 0.2791 -0.72 0.47
## Year2003 0.0226 0.2412 0.09 0.93
## Year2004 -0.4237 0.2414 -1.76 0.08 .
## Year2005 0.0315 0.2261 0.14 0.89
## Year2006 0.1840 0.2318 0.79 0.43
## Year2007 0.0741 0.2315 0.32 0.75
```

```

## Year2008          0.1416      0.2310      0.61      0.54
## Year2009          0.1368      0.2313      0.59      0.55
## Year2010          0.1017      0.2315      0.44      0.66
## Year2011          0.0944      0.2340      0.40      0.69
## Year2012          0.2579      0.2330      1.11      0.27
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.385
## Multiple R-squared:  0.158, Adjusted R-squared:  0.118
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 331 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.307  0.881  0.949  0.905  0.987  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      2.83e-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 1.585 1          1.259
## Year            1.585 15          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
## -1.05659 -0.26782 -0.00678 0.26920 1.14199
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9536 0.2152 4.43 1.3e-05 ***
## LastAuthorFemale1 -0.0198 0.0524 -0.38 0.706
## Year1997 0.1030 0.2842 0.36 0.717
## Year1999 0.3455 0.2261 1.53 0.127
## Year2000 -0.2216 0.2152 -1.03 0.304
## Year2001 0.1940 0.2389 0.81 0.417
## Year2002 -0.1616 0.2668 -0.61 0.545
## Year2003 0.0232 0.2362 0.10 0.922
## Year2004 -0.4045 0.2364 -1.71 0.088 .
## Year2005 0.0492 0.2197 0.22 0.823
## Year2006 0.2035 0.2260 0.90 0.369
## Year2007 0.0801 0.2264 0.35 0.724
```

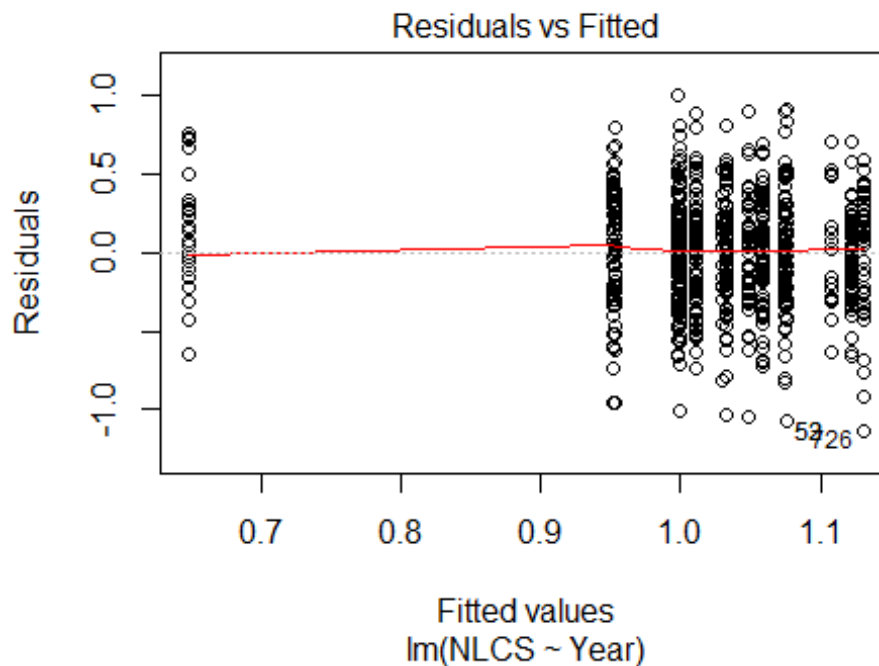


```

## Year2008          0.1604      0.2238      0.72      0.474
## Year2009          0.1522      0.2252      0.68      0.499
## Year2010          0.1164      0.2243      0.52      0.604
## Year2011          0.1122      0.2274      0.49      0.622
## Year2012          0.2644      0.2267      1.17      0.244
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.39
## Multiple R-squared:  0.15,   Adjusted R-squared:  0.109
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 329 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.372  0.881   0.952   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      2.83e-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: 353"
## [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
##   38   30   31   39   44   35   60   61   51   73   66   83   86   91  104
## 2011 2012
##  101  102
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   20   23   16   26   18   19   43   45   35   61   49   58   65   76   81
## 2011 2012
##   73   77

```

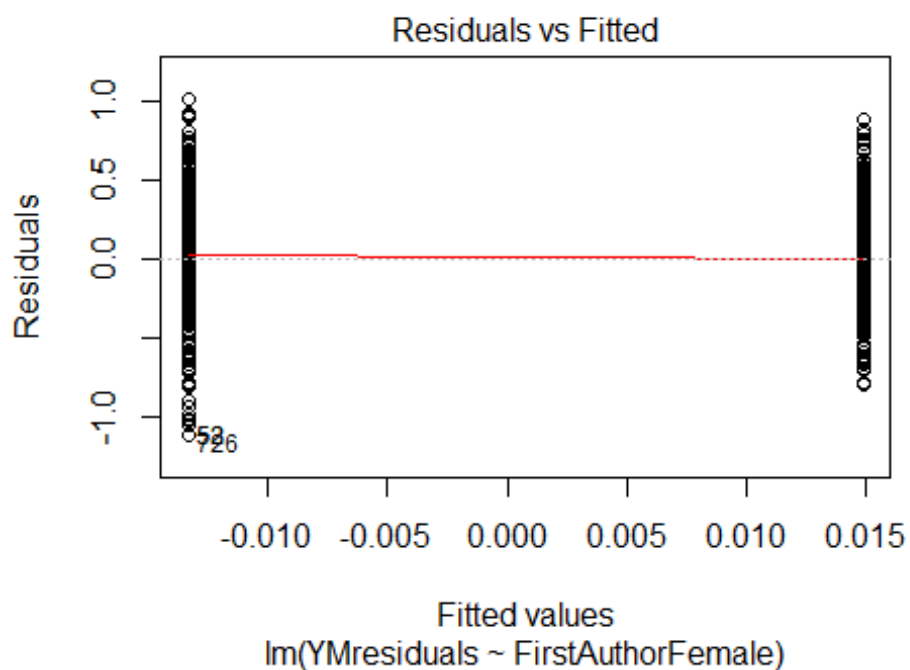
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 14 21 13 21 13 13 36 35 28 51 38 48 57 64 72
## 2011 2012
## 63 63
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS 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 = 8.1, df = 1, p-value = 0.004
## [1] "Female first author team size 2018 geometric mean: 4.94995666967625"
## [1] "Male first author team size 2018 geometric mean: 4.29175812873717"
## Warning in wilcox.test.default(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 = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.56646029525271"
## [1] "Male last author team size 2018 geometric mean: 4.18032194912137"

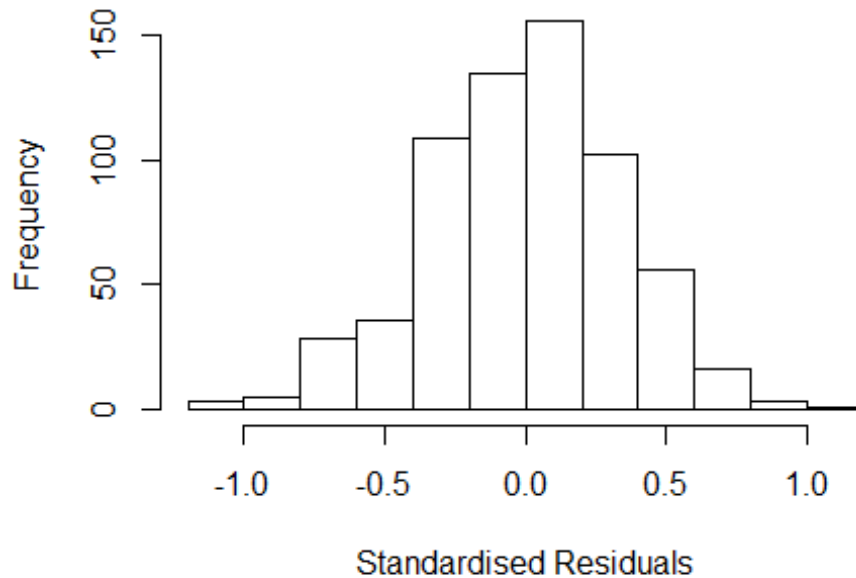
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 300, 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 | 1.141 | 1 | 1.068 |
| LastAuthorFemale | 1.205 | 1 | 1.098 |
| UniqueAuthors | 2.599 | 4 | 1.127 |
| Year | 3.183 | 16 | 1.037 |

Residuals from first and last author and team size



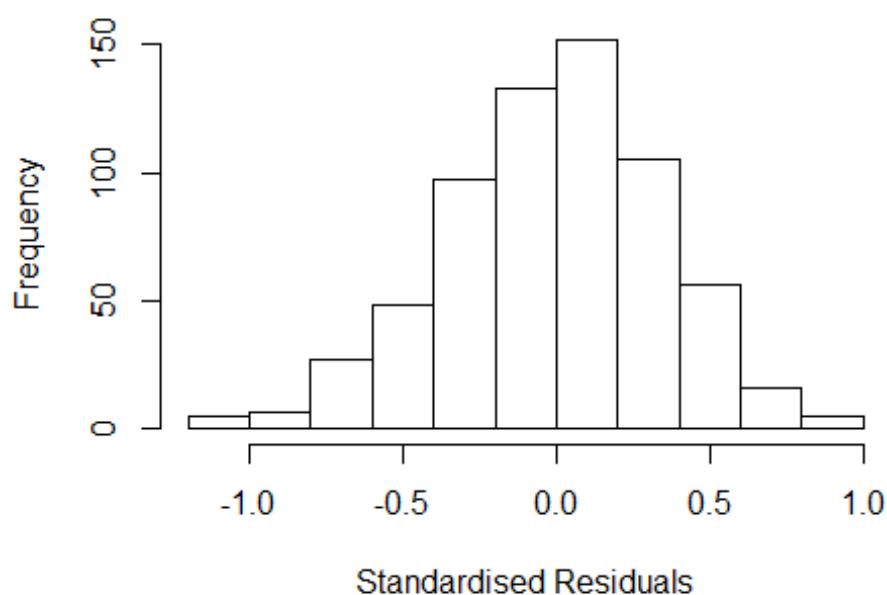
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1168 -0.2222 0.0119 0.2211 1.1577
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.71040 0.12385 5.74 1.5e-08 ***
## FirstAuthorFemale1 -0.01293 0.02805 -0.46 0.6450
## LastAuthorFemale1 0.01255 0.03407 0.37 0.7127
## UniqueAuthors2 0.35570 0.07514 4.73 2.7e-06 ***
## UniqueAuthors3 0.42201 0.07438 5.67 2.1e-08 ***
## UniqueAuthors4 0.44999 0.07487 6.01 3.1e-09 ***
## UniqueAuthors5 0.48739 0.07149 6.82 2.2e-11 ***
## Year1997 0.03038 0.15515 0.20 0.8448
## Year1998 0.03742 0.14851 0.25 0.8011
## Year1999 -0.04199 0.12458 -0.34 0.7362
```

```

## Year2000      -0.14836    0.13533   -1.10    0.2734
## Year2001      0.03753    0.12874    0.29    0.7707
## Year2002     -0.16872    0.12874   -1.31    0.1905
## Year2003     -0.12006    0.11555   -1.04    0.2992
## Year2004     -0.36147    0.12716   -2.84    0.0046 **
## Year2005     -0.14571    0.11173   -1.30    0.1927
## Year2006      0.05073    0.11986    0.42    0.6723
## Year2007     -0.03757    0.11475   -0.33    0.7435
## Year2008      0.00352    0.10808    0.03    0.9740
## Year2009     -0.08786    0.11196   -0.78    0.4329
## Year2010     -0.10071    0.10929   -0.92    0.3571
## Year2011     -0.05867    0.10984   -0.53    0.5934
## Year2012     -0.13641    0.11690   -1.17    0.2437
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.333
## Multiple R-squared:  0.179, Adjusted R-squared:  0.151
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 44 weights are ~= 1. The remaining 606 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.203  0.880  0.955  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      1.54e-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 1.104 1      1.051
## LastAuthorFemale  1.125 1      1.061
## Year              1.233 16      1.007

```

Residuals from first and last author



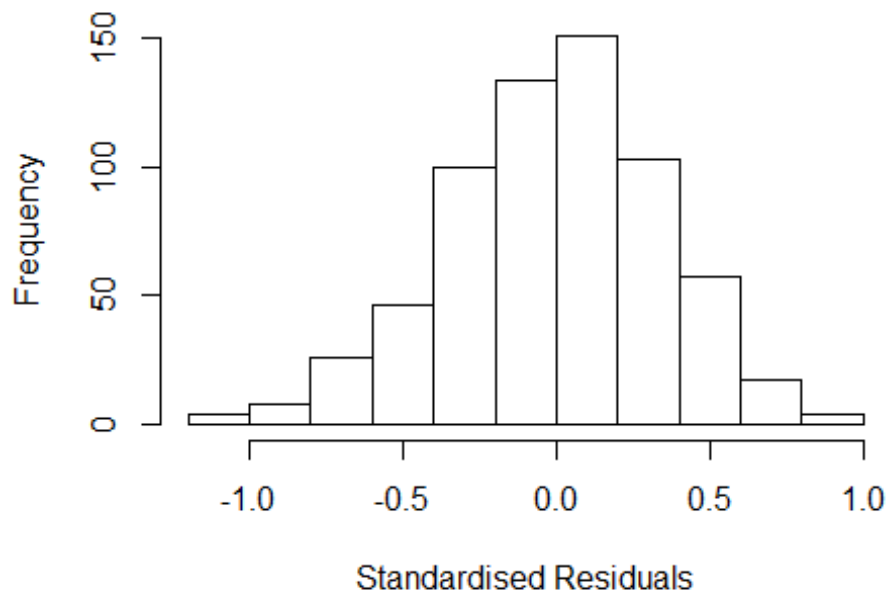
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.17295 -0.23086  0.00562  0.22165  0.89791
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.12023    0.10125   11.06 < 2e-16 ***
## FirstAuthorFemale1  0.00892    0.02895    0.31   0.76
## LastAuthorFemale1  0.02103    0.03527    0.60   0.55
## Year1997         0.01081    0.15878    0.07   0.95
## Year1998        -0.04705    0.14688   -0.32   0.75
## Year1999        -0.06000    0.12140   -0.49   0.62
## Year2000        -0.16414    0.13819   -1.19   0.24
## Year2001         0.01356    0.13457    0.10   0.92
## Year2002        -0.21426    0.14245   -1.50   0.13
## Year2003        -0.15549    0.11863   -1.31   0.19
## Year2004        -0.54928    0.12609   -4.36 1.5e-05 ***
## Year2005        -0.14101    0.11087   -1.27   0.20
```

```

## Year2006          0.05272    0.12380    0.43    0.67
## Year2007         -0.05948    0.11344   -0.52    0.60
## Year2008          0.00652    0.10627    0.06    0.95
## Year2009         -0.08444    0.10997   -0.77    0.44
## Year2010         -0.09558    0.10771   -0.89    0.38
## Year2011         -0.04714    0.10901   -0.43    0.67
## Year2012         -0.12425    0.11102   -1.12    0.26
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.338
## Multiple R-squared:  0.099, Adjusted R-squared:  0.0733
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 51 weights are ~= 1. The remaining 599 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.203  0.867  0.953  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      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 1.102 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
## -1.17401 -0.23204 0.00398 0.22490 0.89169
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.12486 0.09907 11.35 < 2e-16 ***
## FirstAuthorFemale1 0.01142 0.02898 0.39 0.69
## Year1997 0.00814 0.15794 0.05 0.96
## Year1998 -0.04598 0.14628 -0.31 0.75
## Year1999 -0.06252 0.11974 -0.52 0.60
## Year2000 -0.16710 0.13573 -1.23 0.22
## Year2001 0.01521 0.13464 0.11 0.91
## Year2002 -0.21368 0.14224 -1.50 0.13
## Year2003 -0.15574 0.11759 -1.32 0.19
## Year2004 -0.54558 0.12546 -4.35 1.6e-05 ***
## Year2005 -0.14325 0.10951 -1.31 0.19
## Year2006 0.04915 0.12218 0.40 0.69
```

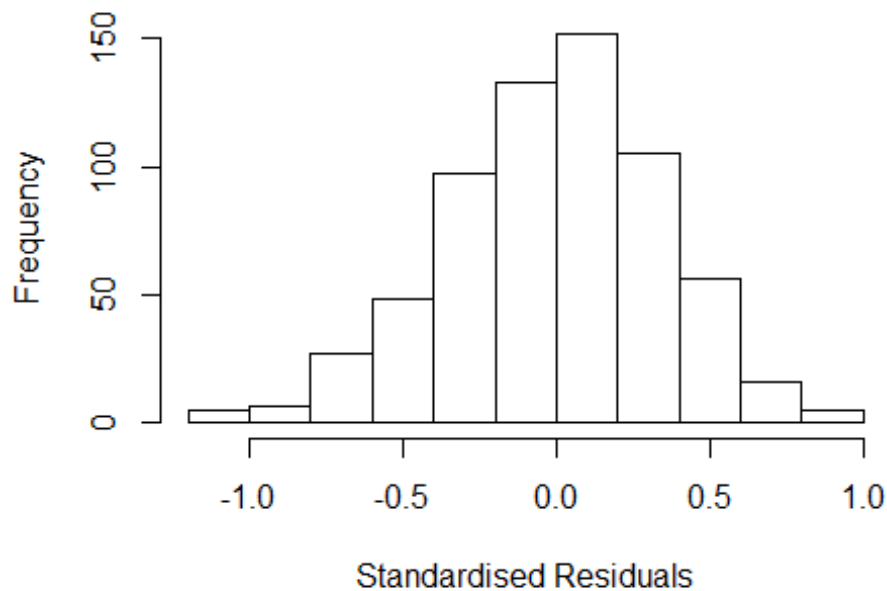


```

## Year2007      -0.06124    0.11213   -0.55    0.59
## Year2008      0.00693    0.10510    0.07    0.95
## Year2009     -0.08764    0.10832   -0.81    0.42
## Year2010     -0.09662    0.10648   -0.91    0.36
## Year2011     -0.04555    0.10817   -0.42    0.67
## Year2012     -0.12689    0.10944   -1.16    0.25
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.339
## Multiple R-squared:  0.0984, Adjusted R-squared:  0.0741
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 50 weights are ~= 1. The remaining 600 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.205  0.868  0.952  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      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 1.123 1      1.060
## Year      1.123 16      1.004

```

Residuals from last author



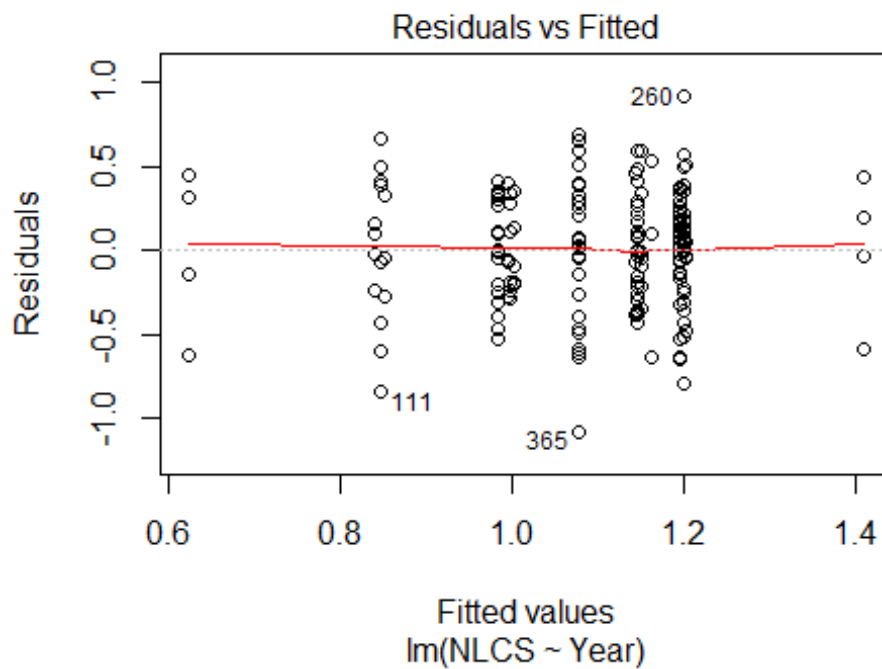
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.17666 -0.23111 0.00603 0.22002 0.89474
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.12326 0.10149 11.07 < 2e-16 ***
## LastAuthorFemale1 0.02243 0.03531 0.64 0.53
## Year1997 0.01366 0.15710 0.09 0.93
## Year1998 -0.04597 0.14709 -0.31 0.75
## Year1999 -0.05885 0.12119 -0.49 0.63
## Year2000 -0.16248 0.13698 -1.19 0.24
## Year2001 0.01571 0.13486 0.12 0.91
## Year2002 -0.21433 0.14301 -1.50 0.13
## Year2003 -0.15375 0.11830 -1.30 0.19
## Year2004 -0.54852 0.12598 -4.35 1.6e-05 ***
## Year2005 -0.13972 0.11081 -1.26 0.21
## Year2006 0.05340 0.12387 0.43 0.67
```

```

## Year2007          -0.05900      0.11360     -0.52      0.60
## Year2008           0.00791      0.10621      0.07      0.94
## Year2009          -0.08244      0.10959     -0.75      0.45
## Year2010          -0.09387      0.10763     -0.87      0.38
## Year2011          -0.04701      0.10923     -0.43      0.67
## Year2012          -0.12390      0.11128     -1.11      0.27
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.337
## Multiple R-squared:  0.099, Adjusted R-squared:  0.0748
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 51 weights are ~= 1. The remaining 599 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.198  0.867  0.953  0.898  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      1.54e-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 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
##    4    4    4    4    7   10    5    9    6    6    9   13   32   24   35
## 2011 2012
##   43   34
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    4    4    3    4    5    3    8    5    4    6    8   27   19   26
## 2011 2012

```

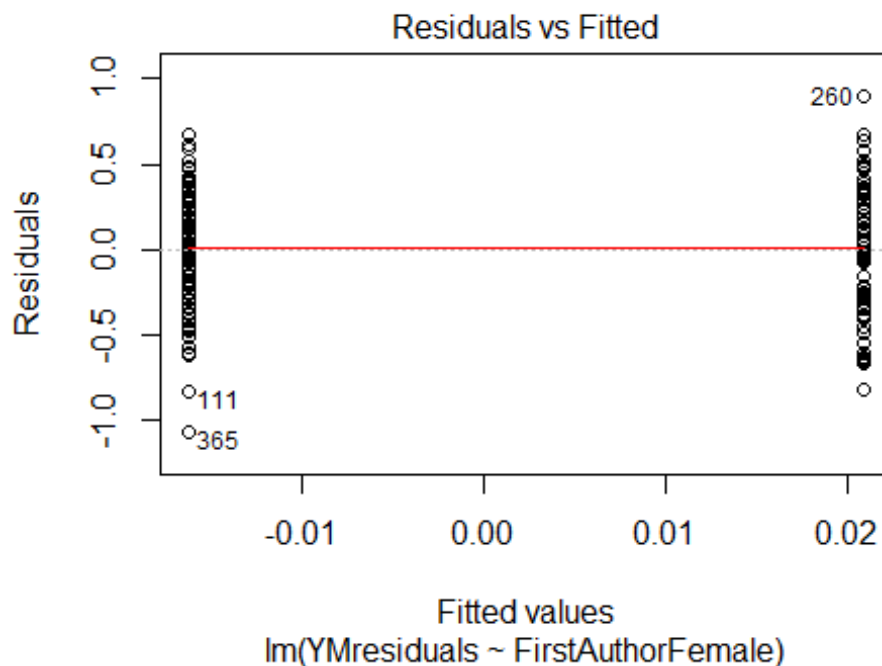
```
## 30 26
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3 3 2 3 4 5 2 7 5 4 6 6 24 16 25
## 2011 2012
## 28 21
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by 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 = 1.8, df = 1, p-value = 0.2
## [1] "Female first author team size 2018 geometric mean: 5.50697406030785"
## [1] "Male first author team size 2018 geometric mean: 5.45926032159334"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 31, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.33433655067999"
## [1] "Male last author team size 2018 geometric mean: 5.5593297457259"

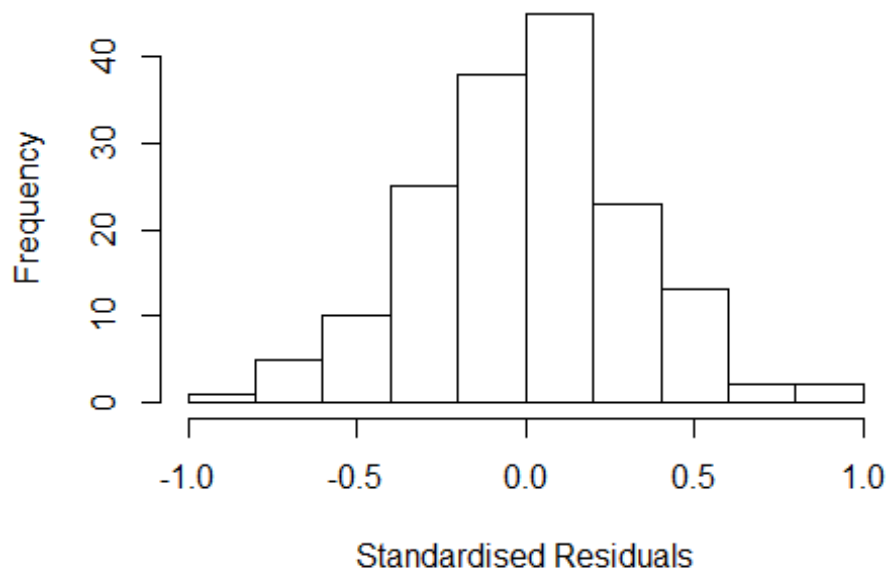
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 27, 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 | 2.368 | 1 | 1.539 |
| ## LastAuthorFemale | 2.893 | 1 | 1.701 |
| ## UniqueAuthors | 28.819 | 4 | 1.522 |
| ## Year | 64.641 | 16 | 1.139 |

Residuals from first and last author and team size



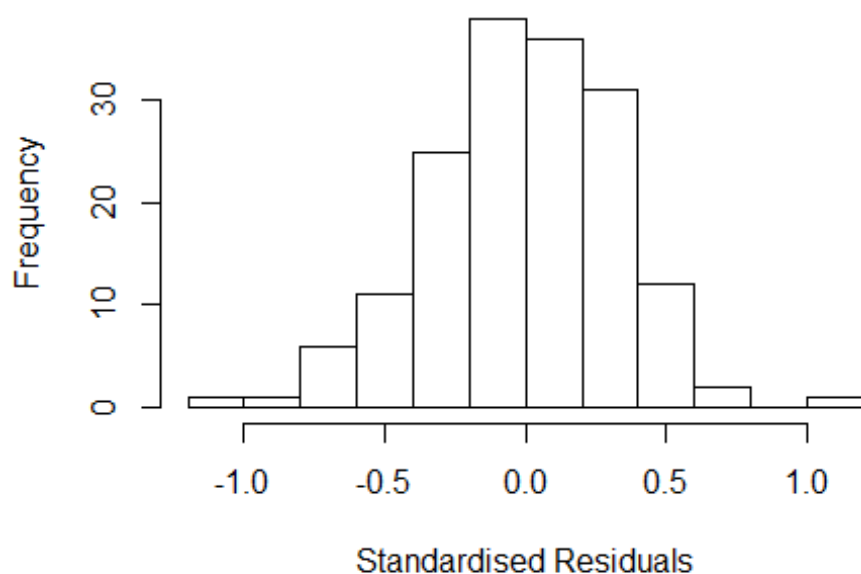
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8321 -0.1914 0.0119 0.1927 0.9925
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5080 0.2441 2.08 0.03926 *
## FirstAuthorFemale1 -0.0299 0.0643 -0.47 0.64214
## LastAuthorFemale1 0.1146 0.0703 1.63 0.10540
## UniqueAuthors2 0.1993 0.2626 0.76 0.44914
## UniqueAuthors3 0.3170 0.2221 1.43 0.15580
## UniqueAuthors4 0.3722 0.2524 1.47 0.14252
## UniqueAuthors5 0.3899 0.2242 1.74 0.08422 .
## Year1997 -0.2412 0.2342 -1.03 0.30476
## Year1998 0.8988 0.1386 6.49 1.4e-09 ***
## Year1999 0.4336 0.4256 1.02 0.31014
```

```

## Year2000          0.1738      0.1801      0.97  0.33619
## Year2001          0.3508      0.1607      2.18  0.03066 *
## Year2002          0.0685      0.1059      0.65  0.51869
## Year2003          0.3234      0.1813      1.78  0.07661 .
## Year2004          0.2033      0.1545      1.32  0.19052
## Year2005         -0.0487      0.1508     -0.32  0.74724
## Year2006          0.1322      0.1546      0.85  0.39425
## Year2007         -0.0858      0.3088     -0.28  0.78159
## Year2008          0.4195      0.1151      3.64  0.00038 ***
## Year2009          0.1326      0.1318      1.01  0.31612
## Year2010          0.2205      0.1171      1.88  0.06178 .
## Year2011          0.2545      0.1221      2.08  0.03891 *
## Year2012          0.2840      0.1661      1.71  0.08943 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.307
## Multiple R-squared:  0.278, Adjusted R-squared:  0.166
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 150 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.276  0.890  0.947  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      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 1.738 1      1.318
## LastAuthorFemale  1.484 1      1.218
## Year              2.478 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
## -1.13309 -0.21628 0.00149 0.21009 1.01172
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8160 0.1249 6.53 1.0e-09 ***
## FirstAuthorFemale1 -0.0083 0.0626 -0.13 0.8946
## LastAuthorFemale1 0.1048 0.0610 1.72 0.0880 .
## Year1997 -0.3130 0.3763 -0.83 0.4069
## Year1998 0.9138 0.1655 5.52 1.5e-07 ***
## Year1999 0.4139 0.4607 0.90 0.3705
## Year2000 0.1507 0.1654 0.91 0.3639
## Year2001 0.3572 0.1968 1.81 0.0716 .
## Year2002 0.0493 0.1479 0.33 0.7392
## Year2003 0.3468 0.1725 2.01 0.0463 *
## Year2004 0.1391 0.1613 0.86 0.3900
## Year2005 0.0020 0.1534 0.01 0.9896
```

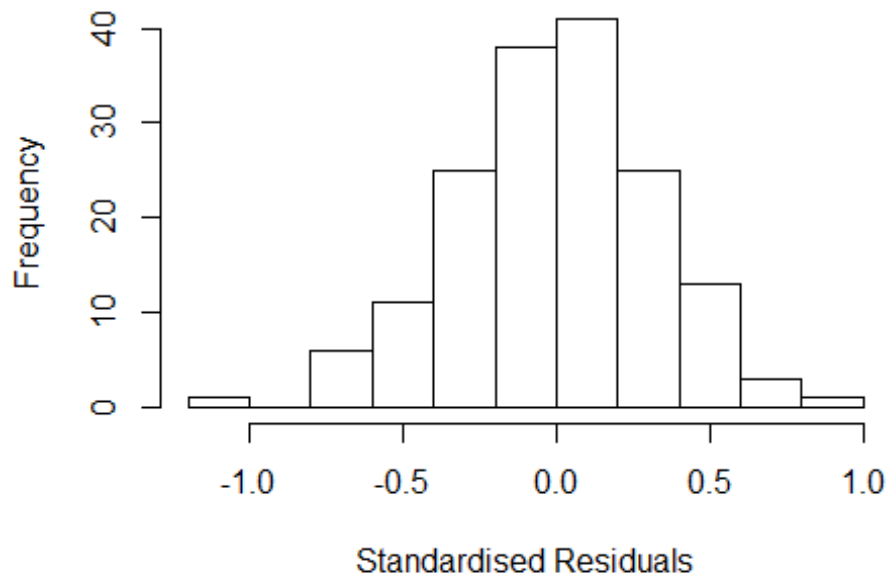


```

## Year2006          0.1873      0.1753      1.07      0.2871
## Year2007         -0.0465      0.4605     -0.10      0.9197
## Year2008          0.4175      0.1379      3.03      0.0029 **
## Year2009          0.1840      0.1485      1.24      0.2172
## Year2010          0.2854      0.1364      2.09      0.0382 *
## Year2011          0.2956      0.1393      2.12      0.0356 *
## Year2012          0.3171      0.1767      1.79      0.0748 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.32
## Multiple R-squared:  0.204, Adjusted R-squared:  0.105
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 143 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##   0.184  0.889  0.949  0.899  0.986  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      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 1.66 1      1.289
## Year              1.66 16      1.016

```

Residuals from first author



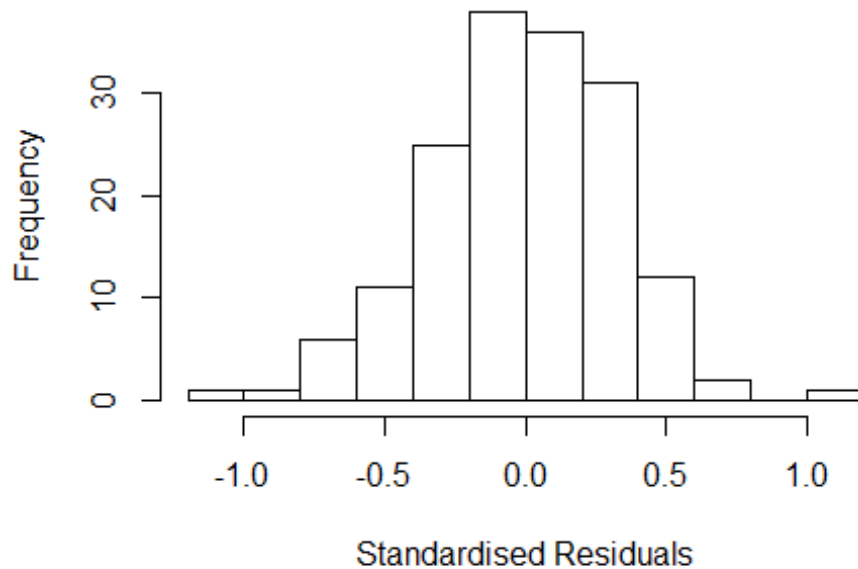
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.13753 -0.22007  0.00496  0.20674  0.95132
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.84756    0.15604    5.43 2.3e-07 ***
## FirstAuthorFemale1 -0.00268    0.06290   -0.04  0.966
## Year1997       -0.34383    0.38237   -0.90  0.370
## Year1998        0.87663    0.19135    4.58 9.8e-06 ***
## Year1999        0.37329    0.46069    0.81  0.419
## Year2000        0.13760    0.20391    0.67  0.501
## Year2001        0.35138    0.21599    1.63  0.106
## Year2002        0.06729    0.19596    0.34  0.732
## Year2003        0.33074    0.19107    1.73  0.086 .
## Year2004        0.15149    0.18641    0.81  0.418
## Year2005       -0.00485    0.17752   -0.03  0.978
## Year2006        0.15105    0.19987    0.76  0.451
```

```

## Year2007      -0.05407    0.50857   -0.11    0.915
## Year2008      0.40108    0.16871    2.38    0.019 *
## Year2009      0.16282    0.17478    0.93    0.353
## Year2010      0.28099    0.16588    1.69    0.092 .
## Year2011      0.31880    0.16884    1.89    0.061 .
## Year2012      0.28997    0.20125    1.44    0.152
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.325
## Multiple R-squared:  0.189, Adjusted R-squared:  0.094
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 147 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.195  0.877   0.951   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      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 1.436 1      1.199
## Year      1.436 16      1.011

```

Residuals from last author



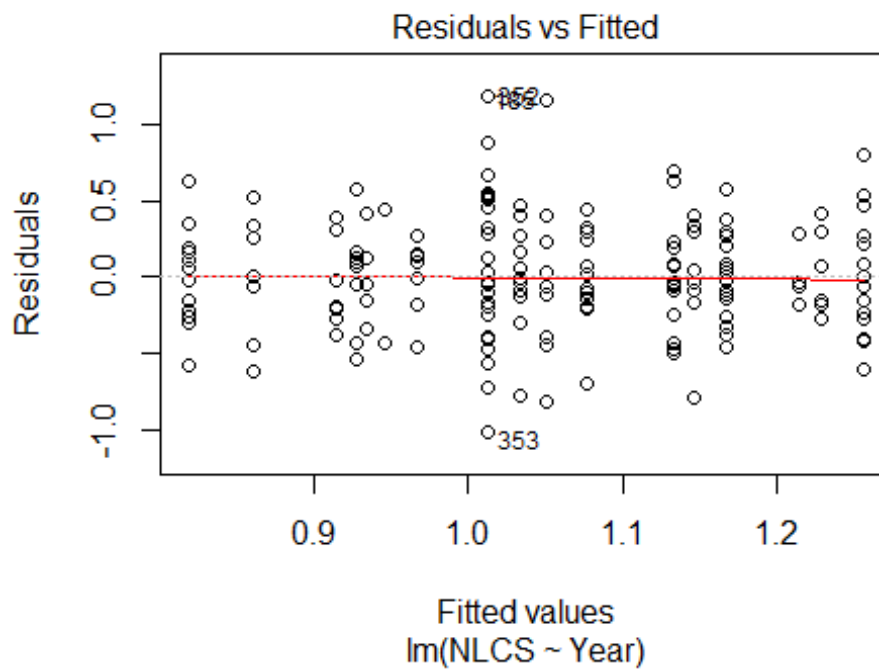
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.130547 -0.218085 0.000235 0.211623 1.006011
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.81596 0.12501 6.53 1.0e-09 ***
## LastAuthorFemale1 0.10476 0.06150 1.70 0.0906 .
## Year1997 -0.31338 0.37838 -0.83 0.4089
## Year1998 0.90554 0.15394 5.88 2.6e-08 ***
## Year1999 0.40700 0.43932 0.93 0.3557
## Year2000 0.14829 0.16406 0.90 0.3675
## Year2001 0.35421 0.19674 1.80 0.0739 .
## Year2002 0.04516 0.14617 0.31 0.7578
## Year2003 0.34644 0.17270 2.01 0.0467 *
## Year2004 0.13598 0.16022 0.85 0.3975
## Year2005 -0.00215 0.14912 -0.01 0.9885
## Year2006 0.18026 0.16873 1.07 0.2871
```

```

## Year2007          -0.05014      0.46943    -0.11    0.9151
## Year2008           0.41330      0.13334      3.10    0.0023 **
## Year2009           0.18083      0.14626      1.24    0.2183
## Year2010           0.28181      0.13680      2.06    0.0412 *
## Year2011           0.29303      0.13968      2.10    0.0376 *
## Year2012           0.31459      0.17577      1.79    0.0756 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.317
## Multiple R-squared:  0.205, Adjusted R-squared:  0.112
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 145 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.178  0.887  0.950  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      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 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
##   14   15   17   20   22   15   14   15   14   13   21   16   23   16   21
## 2011 2012
##   32   31
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    5    8    7   12    4    8    7   12    9   10    8   16   12   15
## 2011 2012

```

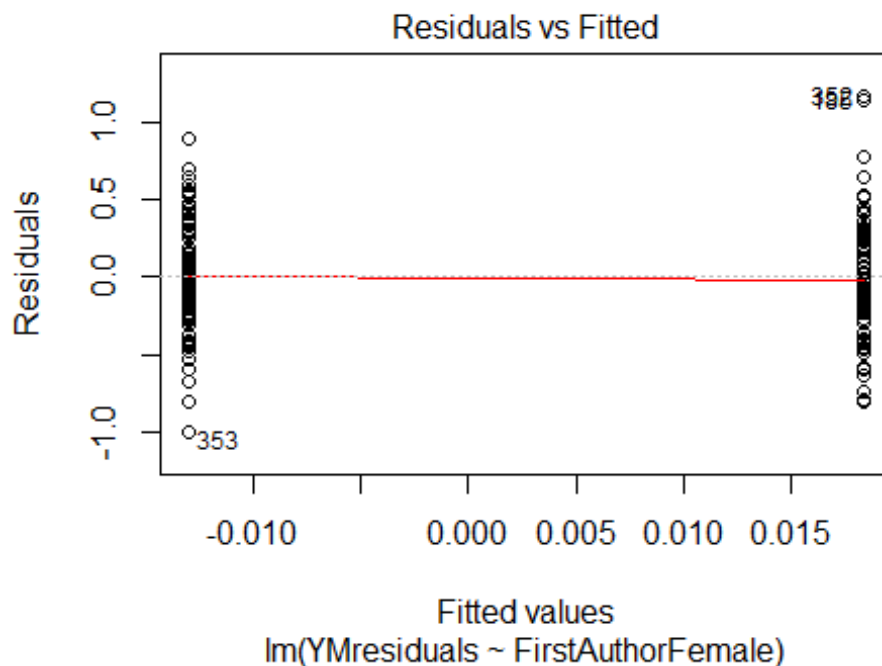
```
## 30 18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2 5 7 7 9 4 8 6 8 9 8 7 14 9 9
## 2011 2012
## 24 17
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 22, df = 16, p-value = 0.1
```



```
##
## 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: 2.28942848510666"
## [1] "Male first author team size 2018 geometric mean: 3.18200494355336"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

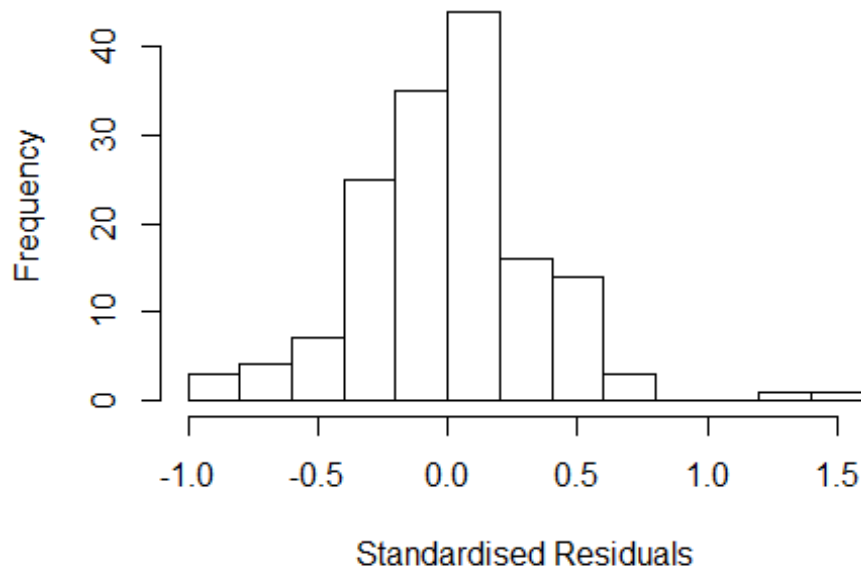
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 24, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.49146187923103"
## [1] "Male last author team size 2018 geometric mean: 3.21970762189899"

## Warning in wilcox.test.default(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] "Regression 1: First author gender, last author gender, team size,
## Year as factors"
##
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.293 1      1.514
## LastAuthorFemale  2.084 1      1.443
## UniqueAuthors    7.116 4      1.278
## Year             20.970 16     1.100
```

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
## -0.83070 -0.20190 0.00835 0.17933 1.48759
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8185 0.5629 1.45 0.148
## FirstAuthorFemale1 -0.0165 0.0669 -0.25 0.806
## LastAuthorFemale1 0.1747 0.1003 1.74 0.084 .
## UniqueAuthors2 0.1214 0.1636 0.74 0.459
## UniqueAuthors3 0.0121 0.1621 0.07 0.941
## UniqueAuthors4 0.1645 0.1821 0.90 0.368
## UniqueAuthors5 0.2849 0.1515 1.88 0.062 .
## Year1997 0.0347 0.5505 0.06 0.950
## Year1998 -0.1132 0.5526 -0.20 0.838
## Year1999 0.2332 0.5542 0.42 0.675
```

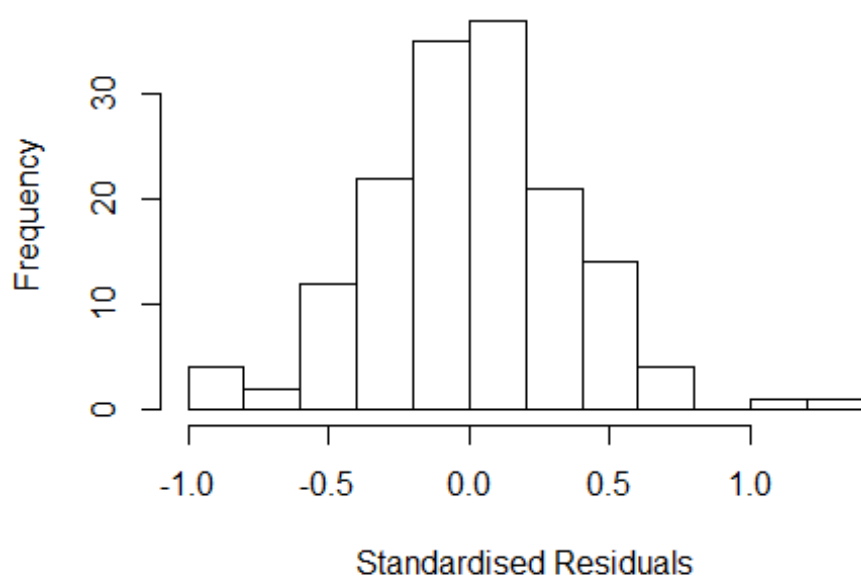


```

## Year2000          0.1206      0.5477      0.22      0.826
## Year2001          0.1949      0.5506      0.35      0.724
## Year2002         -0.0417      0.5556     -0.07      0.940
## Year2003          0.0390      0.5513      0.07      0.944
## Year2004         -0.2088      0.5479     -0.38      0.704
## Year2005         -0.0948      0.5708     -0.17      0.868
## Year2006          0.0853      0.5573      0.15      0.879
## Year2007         -0.1941      0.5617     -0.35      0.730
## Year2008          0.0849      0.5490      0.15      0.877
## Year2009         -0.0453      0.5552     -0.08      0.935
## Year2010          0.3328      0.5545      0.60      0.549
## Year2011          0.0122      0.5537      0.02      0.983
## Year2012          0.1392      0.5472      0.25      0.800
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.326
## Multiple R-squared:  0.249, Adjusted R-squared:  0.122
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 139 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0024 0.8740 0.9540 0.9010 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      6.54e-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 1.712 1      1.308
## LastAuthorFemale 2.100 1      1.449
## Year      3.107 16      1.036

```

Residuals from first and last author



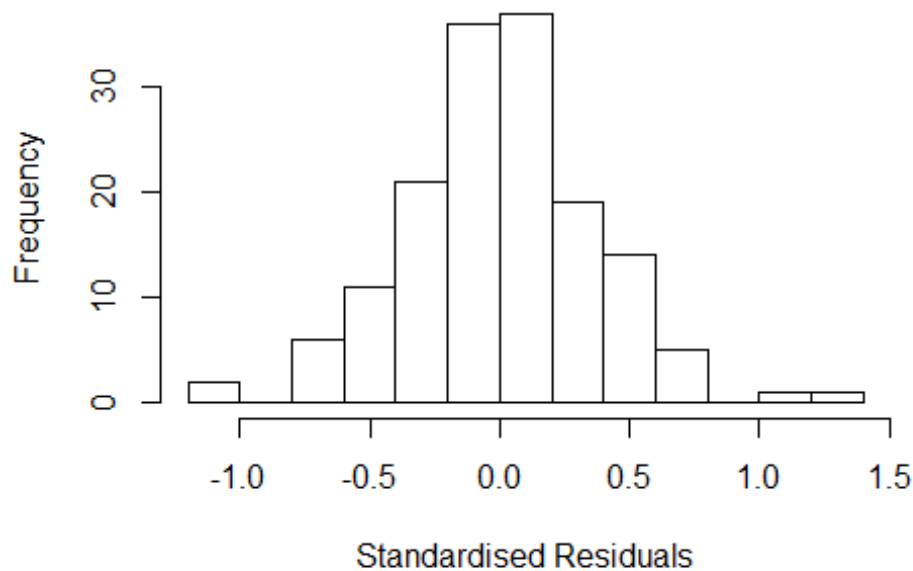
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.98620 -0.20920 0.00338 0.22280 1.25241
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.92142 0.46475 1.98 0.049 *
## FirstAuthorFemale1 0.02358 0.06548 0.36 0.719
## LastAuthorFemale1 0.13249 0.10472 1.27 0.208
## Year1997 -0.01160 0.47593 -0.02 0.981
## Year1998 -0.14095 0.47530 -0.30 0.767
## Year1999 0.26871 0.47538 0.57 0.573
## Year2000 0.18031 0.46905 0.38 0.701
## Year2001 0.25501 0.47801 0.53 0.595
## Year2002 -0.03480 0.47637 -0.07 0.942
## Year2003 0.05304 0.47346 0.11 0.911
## Year2004 -0.18320 0.47025 -0.39 0.697
## Year2005 0.00959 0.49584 0.02 0.985
```

```

## Year2006          0.17037      0.47139      0.36      0.718
## Year2007          -0.12881     0.48577     -0.27     0.791
## Year2008          0.15320     0.46944     0.33     0.745
## Year2009          0.02331     0.47764     0.05     0.961
## Year2010          0.35483     0.47944     0.74     0.461
## Year2011          0.06478     0.47791     0.14     0.892
## Year2012          0.18262     0.46624     0.39     0.696
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.341
## Multiple R-squared:  0.16, Adjusted R-squared:  0.0474
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 138 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.148  0.864  0.959  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      6.54e-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 1.639 1      1.280
## Year              1.639 16      1.016

```

Residuals from first author



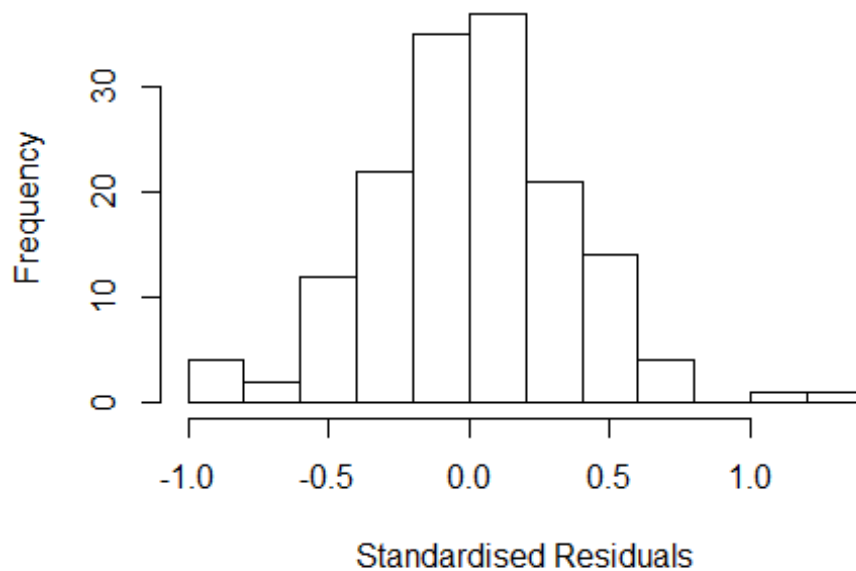
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.00087 -0.22315 0.00601 0.21105 1.20378
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.89600 0.45795 1.96 0.052 .
## FirstAuthorFemale1 0.04900 0.06449 0.76 0.449
## Year1997 -0.00550 0.46948 -0.01 0.991
## Year1998 -0.06805 0.46300 -0.15 0.883
## Year1999 0.30307 0.46718 0.65 0.518
## Year2000 0.19715 0.46311 0.43 0.671
## Year2001 0.31532 0.46650 0.68 0.500
## Year2002 -0.00115 0.47125 0.00 0.998
## Year2003 0.07384 0.46634 0.16 0.874
## Year2004 -0.15542 0.46170 -0.34 0.737
## Year2005 0.05822 0.49234 0.12 0.906
## Year2006 0.18237 0.46606 0.39 0.696
```

```

## Year2007          -0.09814    0.48216   -0.20    0.839
## Year2008          0.16995    0.46423    0.37    0.715
## Year2009          0.07407    0.46971    0.16    0.875
## Year2010          0.39599    0.47306    0.84    0.404
## Year2011          0.10486    0.47072    0.22    0.824
## Year2012          0.21805    0.45796    0.48    0.635
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.346
## Multiple R-squared:  0.143, Adjusted R-squared:  0.0349
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 130 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.201  0.860  0.950  0.896  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      6.54e-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 1.906 1          1.381
## Year            1.906 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
## -0.99540 -0.20287 0.00257 0.21413 1.27662
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.94500 0.45891 2.06 0.041 *
## LastAuthorFemale1 0.14030 0.09943 1.41 0.161
## Year1997 -0.01714 0.47516 -0.04 0.971
## Year1998 -0.15379 0.47395 -0.32 0.746
## Year1999 0.25461 0.47234 0.54 0.591
## Year2000 0.16491 0.46591 0.35 0.724
## Year2001 0.22943 0.47118 0.49 0.627
## Year2002 -0.05335 0.47258 -0.11 0.910
## Year2003 0.03335 0.46994 0.07 0.944
## Year2004 -0.19386 0.46840 -0.41 0.680
## Year2005 -0.01462 0.48896 -0.03 0.976
## Year2006 0.15355 0.46786 0.33 0.743
```

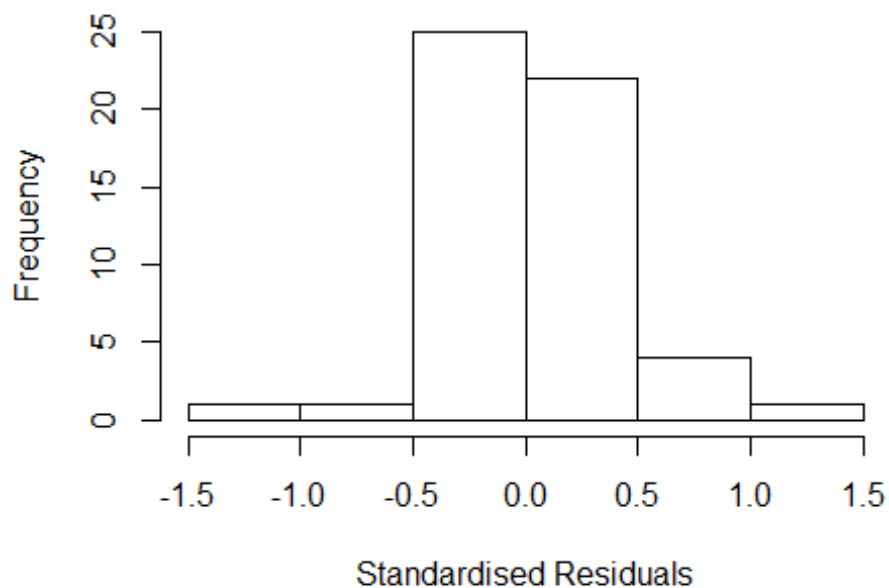
```

## Year2007          -0.13931      0.48231    -0.29      0.773
## Year2008           0.13887      0.46797      0.30      0.767
## Year2009           0.00157      0.46943      0.00      0.997
## Year2010           0.34765      0.47919      0.73      0.469
## Year2011           0.05040      0.47426      0.11      0.916
## Year2012           0.17043      0.46460      0.37      0.714
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.341
## Multiple R-squared:  0.16,   Adjusted R-squared:  0.0547
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 139 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.131  0.865  0.957   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      6.54e-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: 153"
## [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
##    3    2    3    2    1    1    3    7    5    2    5    5    7    8    12
## 2011 2012
##    8    7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    2    1    1    0    1    3    6    3    2    5    5    3    7    9
## 2011 2012

```

```
##      6      6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      1      0      1      0      1      3      2      3      2      5      5      3      7      9
## 2011 2012
##      6      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: 2.44948974278318"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -1.182e+15  1          NaN
## LastAuthorFemale  3.817e+02  1          19.54
## UniqueAuthors    -6.624e+30  4          NaN
## Year              -1.677e+33 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,
```



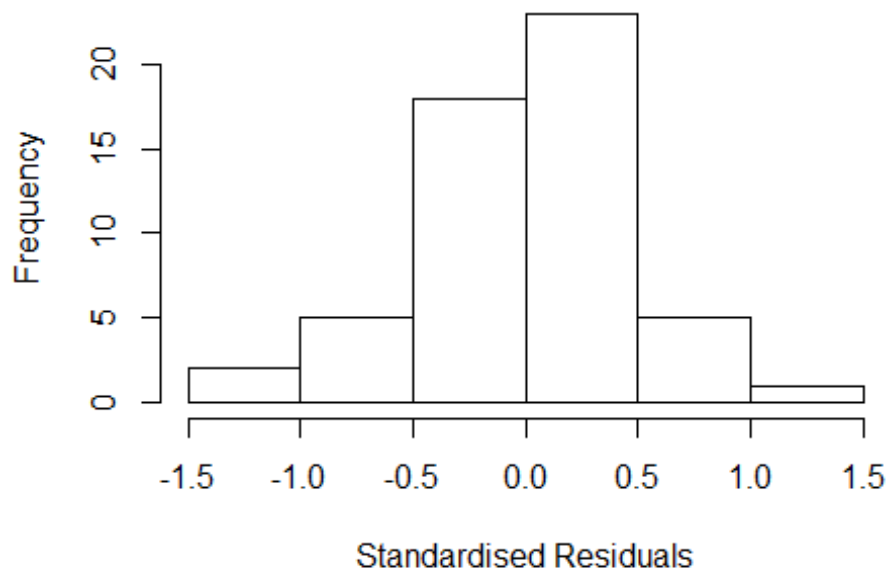
```

##      k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q      Median        3Q        Max
## -1.12120 -0.26198  0.00143  0.21250  1.15650
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.7591      0.5079   1.49  0.14452
## FirstAuthorFemale1 -0.3257      0.1818  -1.79  0.08233 .
## LastAuthorFemale1  0.1123      0.1495   0.75  0.45803
## UniqueAuthors2     0.4440      0.4495   0.99  0.33046
## UniqueAuthors3     0.4665      0.5187   0.90  0.37496
## UniqueAuthors4     0.6845      0.5662   1.21  0.23528
## UniqueAuthors5     0.2100      0.5596   0.38  0.70991
## Year1997           0.7819      0.1941   4.03  0.00031 ***
## Year1999          -0.4476      0.2417  -1.85  0.07297 .
## Year2001           0.0791      0.3142   0.25  0.80277
## Year2002          -0.2924      0.3072  -0.95  0.34812
## Year2003          -0.0420      0.2042  -0.21  0.83810
## Year2004          -0.0744      0.3344  -0.22  0.82532
## Year2005           0.1164      0.3286   0.35  0.72540
## Year2006           0.2438      0.5377   0.45  0.65328
## Year2007          -0.2501      0.2560  -0.98  0.33578
## Year2008          -0.2253      0.2796  -0.81  0.42626
## Year2009          -0.2120      0.1562  -1.36  0.18387
## Year2010          -0.0196      0.2643  -0.07  0.94142
## Year2011          -0.3152      0.2122  -1.48  0.14704
## Year2012           0.0198      0.1849   0.11  0.91550
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.45
## Multiple R-squared:  0.355, Adjusted R-squared:  -0.0363
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 44 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.489  0.906  0.965  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      1.85e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##      trace.lev      mts      compute.rd

```

```
##          0          1000          0
##          psi          subsampling          cov
##          "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##          "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -8.525e+26 1 NaN
## LastAuthorFemale -7.824e+12 1 NaN
## Year 2.445e+15 14 3.545
```

Residuals from first and last author



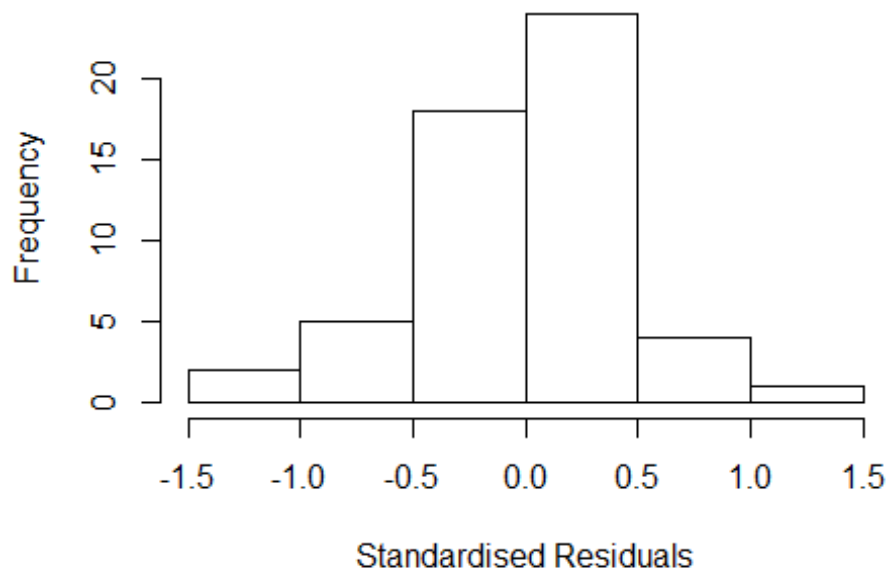
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1399 -0.2139 0.0607 0.1744 1.3896
##
## Coefficients:
```

```

##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0453    0.1829   5.71 1.5e-06 ***
## FirstAuthorFemale1 -0.2405    0.1948  -1.24   0.22
## LastAuthorFemale1  0.1105    0.1248   0.89   0.38
## Year1997          0.9397    0.1829   5.14 9.2e-06 ***
## Year1999          -0.2673    0.1829  -1.46   0.15
## Year2001           0.3923    0.2369   1.66   0.11
## Year2002          -0.4290    0.3212  -1.34   0.19
## Year2003           0.1630    0.2084   0.78   0.44
## Year2004           0.1874    0.3135   0.60   0.55
## Year2005           0.0522    0.1912   0.27   0.79
## Year2006           0.3352    0.5203   0.64   0.52
## Year2007          -0.0849    0.2643  -0.32   0.75
## Year2008          -0.0346    0.3258  -0.11   0.92
## Year2009          -0.1453    0.1879  -0.77   0.44
## Year2010           0.0668    0.2436   0.27   0.79
## Year2011          -0.2180    0.1802  -1.21   0.23
## Year2012           0.1136    0.2175   0.52   0.60
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.443
## Multiple R-squared:  0.264, Adjusted R-squared:  -0.0536
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 47 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.304  0.876  0.972  0.911  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      1.85e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500           50           2           1           1000      200
##      trace.lev      mts      compute.rd
##      0             1000      0
##      psi          subsampling          cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.887e+14 1      1.374e+07
## Year              1.887e+14 14      3.235e+00

```

Residuals from first author



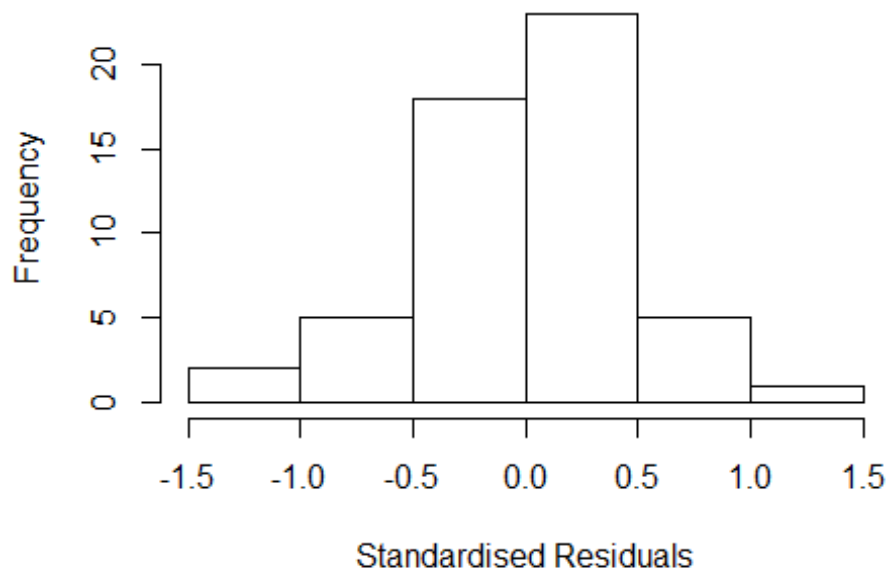
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1742 -0.1713 0.0526 0.1825 1.3702
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1353 0.1549 7.33 8.9e-09 ***
## FirstAuthorFemale1 -0.1996 0.1842 -1.08 0.285
## Year1997 0.8497 0.1549 5.48 2.9e-06 ***
## Year1999 -0.3573 0.1549 -2.31 0.027 *
## Year2001 0.2613 0.1549 1.69 0.100 .
## Year2002 -0.4686 0.2936 -1.60 0.119
## Year2003 0.0873 0.1895 0.46 0.648
## Year2004 0.0976 0.2971 0.33 0.744
## Year2005 -0.0378 0.1647 -0.23 0.820
## Year2006 0.2385 0.5100 0.47 0.643
## Year2007 -0.1555 0.2568 -0.61 0.549
## Year2008 -0.1243 0.3099 -0.40 0.691
```

```

## Year2009          -0.2160      0.1906   -1.13    0.264
## Year2010          -0.0249      0.2302   -0.11    0.914
## Year2011          -0.2699      0.1769   -1.53    0.135
## Year2012           0.0247      0.2033    0.12    0.904
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.44
## Multiple R-squared:  0.258, Adjusted R-squared:  -0.0343
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 48 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.312  0.887   0.973   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      1.85e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 18.55 1      4.307
## Year             18.55 14      1.110

```

Residuals from last author



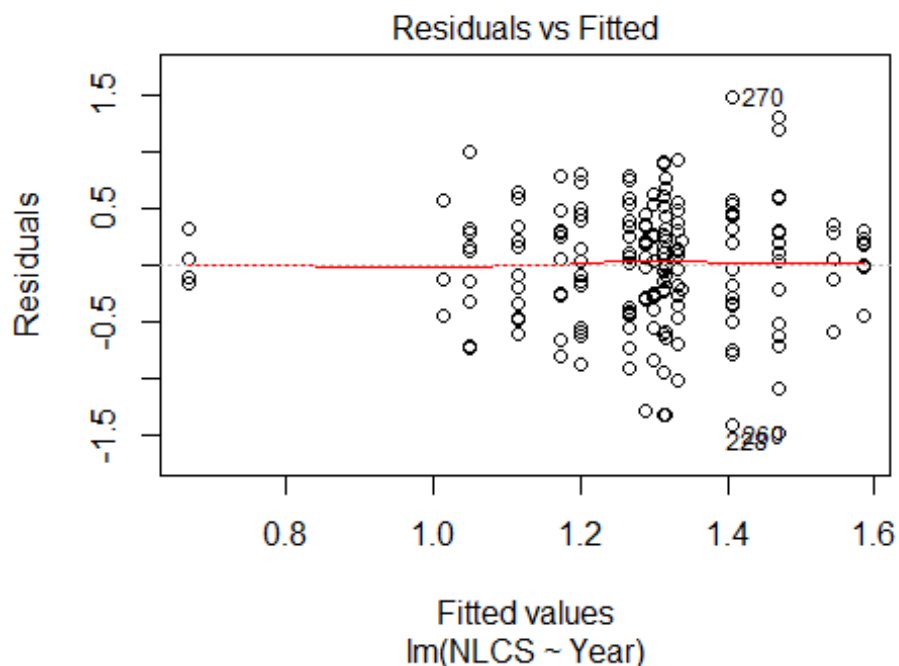
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3368 -0.2229 0.0396 0.2211 1.4397
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.01e+00 1.33e-01 7.58 4.1e-09 ***
## LastAuthorFemale1 2.64e-02 1.23e-01 0.21 0.832
## Year1997 9.76e-01 1.33e-01 7.33 8.9e-09 ***
## Year1999 -2.31e-01 1.33e-01 -1.74 0.091 .
## Year2001 1.88e-01 1.33e-01 1.41 0.167
## Year2002 -5.12e-01 3.40e-01 -1.50 0.141
## Year2003 6.78e-04 1.27e-01 0.01 0.996
## Year2004 2.24e-01 2.86e-01 0.78 0.439
## Year2005 8.84e-02 1.44e-01 0.61 0.544
## Year2006 3.28e-01 5.00e-01 0.66 0.516
## Year2007 -9.89e-02 2.60e-01 -0.38 0.706
## Year2008 1.86e-03 2.99e-01 0.01 0.995
```

```

## Year2009          -1.83e-01    1.55e-01    -1.18    0.246
## Year2010           1.07e-03    2.06e-01     0.01    0.996
## Year2011          -1.53e-01    1.19e-01    -1.29    0.206
## Year2012           9.73e-06    1.70e-01     0.00    1.000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.44
## Multiple R-squared:  0.247, Adjusted R-squared:  -0.0503
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 46 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.263  0.876  0.969  0.904  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      1.85e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  54"
## [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
##    4   11   8   11   17   12   13   17   3    5   16   14   19   17   18
## 2011 2012
##   12   19
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    9    5    8   15   12   13   14    3    5   12   13   15   17   15
## 2011 2012
##   12   18
##

```

```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      8      5      8     15     10     13     11      3      5     12     13     15     15     15
## 2011 2012
##     11     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 = 18, df = 16, p-value = 0.3
```

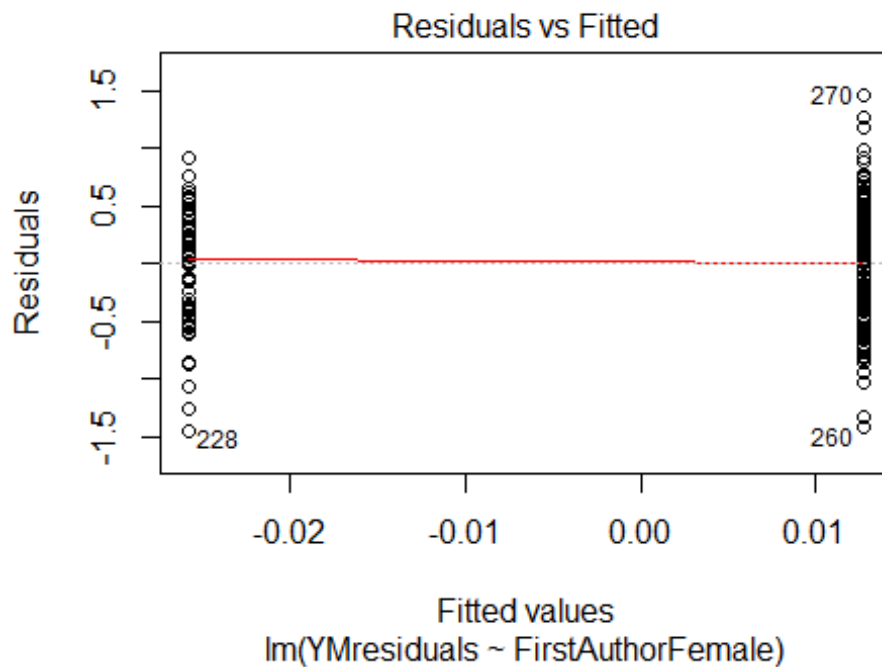


```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.18, df = 1, p-value = 0.7
##
## [1] "Female first author team size 2018 geometric mean: 1.73205080756888"
## [1] "Male first author team size 2018 geometric mean: 2.03393700979443"
##
## Warning in wilcox.test.default(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 = 0.9
## 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: 2.03393700979443"

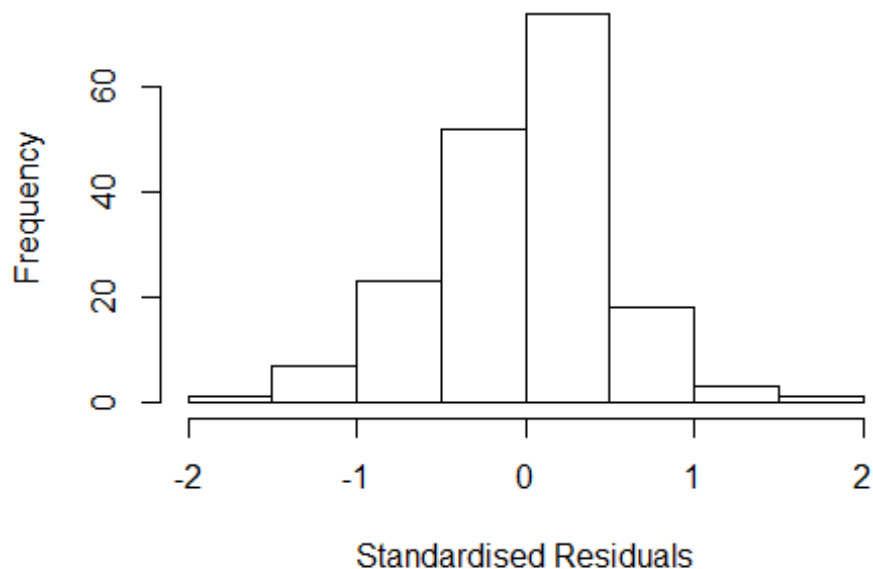
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 6, p-value = 0.9
## 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 | 6.815 | 1 | 2.611 |
| LastAuthorFemale | 2.765 | 1 | 1.663 |
| UniqueAuthors | 139.678 | 4 | 1.854 |
| Year | 151.790 | 16 | 1.170 |

Residuals from first and last author and team size



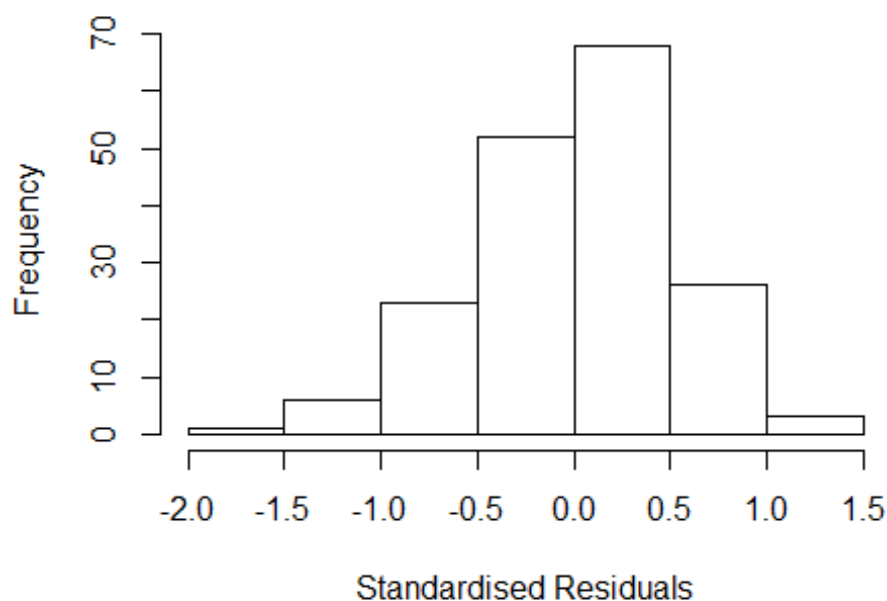
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5210 -0.3167 0.0486 0.3041 1.5265
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.17299 0.18804 6.24 4e-09 ***
## FirstAuthorFemale1 -0.03204 0.10273 -0.31 0.7555
## LastAuthorFemale1 -0.03518 0.10122 -0.35 0.7286
## UniqueAuthors2 0.16451 0.09956 1.65 0.1005
## UniqueAuthors3 -0.07688 0.17035 -0.45 0.6524
## UniqueAuthors4 0.26251 0.13930 1.88 0.0614 .
## UniqueAuthors5 0.51754 0.17557 2.95 0.0037 **
## Year1997 -0.17156 0.26678 -0.64 0.5211
## Year1998 -0.52029 0.19261 -2.70 0.0077 **
## Year1999 0.37768 0.20396 1.85 0.0660 .
```

```

## Year2000      0.12890    0.24050    0.54    0.5927
## Year2001     -0.00778    0.21142   -0.04    0.9707
## Year2002      0.08206    0.22805    0.36    0.7195
## Year2003     -0.10011    0.23813   -0.42    0.6748
## Year2004     -0.16967    0.32528   -0.52    0.6027
## Year2005      0.30510    0.23238    1.31    0.1911
## Year2006     -0.10416    0.22033   -0.47    0.6371
## Year2007      0.15662    0.22663    0.69    0.4905
## Year2008      0.12996    0.25005    0.52    0.6040
## Year2009      0.06462    0.24057    0.27    0.7886
## Year2010      0.22989    0.26584    0.86    0.3885
## Year2011      0.05936    0.20932    0.28    0.7771
## Year2012      0.18352    0.23184    0.79    0.4298
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.528
## Multiple R-squared:  0.145, Adjusted R-squared:  0.0239
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 165 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.384  0.895  0.960  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      5.59e-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 1.853 1      1.361
## LastAuthorFemale  1.619 1      1.272
## Year              1.720 16      1.017

```

Residuals from first and last author



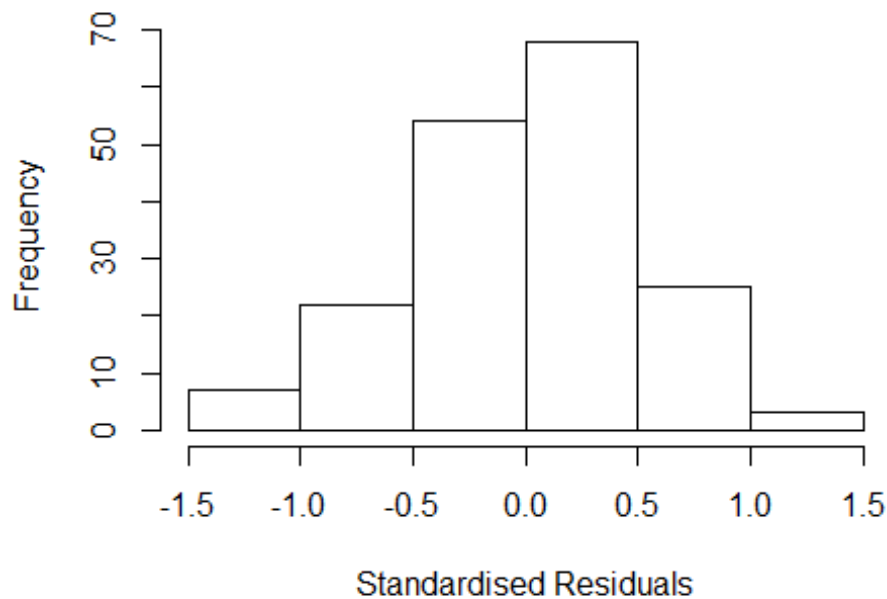
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5504 -0.3516 0.0412 0.2963 1.4662
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.337500 0.159530 8.38 2.5e-14 ***
## FirstAuthorFemale1 0.032204 0.102432 0.31 0.75363
## LastAuthorFemale1 -0.079640 0.102937 -0.77 0.44026
## Year1997 -0.253643 0.259362 -0.98 0.32957
## Year1998 -0.669608 0.177345 -3.78 0.00022 ***
## Year1999 0.280062 0.194681 1.44 0.15223
## Year2000 0.001245 0.213744 0.01 0.99536
## Year2001 -0.150536 0.200733 -0.75 0.45440
## Year2002 -0.029454 0.213899 -0.14 0.89065
## Year2003 -0.258951 0.235704 -1.10 0.27358
## Year2004 -0.334219 0.311126 -1.07 0.28434
## Year2005 0.215260 0.222575 0.97 0.33494
```

```

## Year2006      -0.219497    0.205840    -1.07    0.28787
## Year2007      0.054054    0.221914     0.24    0.80787
## Year2008      0.016113    0.226276     0.07    0.94332
## Year2009     -0.047927    0.224174    -0.21    0.83098
## Year2010      0.180651    0.287212     0.63    0.53026
## Year2011      0.000633    0.197497     0.00    0.99745
## Year2012      0.079335    0.221083     0.36    0.72018
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.527
## Multiple R-squared:  0.115, Adjusted R-squared:  0.0156
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 165 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.368  0.890   0.958   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      5.59e-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 1.367 1      1.169
## Year              1.367 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
## -1.4919 -0.3463 0.0414 0.2989 1.4701
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.33750 0.15947 8.39 2.4e-14 ***
## FirstAuthorFemale1 -0.00424 0.08970 -0.05 0.96238
## Year1997 -0.24774 0.25782 -0.96 0.33805
## Year1998 -0.67151 0.18078 -3.71 0.00028 ***
## Year1999 0.25479 0.19680 1.29 0.19728
## Year2000 0.00521 0.21518 0.02 0.98072
## Year2001 -0.16459 0.19923 -0.83 0.40996
## Year2002 -0.03282 0.21246 -0.15 0.87743
## Year2003 -0.26282 0.23737 -1.11 0.26985
## Year2004 -0.33406 0.31080 -1.07 0.28405
## Year2005 0.21514 0.22248 0.97 0.33498
## Year2006 -0.22350 0.20349 -1.10 0.27370
```

```

## Year2007          0.03159    0.22177    0.14  0.88690
## Year2008          0.01528    0.22733    0.07  0.94649
## Year2009         -0.06251    0.22078   -0.28  0.77743
## Year2010          0.15861    0.28174    0.56  0.57423
## Year2011         -0.00862    0.19713   -0.04  0.96519
## Year2012          0.07539    0.22112    0.34  0.73357
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.531
## Multiple R-squared:  0.111, Adjusted R-squared:  0.0175
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 161 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.410  0.889  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      5.59e-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 1.241  1          1.114
## Year              1.241 16          1.007
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.5205 -0.3565  0.0392  0.3052  1.4596

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.33750    0.15933   8.39 2.3e-14 ***
## LastAuthorFemale1 -0.06153    0.09010  -0.68 0.49565
## Year1997       -0.24842    0.25643  -0.97 0.33412
## Year1998       -0.66066    0.17733  -3.73 0.00027 ***
## Year1999        0.28937    0.19096   1.52 0.13164
## Year2000        0.01291    0.20702   0.06 0.95035
## Year2001       -0.14676    0.20106  -0.73 0.46651
## Year2002       -0.01948    0.21162  -0.09 0.92679
## Year2003       -0.24767    0.22885  -1.08 0.28076
## Year2004       -0.33373    0.31013  -1.08 0.28349
## Year2005        0.21489    0.22228   0.97 0.33511
## Year2006       -0.21476    0.20514  -1.05 0.29671
## Year2007        0.05744    0.21826   0.26 0.79274
## Year2008        0.01910    0.22757   0.08 0.93320
## Year2009       -0.04640    0.22293  -0.21 0.83537
## Year2010        0.18296    0.28565   0.64 0.52275
## Year2011        0.00583    0.19838   0.03 0.97661
## Year2012        0.08592    0.22085   0.39 0.69775
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.538
## Multiple R-squared:  0.113, Adjusted R-squared:  0.0193
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 163 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.404  0.892  0.959  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      5.59e-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: 179"
## [1] ""

```



```

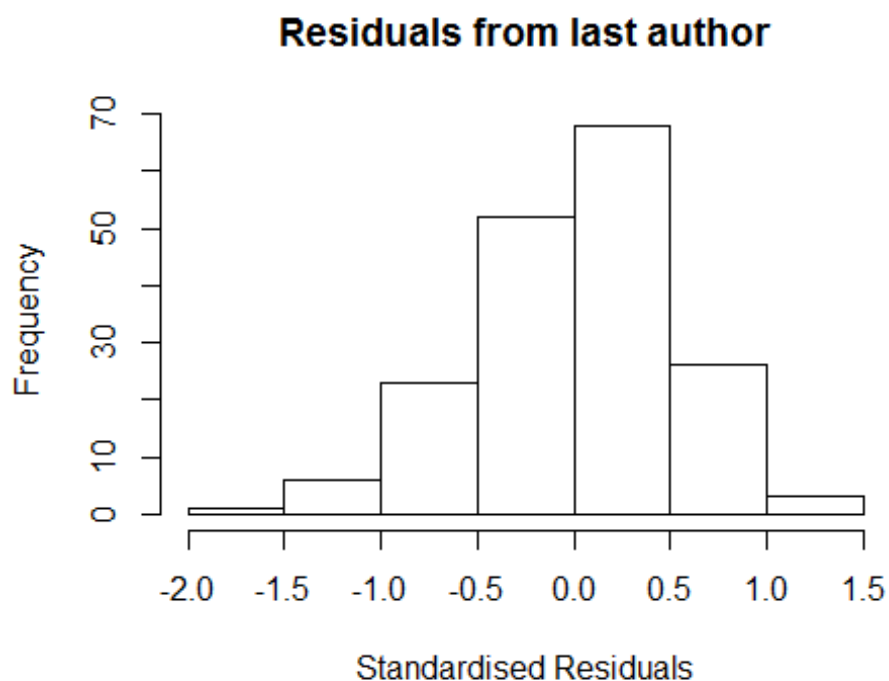
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1401"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    4    2    6    5    5    2    3    3    1    7    9   10    5    9   13
## 2012
##    5
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    4    2    5    4    4    2    3    3    1    6    7    9    4    9   11
## 2012
##    5
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    4    2    5    3    4    2    3    3    1    6    7    9    4    9   10
## 2012
##    3
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.6207413942089"
## [1] "Male first author team size 2018 geometric mean: 2.03964890265551"

## Warning in wilcox.test.default(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] "Female last author team size 2018 geometric mean: 2.91295063024394"
## [1] "Male last author team size 2018 geometric mean: 1.782602457966"

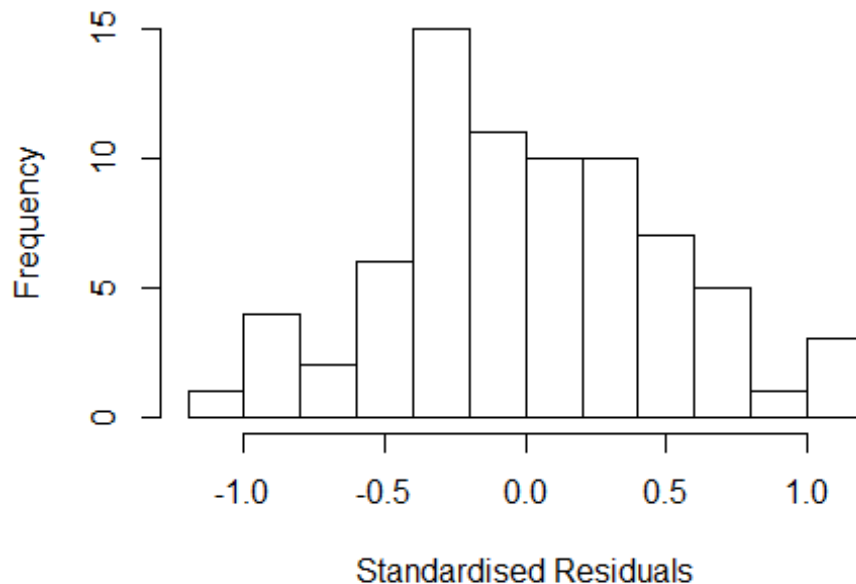
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## 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 1.429e+01  1      3.781e+00
## LastAuthorFemale  4.194e+13  1      6.476e+06
## UniqueAuthors    8.921e+15  4      9.858e+01
## Year              3.979e+16 15      3.575e+00
```

Residuals from first and last author and team size



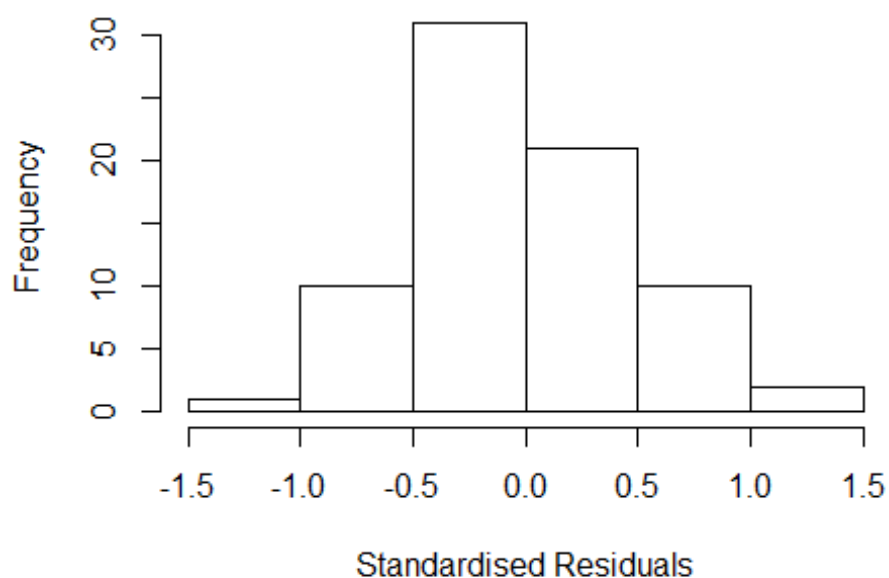
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.10e+00 -3.36e-01 -5.55e-17 3.81e-01 1.14e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3058 0.6100 2.14 0.037 *
## FirstAuthorFemale1 0.3507 0.2178 1.61 0.113
## LastAuthorFemale1 0.0335 0.1866 0.18 0.858
## UniqueAuthors2 0.1516 0.1680 0.90 0.371
## UniqueAuthors3 0.1077 0.3862 0.28 0.781
## UniqueAuthors4 0.4917 0.3691 1.33 0.189
## UniqueAuthors5 -0.1138 0.2966 -0.38 0.703
## Year1998 -0.5769 0.7076 -0.82 0.419
## Year1999 -0.2146 0.6475 -0.33 0.742
## Year2000 -0.9268 0.9166 -1.01 0.317
```

```

## Year2001          0.5961      0.5962      1.00      0.322
## Year2002          0.2902      0.6969      0.42      0.679
## Year2003         -0.6217      0.5950     -1.04      0.301
## Year2004         -0.6463      0.5141     -1.26      0.214
## Year2005         -1.2461      0.6425     -1.94      0.058 .
## Year2006         -0.3621      0.5568     -0.65      0.518
## Year2007         -0.4419      0.6044     -0.73      0.468
## Year2008         -0.4498      0.6230     -0.72      0.474
## Year2009          0.1563      0.5951      0.26      0.794
## Year2010         -0.5663      0.5420     -1.04      0.301
## Year2011         -0.6270      0.4949     -1.27      0.211
## Year2012          0.1263      0.6537      0.19      0.848
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.54
## Multiple R-squared:  0.437, Adjusted R-squared:  0.213
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 69 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.636  0.916  0.956  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      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 21.62 1      4.650
## LastAuthorFemale  12.26 1      3.501
## Year              90.59 15      1.162

```

Residuals from first and last author



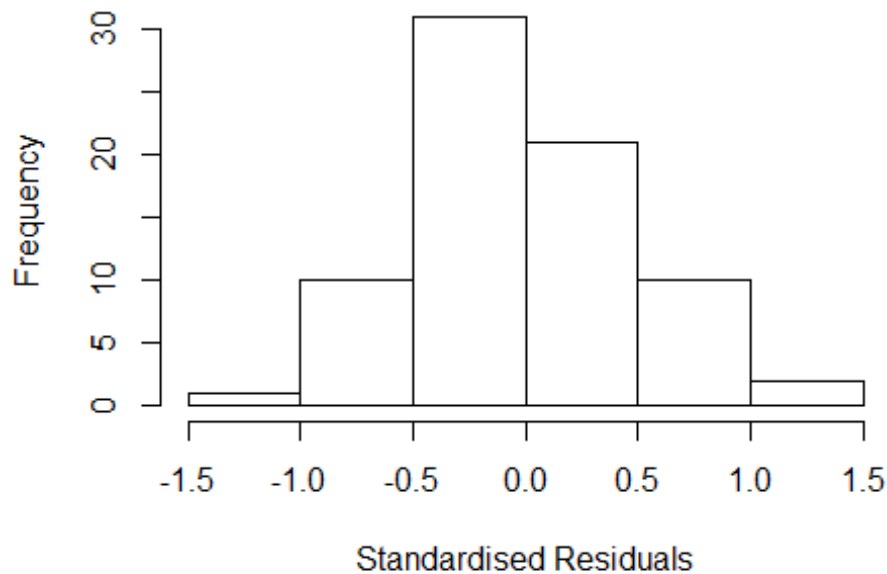
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1941 -0.3754 -0.0368 0.3869 1.2207
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4103 0.4283 3.29 0.0017 **
## FirstAuthorFemale1 0.3459 0.1662 2.08 0.0419 *
## LastAuthorFemale1 0.0270 0.1681 0.16 0.8732
## Year1998 -0.6032 0.5990 -1.01 0.3182
## Year1999 -0.2290 0.4411 -0.52 0.6056
## Year2000 -1.0060 0.7335 -1.37 0.1756
## Year2001 0.5686 0.4670 1.22 0.2284
## Year2002 0.1913 0.5008 0.38 0.7039
## Year2003 -0.6719 0.4123 -1.63 0.1087
## Year2004 -0.5515 0.4627 -1.19 0.2383
## Year2005 -1.1942 0.4140 -2.88 0.0055 **
## Year2006 -0.3906 0.4510 -0.87 0.3901
```

```

## Year2007          -0.4932      0.4609   -1.07    0.2891
## Year2008          -0.5349      0.4488   -1.19    0.2383
## Year2009           0.1723      0.4008    0.43    0.6689
## Year2010          -0.5304      0.4420   -1.20    0.2351
## Year2011          -0.5475      0.4675   -1.17    0.2464
## Year2012           0.0259      0.4738    0.05    0.9566
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.576
## Multiple R-squared:  0.409, Adjusted R-squared:  0.233
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.633  0.910  0.956  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 15.15 1      3.892
## Year              15.15 15      1.095

```

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
## -1.2039 -0.3758 -0.0393 0.3922 1.2325
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4281 0.4382 3.26 0.0019 **
## FirstAuthorFemale1 0.3648 0.1344 2.72 0.0087 **
## Year1998 -0.6305 0.5836 -1.08 0.2844
## Year1999 -0.2511 0.4450 -0.56 0.5747
## Year2000 -1.0356 0.7451 -1.39 0.1698
## Year2001 0.5620 0.4743 1.18 0.2409
## Year2002 0.1775 0.5088 0.35 0.7284
## Year2003 -0.6870 0.4233 -1.62 0.1100
## Year2004 -0.5668 0.4716 -1.20 0.2343
## Year2005 -1.2309 0.4048 -3.04 0.0035 **
## Year2006 -0.3999 0.4616 -0.87 0.3899
## Year2007 -0.5089 0.4707 -1.08 0.2841
```

```

## Year2008          -0.5491      0.4605   -1.19    0.2380
## Year2009           0.1537      0.4117    0.37    0.7102
## Year2010          -0.5442      0.4526   -1.20    0.2342
## Year2011          -0.5611      0.4780   -1.17    0.2452
## Year2012           0.0102      0.4841    0.02    0.9833
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.561
## Multiple R-squared:  0.412, Adjusted R-squared:  0.25
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.609  0.906  0.956  0.925  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 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 8.627 1          2.937
## Year            8.627 15          1.074

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.1123 -0.3636 -0.0341  0.3596  1.2314
##

```



```

## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.4055      0.3632   3.87 0.00028 ***
## LastAuthorFemale1 0.2958      0.1426   2.08 0.04243 *
## Year1998         -0.4255      0.7799  -0.55 0.58743
## Year1999         -0.0550      0.4032  -0.14 0.89199
## Year2000         -1.0119      0.7042  -1.44 0.15611
## Year2001          0.5442      0.4221   1.29 0.20243
## Year2002          0.2346      0.4273   0.55 0.58516
## Year2003         -0.6414      0.3456  -1.86 0.06857 .
## Year2004         -0.5218      0.3947  -1.32 0.19128
## Year2005         -0.8435      0.3632  -2.32 0.02375 *
## Year2006         -0.3961      0.3952  -1.00 0.32033
## Year2007         -0.4503      0.4105  -1.10 0.27724
## Year2008         -0.5097      0.3936  -1.29 0.20050
## Year2009          0.3019      0.3447   0.88 0.38469
## Year2010         -0.4908      0.3808  -1.29 0.20254
## Year2011         -0.4866      0.4044  -1.20 0.23370
## Year2012          0.0533      0.4341   0.12 0.90276
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.563
## Multiple R-squared:  0.387, Adjusted R-squared:  0.218
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~ = 1. The remaining 69 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.611  0.887   0.958   0.924   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      1.33e-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] "#####"

```

```

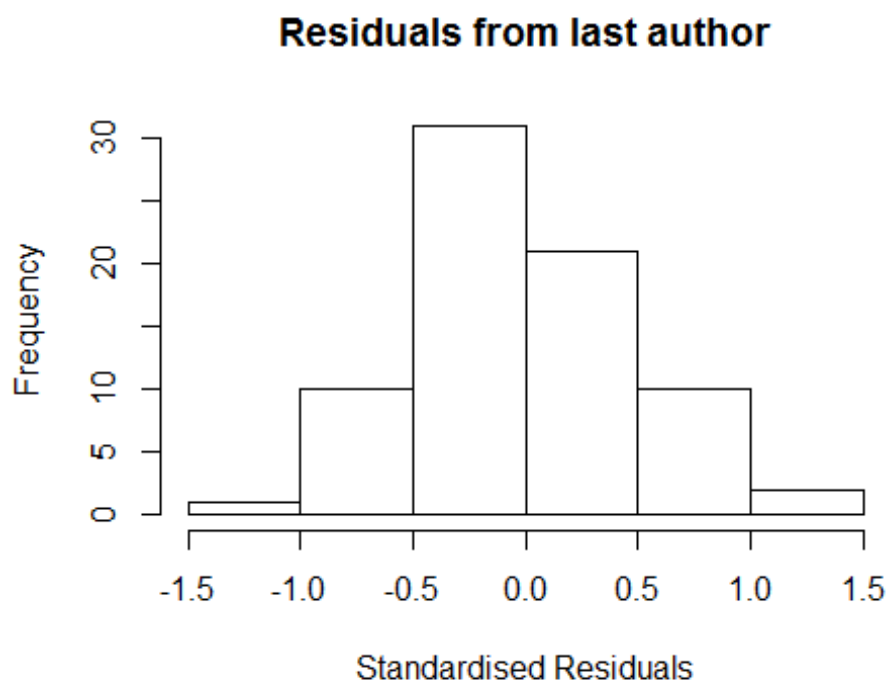
## [1] "Analysis of AJSC 1402"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    5    1    2    2    2    3    5    6    5    6    4    7    6    6
## 2011 2012
##    6    11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    5    1    2    1    2    3    3    4    5    4    4    6    4    6
## 2011 2012
##    4    10
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    5    1    2    1    2    2    3    4    4    4    3    6    4    6
## 2011 2012
##    4    7
## [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: 1.8881750225898"

## 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 = 4.5, p-value = 0.3
## 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: 1.8881750225898"

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

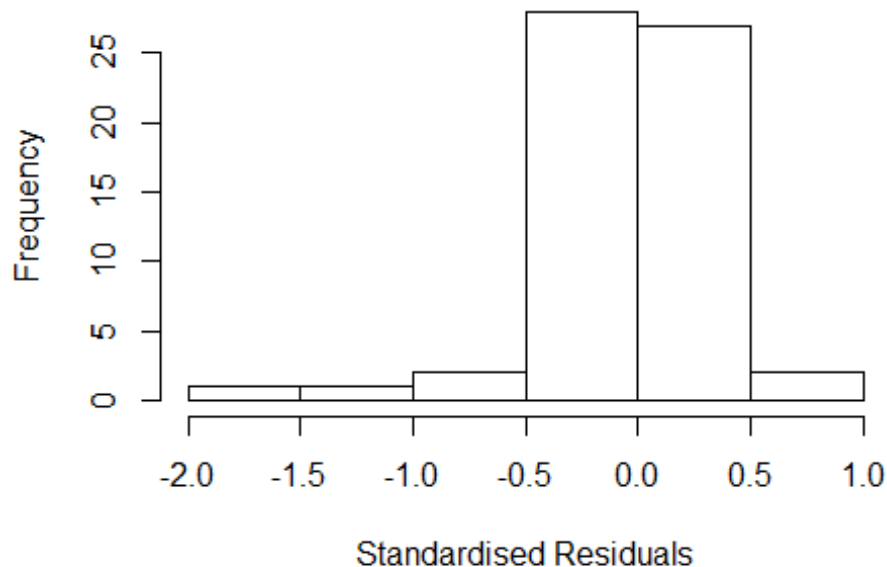
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4.5, 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 | -4.647e+15 | 1 | NaN |
| ## LastAuthorFemale | 1.075e+16 | 1 | 1.037e+08 |
| ## UniqueAuthors | 3.095e+43 | 3 | 1.772e+07 |
| ## Year | 6.202e+58 | 16 | 6.875e+01 |

Residuals from first and last author and team size



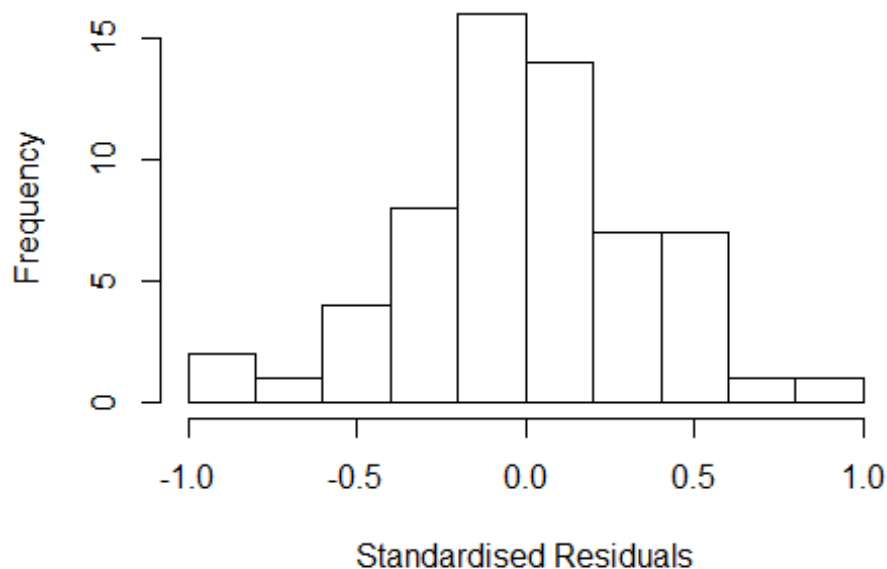
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.51e+00 -1.50e-01 1.11e-16 1.20e-01 7.75e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.65824 0.28319 2.32 0.0254 *
## FirstAuthorFemale1 -0.00414 0.17139 -0.02 0.9809
## LastAuthorFemale1 -0.22439 0.17815 -1.26 0.2153
## UniqueAuthors2 0.37526 0.14127 2.66 0.0114 *
## UniqueAuthors3 -0.61163 0.22307 -2.74 0.0092 **
## UniqueAuthors4 -0.10958 0.14098 -0.78 0.4417
## Year1997 0.30411 0.28332 1.07 0.2897
## Year1998 -0.36224 0.28319 -1.28 0.2084
## Year1999 1.07679 0.24306 4.43 7.4e-05 ***
## Year2000 0.15729 0.24216 0.65 0.5198
```

```

## Year2001      0.43513      0.34056      1.28      0.2089
## Year2002     -0.07961      0.27058     -0.29      0.7702
## Year2003      0.05300      0.30529      0.17      0.8631
## Year2004      0.00780      0.30070      0.03      0.9794
## Year2005      0.16018      0.29267      0.55      0.5873
## Year2006      0.32724      0.30767      1.06      0.2940
## Year2007      0.52235      0.31368      1.67      0.1039
## Year2008      0.32611      0.33385      0.98      0.3347
## Year2009      0.45839      0.37981      1.21      0.2348
## Year2010      0.66128      0.28937      2.29      0.0278 *
## Year2011      0.82733      0.32387      2.55      0.0147 *
## Year2012      0.04342      0.35842      0.12      0.9042
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.357
## Multiple R-squared:  0.596, Adjusted R-squared:  0.378
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 50 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0335 0.9220 0.9740 0.9080 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      1.64e-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 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.643e+14 1      1.282e+07
## LastAuthorFemale  2.199e+14 1      1.483e+07
## Year              5.597e+14 16      2.890e+00

```

Residuals from first and last author



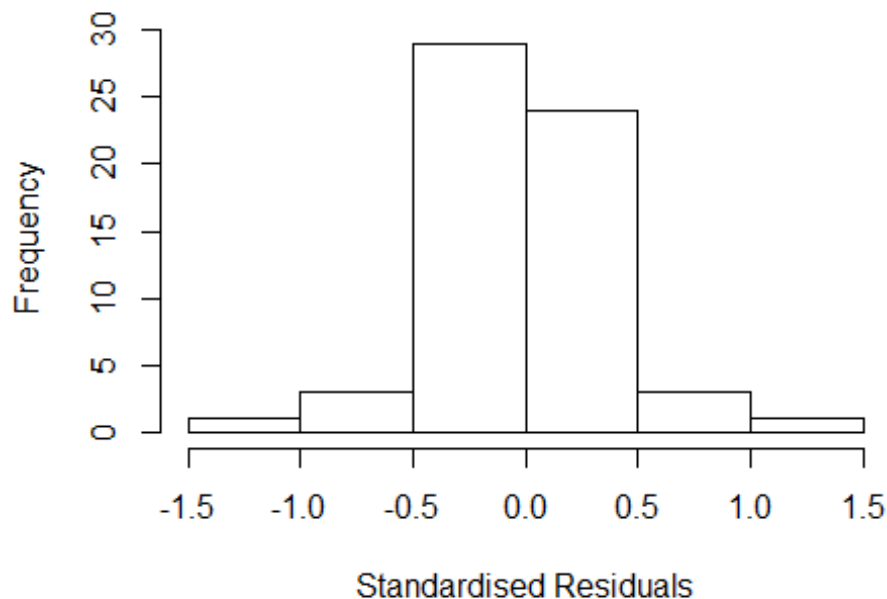
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -9.90e-01 -1.95e-01 5.55e-17 2.03e-01 9.57e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.546 0.342 1.60 0.11747
## FirstAuthorFemale1 0.183 0.190 0.96 0.34016
## LastAuthorFemale1 -0.258 0.220 -1.17 0.24695
## Year1997 0.541 0.333 1.62 0.11224
## Year1998 -0.250 0.342 -0.73 0.46818
## Year1999 1.035 0.285 3.63 0.00075 ***
## Year2000 0.116 0.284 0.41 0.68617
## Year2001 0.735 0.383 1.92 0.06157 .
## Year2002 0.143 0.309 0.46 0.64542
## Year2003 0.224 0.371 0.60 0.54921
## Year2004 0.215 0.348 0.62 0.54039
## Year2005 0.234 0.351 0.67 0.50797
```

```

## Year2006          0.491      0.349      1.41  0.16644
## Year2007          0.817      0.364      2.24  0.03017 *
## Year2008          0.483      0.419      1.15  0.25488
## Year2009          0.531      0.420      1.26  0.21293
## Year2010          0.737      0.350      2.10  0.04139 *
## Year2011          0.794      0.445      1.78  0.08167 .
## Year2012          0.349      0.417      0.84  0.40672
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.414
## Multiple R-squared:  0.393, Adjusted R-squared:  0.132
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 52 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.547  0.898  0.967  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      1.64e-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 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.632e+14 1      1.278e+07
## Year      1.632e+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
## -1.055 -0.184 -0.023 0.198 1.003
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.47998 0.38304 1.25 0.2169
## FirstAuthorFemale1 0.00804 0.14168 0.06 0.9550
## Year1997 0.53591 0.38628 1.39 0.1725
## Year1998 -0.18398 0.38304 -0.48 0.6334
## Year1999 1.01847 0.32639 3.12 0.0032 **
## Year2000 0.09897 0.32572 0.30 0.7627
## Year2001 0.97647 0.33904 2.88 0.0062 **
## Year2002 0.16800 0.34879 0.48 0.6325
## Year2003 0.33848 0.43132 0.78 0.4369
## Year2004 0.28087 0.38849 0.72 0.4736
## Year2005 0.27924 0.39704 0.70 0.4857
## Year2006 0.53360 0.38620 1.38 0.1742
```

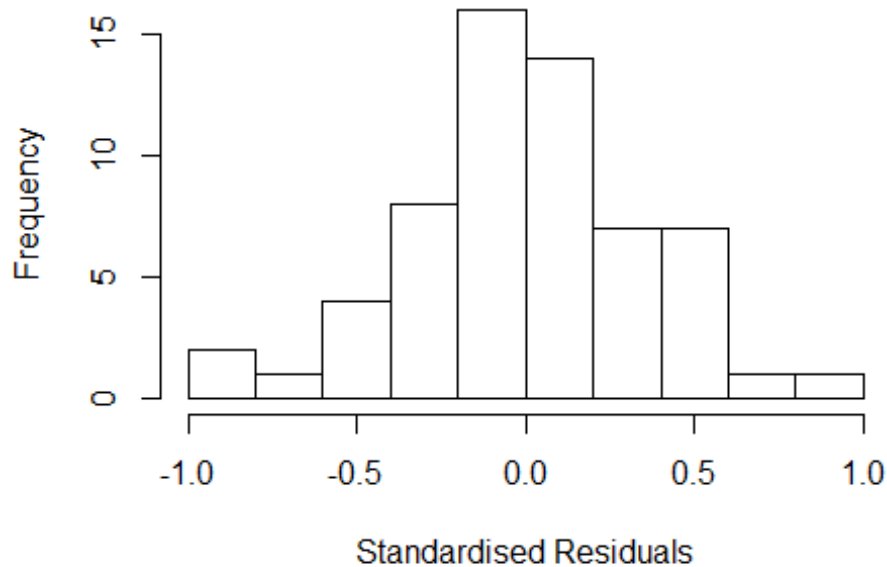


```

## Year2007          0.94346    0.37039    2.55    0.0145 *
## Year2008          0.54652    0.45426    1.20    0.2355
## Year2009          0.59808    0.45659    1.31    0.1972
## Year2010          0.77438    0.38635    2.00    0.0514 .
## Year2011          0.92574    0.48029    1.93    0.0605 .
## Year2012          0.36863    0.43711    0.84    0.4037
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.387, Adjusted R-squared:  0.145
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 51 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.471  0.890   0.961   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      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"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale -1.359e+15  1      NaN
## Year -1.359e+15 16      NaN

```

Residuals from last author



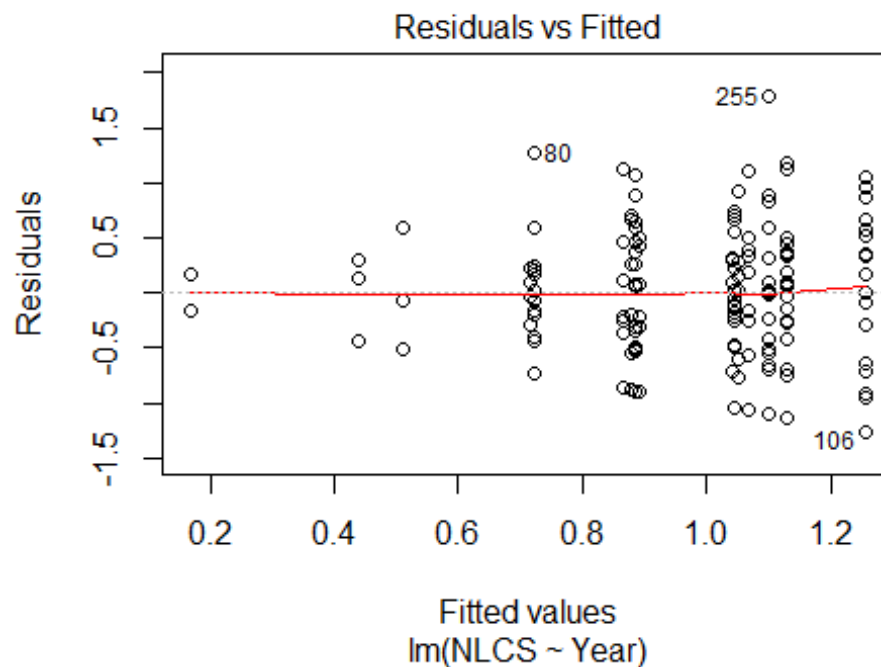
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.06e+00 -1.99e-01 5.55e-17 1.49e-01 9.81e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.577 0.321 1.80 0.0790 .
## LastAuthorFemale1 -0.117 0.134 -0.87 0.3881
## Year1997 0.488 0.307 1.59 0.1191
## Year1998 -0.281 0.321 -0.88 0.3858
## Year1999 1.046 0.272 3.84 0.0004 ***
## Year2000 0.127 0.272 0.47 0.6432
## Year2001 0.888 0.334 2.66 0.0111 *
## Year2002 0.133 0.296 0.45 0.6544
## Year2003 0.244 0.387 0.63 0.5323
## Year2004 0.184 0.327 0.56 0.5770
## Year2005 0.214 0.331 0.65 0.5215
## Year2006 0.472 0.335 1.41 0.1659
```

```

## Year2007          0.849      0.337      2.52      0.0154 *
## Year2008          0.449      0.403      1.11      0.2715
## Year2009          0.501      0.405      1.24      0.2231
## Year2010          0.723      0.338      2.14      0.0381 *
## Year2011          0.831      0.449      1.85      0.0711 .
## Year2012          0.294      0.381      0.77      0.4442
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.393, Adjusted R-squared:  0.153
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 52 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.469  0.890  0.974  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      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] "Sample size for the above analysis: 61"
## [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
##    1    5    2    6    3    6    7    7    4    7   15   17   26   21   21
## 2011 2012
##   26   22
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    3    2    6    3    4    6    7    4    7   12   14   19   15   17
## 2011 2012

```

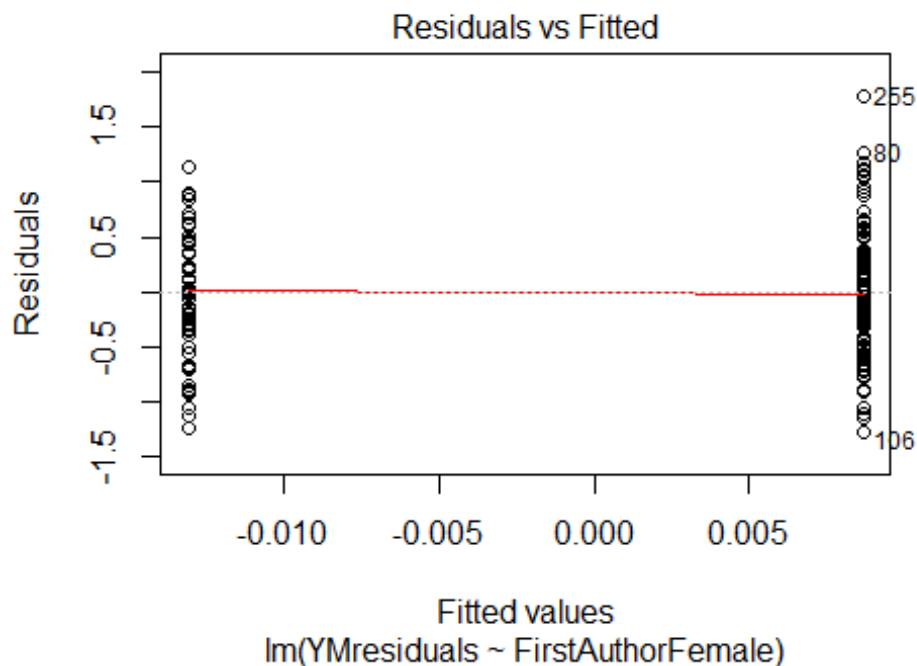
```
## 21 18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 0 3 2 6 3 4 5 5 3 7 11 13 19 15 14
## 2011 2012
## 17 17
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 9.4, df = 15, p-value = 0.9
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.46, df = 1, p-value = 0.5
## [1] "Female first author team size 2018 geometric mean: 2"
## [1] "Male first author team size 2018 geometric mean: 1.72637320145005"
## Warning in wilcox.test.default(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 = 0.6
## 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: 1.79685106477547"

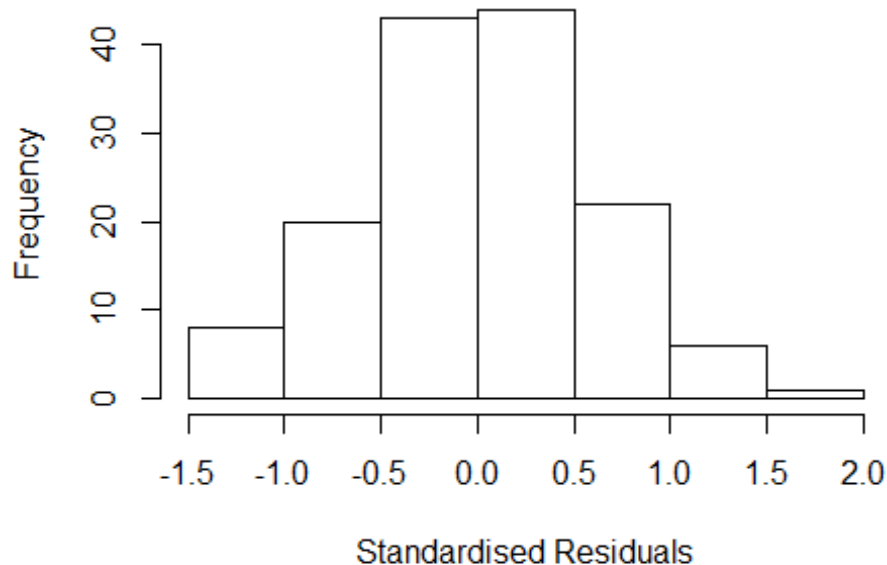
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## 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 | 2.466 | 1 | 1.570 |
| LastAuthorFemale | 3.337 | 1 | 1.827 |
| UniqueAuthors | 2.692 | 2 | 1.281 |
| Year | 5.572 | 15 | 1.059 |

Residuals from first and last author and team size



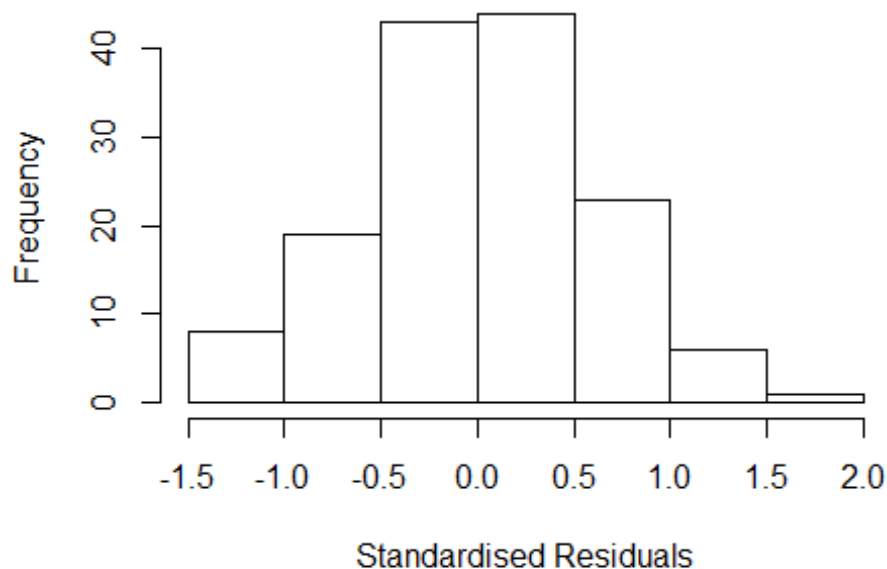
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3359 -0.4129 0.0113 0.3563 1.8132
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.44257 0.19404 2.28 0.0243 *
## FirstAuthorFemale1 0.06057 0.14630 0.41 0.6796
## LastAuthorFemale1 -0.01696 0.16306 -0.10 0.9173
## UniqueAuthors2 0.00821 0.12834 0.06 0.9491
## UniqueAuthors3 -0.07245 0.14765 -0.49 0.6245
## Year1998 -0.26065 0.27143 -0.96 0.3388
## Year1999 0.43076 0.35006 1.23 0.2208
## Year2000 0.03294 0.34469 0.10 0.9240
## Year2001 0.64566 0.32134 2.01 0.0467 *
## Year2002 0.59862 0.37267 1.61 0.1107
```

```

## Year2003          0.20730      0.36281      0.57      0.5688
## Year2004          0.38985      0.25284      1.54      0.1256
## Year2005          0.44375      0.26648      1.67      0.0984 .
## Year2006          0.69366      0.25107      2.76      0.0066 **
## Year2007          0.30239      0.23282      1.30      0.1964
## Year2008          0.84974      0.28435      2.99      0.0034 **
## Year2009          0.59156      0.25288      2.34      0.0209 *
## Year2010          0.35859      0.26712      1.34      0.1819
## Year2011          0.62269      0.26942      2.31      0.0225 *
## Year2012          0.62727      0.28322      2.21      0.0286 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.635
## Multiple R-squared:  0.128, Adjusted R-squared:  -0.00532
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  10 weights are ~ = 1. The remaining 134 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.394  0.894  0.957   0.922  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 2.552 1      1.598
## LastAuthorFemale  3.443 1      1.856
## Year              2.179 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
## -1.32920 -0.42270 0.00778 0.36031 1.84090
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.4453 0.1907 2.33 0.0211 *
## FirstAuthorFemale1 0.0653 0.1483 0.44 0.6603
## LastAuthorFemale1 -0.0171 0.1653 -0.10 0.9179
## Year1998 -0.3019 0.2444 -1.24 0.2191
## Year1999 0.4284 0.3515 1.22 0.2251
## Year2000 0.0269 0.3431 0.08 0.9377
## Year2001 0.6026 0.2932 2.06 0.0419 *
## Year2002 0.5970 0.3710 1.61 0.1101
## Year2003 0.1989 0.3617 0.55 0.5834
## Year2004 0.3356 0.2242 1.50 0.1368
## Year2005 0.4419 0.2638 1.67 0.0964 .
## Year2006 0.6857 0.2485 2.76 0.0067 **
```

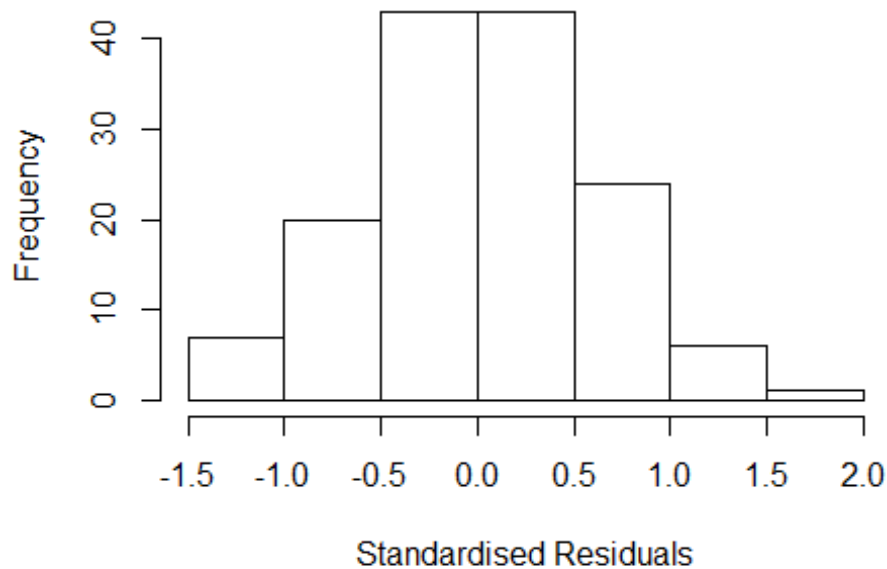


```

## Year2007          0.2932      0.2313      1.27      0.2073
## Year2008          0.8357      0.2800      2.98      0.0034 **
## Year2009          0.5804      0.2429      2.39      0.0184 *
## Year2010          0.3531      0.2630      1.34      0.1818
## Year2011          0.6162      0.2688      2.29      0.0236 *
## Year2012          0.5968      0.2728      2.19      0.0305 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.623
## Multiple R-squared:  0.128, Adjusted R-squared:  0.0101
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 135 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.363  0.891  0.954  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      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 1.415 1      1.189
## Year              1.415 15      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
## -1.337 -0.422 0.008 0.369 1.845
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.4453 0.1908 2.33 0.0211 *
## FirstAuthorFemale1 0.0558 0.1108 0.50 0.6154
## Year1998 -0.3057 0.2439 -1.25 0.2124
## Year1999 0.4222 0.3399 1.24 0.2165
## Year2000 0.0216 0.3382 0.06 0.9491
## Year2001 0.6005 0.2920 2.06 0.0418 *
## Year2002 0.5969 0.3713 1.61 0.1104
## Year2003 0.1947 0.3609 0.54 0.5906
## Year2004 0.3306 0.2196 1.51 0.1346
## Year2005 0.4430 0.2640 1.68 0.0958 .
## Year2006 0.6874 0.2479 2.77 0.0064 **
## Year2007 0.2922 0.2315 1.26 0.2092
```

```

## Year2008          0.8360      0.2805      2.98      0.0034 **
## Year2009          0.5770      0.2402      2.40      0.0177 *
## Year2010          0.3522      0.2604      1.35      0.1786
## Year2011          0.6129      0.2606      2.35      0.0202 *
## Year2012          0.5929      0.2690      2.20      0.0293 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.622
## Multiple R-squared:  0.128, Adjusted R-squared:  0.0181
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 135 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.358  0.893  0.954  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      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 1.965 1          1.402
## Year            1.965 15          1.023

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.3194 -0.4345  0.0198  0.3734  1.8309
##

```

```

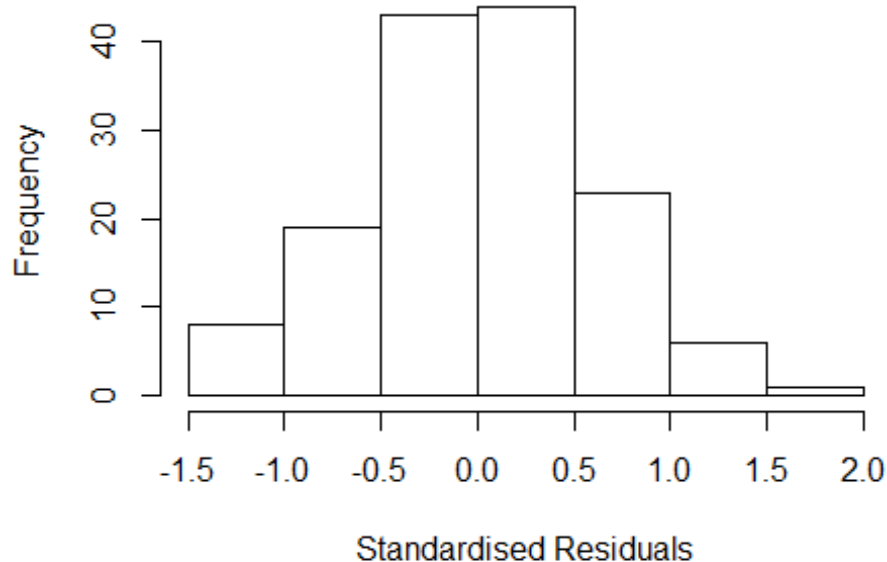
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.4452    0.1906   2.34  0.0211 *
## LastAuthorFemale1 0.0260    0.1243   0.21  0.8348
## Year1998         -0.2907    0.2390  -1.22  0.2262
## Year1999          0.4275    0.3507   1.22  0.2251
## Year2000          0.0421    0.3463   0.12  0.9034
## Year2001          0.6086    0.2926   2.08  0.0395 *
## Year2002          0.5972    0.3711   1.61  0.1101
## Year2003          0.2108    0.3576   0.59  0.5567
## Year2004          0.3508    0.2165   1.62  0.1077
## Year2005          0.4597    0.2654   1.73  0.0857 .
## Year2006          0.7061    0.2440   2.89  0.0045 **
## Year2007          0.3042    0.2303   1.32  0.1890
## Year2008          0.8482    0.2781   3.05  0.0028 **
## Year2009          0.5903    0.2414   2.44  0.0159 *
## Year2010          0.3687    0.2625   1.40  0.1626
## Year2011          0.6331    0.2724   2.32  0.0217 *
## Year2012          0.6069    0.2718   2.23  0.0273 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.628
## Multiple R-squared:  0.126, Adjusted R-squared:  0.0154
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 132 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.376  0.895   0.953   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      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 1404"
## [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
##    2    5    4    5    3    3    2    4    1    5    3    4    7    8    5
## 2012
##    5
##
## 1996 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    2    4    1    2    0    3    2    3    1    4    2    4    6    5    4
## 2012
##    1
##
## 1996 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    3    1    1    0    2    2    3    1    4    2    2    6    5    4
## 2012
##    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: 4"

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

Residuals from last author



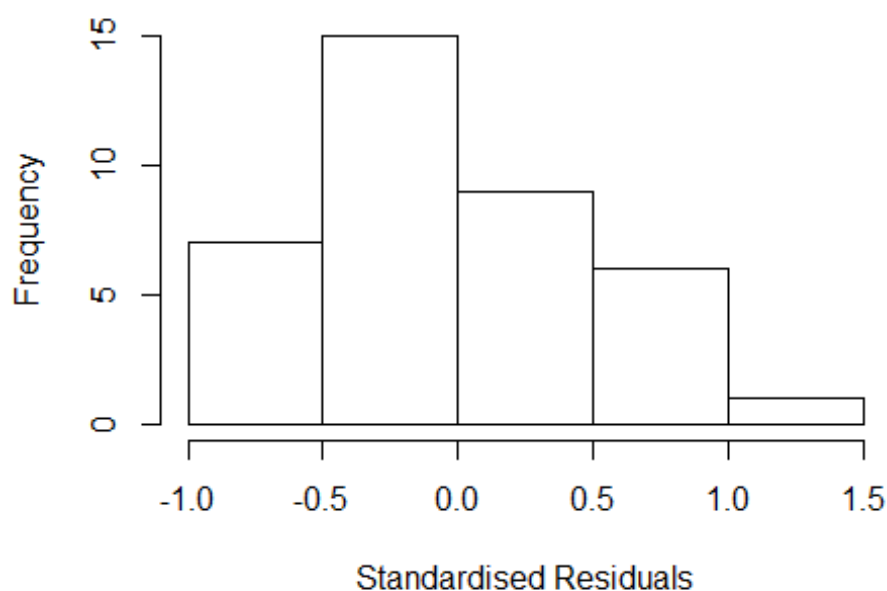
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 2.82842712474619"
## [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 12.126 1          3.482
## LastAuthorFemale  5.983 1          2.446
## Year              10.243 14         1.087
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8.66e-01 -4.36e-01  2.22e-16  3.83e-01  1.48e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.55e+00   5.62e-08  2.75e+07  <2e-16 ***
## FirstAuthorFemale1 -1.24e+00   3.56e-01 -3.49e+00  0.0022 **
## LastAuthorFemale1  1.33e-01   3.56e-01  3.70e-01  0.7128
## Year1998         6.00e-02   2.65e-01  2.30e-01  0.8229
## Year1999        -1.55e+00   5.52e-08 -2.80e+07  <2e-16 ***
## Year2000        -1.10e-01   4.04e-08 -2.73e+06  <2e-16 ***
## Year2002        -1.09e-01   2.64e-01 -4.10e-01  0.6847
## Year2003         5.94e-01   5.61e-01  1.06e+00  0.3019
## Year2004        -1.12e+00   4.00e-01 -2.81e+00  0.0106 *
## Year2005        -1.12e+00   5.38e-08 -2.07e+07  <2e-16 ***
## Year2006         2.05e-01   3.13e-01  6.60e-01  0.5185
## Year2007        -1.46e-01   3.27e-01 -4.50e-01  0.6589
## Year2008        -3.68e-01   5.68e-01 -6.50e-01  0.5240
## Year2009        -2.59e-01   3.12e-01 -8.30e-01  0.4155
## Year2010        -6.81e-01   3.24e-01 -2.10e+00  0.0479 *
## Year2011         3.67e-01   3.08e-01  1.19e+00  0.2466
## Year2012        -1.19e+00   4.04e-08 -2.94e+07  <2e-16 ***
```

```

## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.863
## Multiple R-squared:  0.463, Adjusted R-squared:  0.0544
## Convergence in 8 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 32 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.751  0.948  0.975  0.959  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           2.63e-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"
## 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

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8.72e-01 -4.22e-01  2.78e-16  4.15e-01  1.48e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.5470     0.0000      Inf <2e-16 ***
## FirstAuthorFemale1 -1.1342     0.2349    -4.83  8e-05 ***
## Year1998           0.1023     0.2359     0.43   0.67
## Year1999          -1.5470     0.0000    -Inf <2e-16 ***
## Year2000           -0.1100     0.0000    -Inf <2e-16 ***
## Year2002           -0.0969     0.2723    -0.36   0.73
## Year2003           0.5517     0.6230     0.89   0.39
## Year2004          -1.1255     0.4023    -2.80   0.01 *
```



```

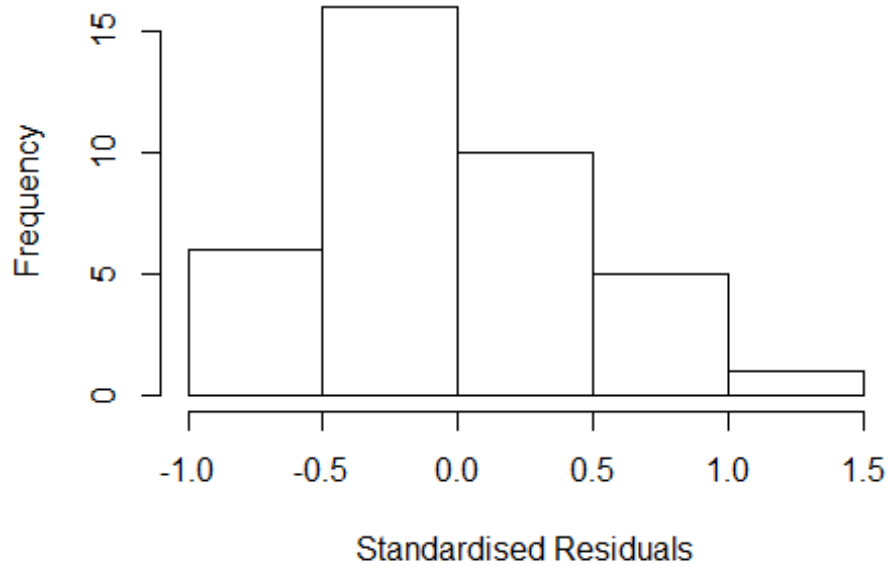
## Year2005          -1.1160      0.0000      -Inf    <2e-16 ***
## Year2006           0.2117      0.3120      0.68     0.50
## Year2007          -0.1465      0.3283     -0.45     0.66
## Year2008          -0.3680      0.5735     -0.64     0.53
## Year2009          -0.2630      0.3116     -0.84     0.41
## Year2010          -0.6749      0.3257     -2.07     0.05 .
## Year2011           0.3170      0.2692      1.18     0.25
## Year2012          -1.1860      0.0000      -Inf    <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.836
## Multiple R-squared:  0.463, Adjusted R-squared:  0.0966
## Convergence in 9 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~ = 1. The remaining 31 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.735  0.936  0.973  0.955  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      2.63e-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"

## 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

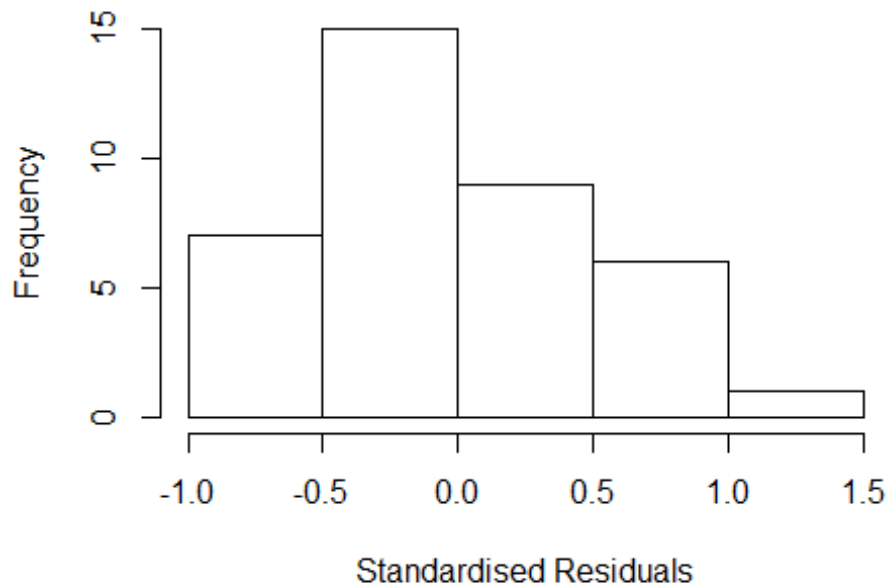
```

Residuals from first author



| | | | |
|---------------------|------|----|-------------------------|
| ## | GVIF | Df | $GVIF^{1/(2 \cdot Df)}$ |
| ## LastAuthorFemale | NaN | 1 | NaN |
| ## Year | NaN | 14 | NaN |

Residuals from last author



```

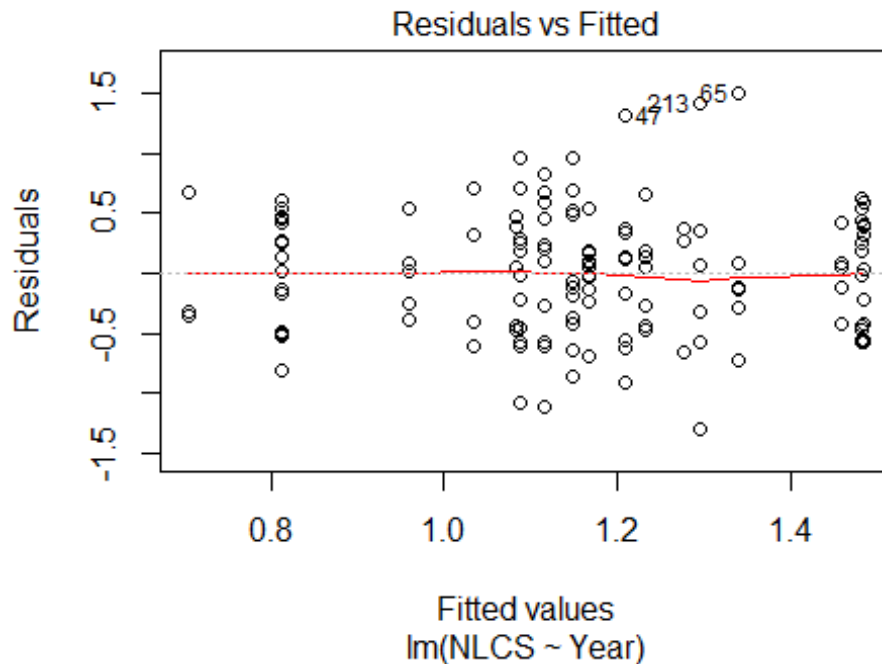
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0040 -0.4191 -0.0824 0.3740 1.4884
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5470 0.0000 Inf <2e-16 ***
## LastAuthorFemale1 -0.5951 0.2573 -2.31 0.030 *
## Year1998 0.2993 0.1508 1.98 0.060 .
## Year1999 -1.5470 0.0000 -Inf <2e-16 ***
## Year2000 -0.1100 0.0000 -Inf <2e-16 ***
## Year2002 -0.3665 0.1372 -2.67 0.014 *
## Year2003 -0.2850 1.0507 -0.27 0.789
## Year2004 -1.1293 0.4101 -2.75 0.012 *
## Year2005 -1.1160 0.0000 -Inf <2e-16 ***
## Year2006 0.0639 0.3538 0.18 0.858
## Year2007 -0.1465 0.3305 -0.44 0.662
## Year2008 -0.3680 0.5849 -0.63 0.536
## Year2009 -0.2714 0.3207 -0.85 0.406
## Year2010 -0.8014 0.3222 -2.49 0.021 *
## Year2011 -0.4383 0.4031 -1.09 0.289
## Year2012 -1.1860 0.0000 -Inf <2e-16 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.789
## Multiple R-squared: 0.37, Adjusted R-squared: -0.06
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 31 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.702 0.925 0.972 0.940 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-03 1.82e-12
## warn.limit.reject warn.limit.meanrw
## 5.00e-01 5.00e-01
## nResample max.it best.r.s k.fast.s k.max maxit.scale

```

```

##           500           50           2           1           1000           200
## trace.lev      mts compute.rd
##           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 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
##    5    3    3    5    6   14    9    5    8    9   13   13   19   15   15
## 2011 2012
##    9   14
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    3    3    5    4   10    9    5    7    8   11   11   16   13   13
## 2011 2012
##    7   10
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    3    3    5    4   10    9    5    5    8   11   10   15   13   12
## 2011 2012
##    7    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 = 17, df = 16, p-value = 0.4

```



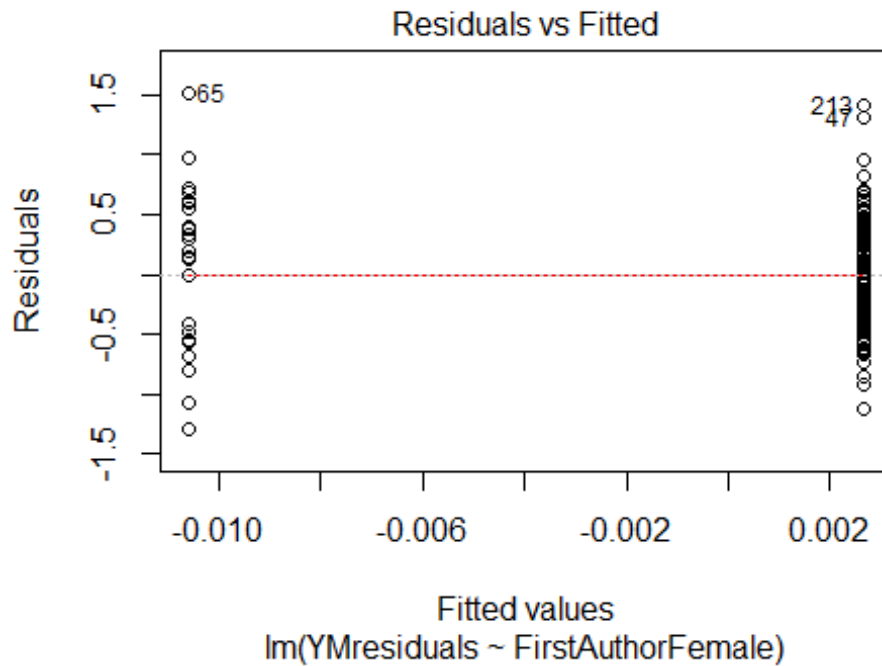
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 6.7, df = 1, p-value = 0.01

## [1] "Female first author team size 2018 geometric mean: 1.73851050644476"
## [1] "Male first author team size 2018 geometric mean: 1.97188300730021"

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

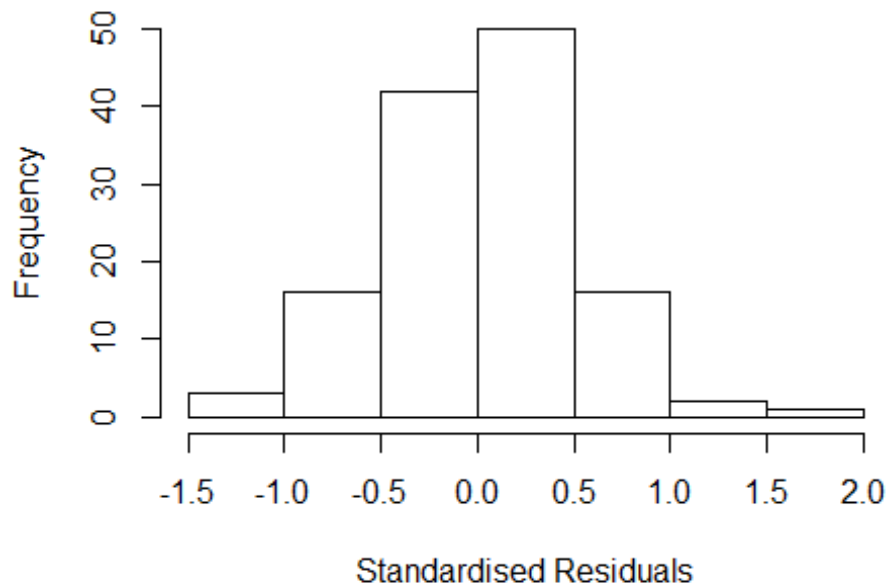
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 34, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.8612097182042"
## [1] "Male last author team size 2018 geometric mean: 1.89807715340301"

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



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 43, 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  4.565  1      2.137
## LastAuthorFemale   6.714  1      2.591
## UniqueAuthors    278.204  4      2.021
## Year              686.509 16      1.226
```

Residuals from first and last author and team size



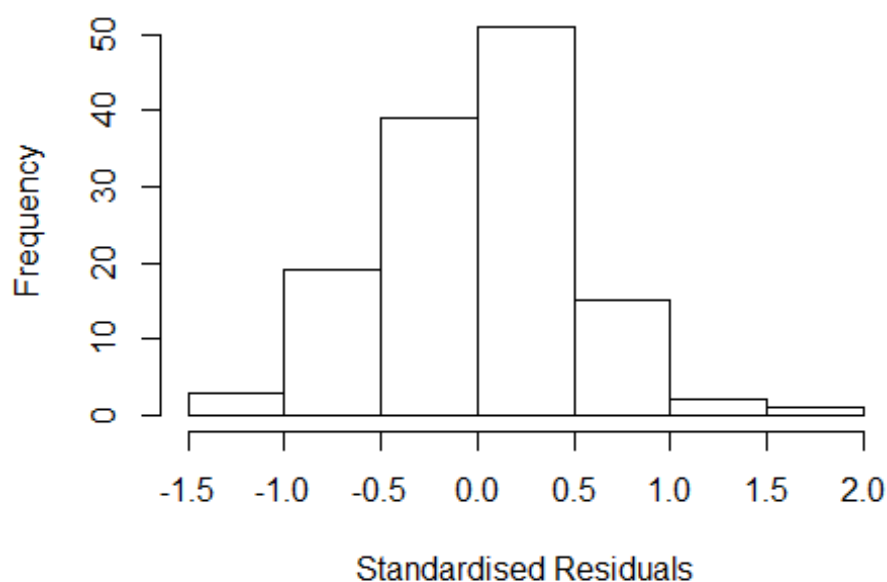
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1846 -0.3922 0.0215 0.3328 1.7077
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2634 0.2278 5.55 2.1e-07 ***
## FirstAuthorFemale1 -0.1106 0.2619 -0.42 0.67
## LastAuthorFemale1 0.0659 0.1930 0.34 0.73
## UniqueAuthors2 0.0613 0.1303 0.47 0.64
## UniqueAuthors3 -0.1634 0.1864 -0.88 0.38
## UniqueAuthors4 0.2343 0.6159 0.38 0.70
## UniqueAuthors5 0.0716 0.2570 0.28 0.78
## Year1997 -0.5892 0.4577 -1.29 0.20
## Year1998 0.0423 0.3797 0.11 0.91
## Year1999 0.1590 0.2541 0.63 0.53
```

```

## Year2000          -0.2310      0.4110   -0.56      0.58
## Year2001           0.2138      0.2736    0.78      0.44
## Year2002          -0.0835      0.2960   -0.28      0.78
## Year2003          -0.1940      0.3426   -0.57      0.57
## Year2004          -0.0797      0.3353   -0.24      0.81
## Year2005          -0.0338      0.2824   -0.12      0.91
## Year2006          -0.1254      0.2584   -0.49      0.63
## Year2007          -0.0340      0.3138   -0.11      0.91
## Year2008          -0.4241      0.2672   -1.59      0.12
## Year2009          -0.1500      0.2681   -0.56      0.58
## Year2010           0.0847      0.3239    0.26      0.79
## Year2011          -0.0847      0.3817   -0.22      0.82
## Year2012           0.3396      0.2505    1.36      0.18
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.51
## Multiple R-squared:  0.17,   Adjusted R-squared:  -0.000751
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 121 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.239  0.897  0.949   0.912  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      7.69e-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 3.082 1          1.756
## LastAuthorFemale  3.107 1          1.763
## Year              3.802 16          1.043

```


Residuals from first and last author



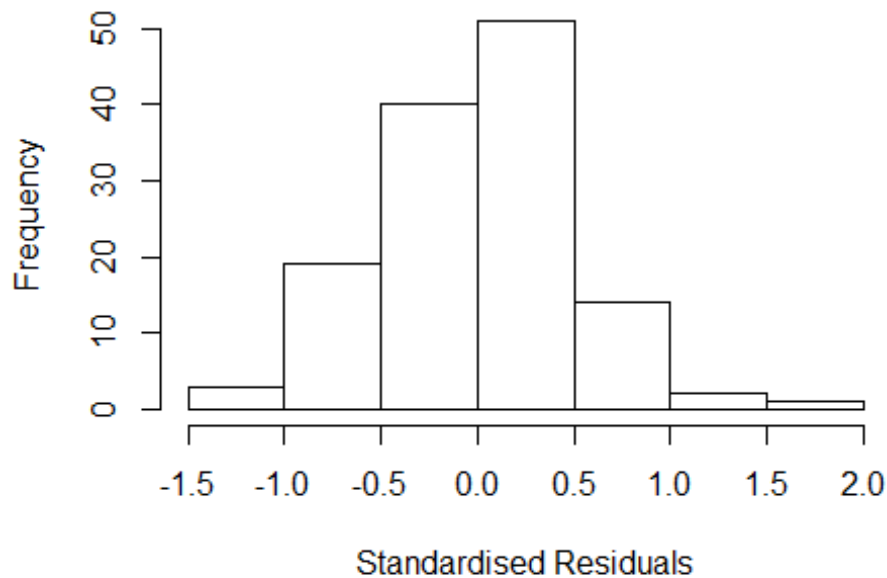
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2695 -0.4242  0.0538  0.3550  1.6559
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.255445   0.219583    5.72 9.2e-08 ***
## FirstAuthorFemale1 -0.069867   0.223534   -0.31  0.76
## LastAuthorFemale1  0.040978   0.163623    0.25  0.80
## Year1997        -0.580930   0.382859   -1.52  0.13
## Year1998         0.046241   0.369103    0.13  0.90
## Year1999         0.195852   0.252315    0.78  0.44
## Year2000        -0.223888   0.393092   -0.57  0.57
## Year2001         0.231352   0.264630    0.87  0.38
## Year2002        -0.080520   0.296357   -0.27  0.79
## Year2003        -0.169333   0.293126   -0.58  0.56
## Year2004         0.000568   0.353345    0.00  1.00
## Year2005        -0.016282   0.250692   -0.06  0.95
```

```

## Year2006      -0.086231    0.236128   -0.37    0.72
## Year2007      -0.048032    0.295261   -0.16    0.87
## Year2008      -0.419197    0.253248   -1.66    0.10
## Year2009      -0.141523    0.260690   -0.54    0.59
## Year2010       0.014040    0.286234    0.05    0.96
## Year2011      -0.007678    0.402101   -0.02    0.98
## Year2012       0.341857    0.251182    1.36    0.18
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.542
## Multiple R-squared:  0.149, Adjusted R-squared:  0.0114
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 120 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.330  0.909   0.945   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      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 1.722 1      1.312
## Year              1.722 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
## -1.2719 -0.4306 0.0505 0.3582 1.6221
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.262852 0.226916 5.57 1.8e-07 ***
## FirstAuthorFemale1 -0.043704 0.170701 -0.26 0.80
## Year1997 -0.589866 0.388497 -1.52 0.13
## Year1998 0.040114 0.376805 0.11 0.92
## Year1999 0.195887 0.260683 0.75 0.45
## Year2000 -0.229327 0.402701 -0.57 0.57
## Year2001 0.231262 0.267631 0.86 0.39
## Year2002 -0.083745 0.302033 -0.28 0.78
## Year2003 -0.176638 0.298818 -0.59 0.56
## Year2004 0.000771 0.333847 0.00 1.00
## Year2005 -0.027400 0.254193 -0.11 0.91
## Year2006 -0.083617 0.241466 -0.35 0.73
```

```

## Year2007          -0.049766    0.296532   -0.17     0.87
## Year2008          -0.420592    0.258886   -1.62     0.11
## Year2009          -0.147936    0.269047   -0.55     0.58
## Year2010           0.009017    0.288754    0.03     0.98
## Year2011           0.004788    0.393530    0.01     0.99
## Year2012           0.343804    0.258576    1.33     0.19
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.531
## Multiple R-squared:  0.15,   Adjusted R-squared:  0.0207
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 119 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.330  0.900   0.941   0.915  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      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  1.6  1          1.265
## Year              1.6 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
## -1.2647 -0.4272  0.0539  0.3739  1.5893

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.23680    0.19920     6.21 9.3e-09 ***
## LastAuthorFemale1 0.00840    0.12216     0.07  0.945
## Year1997       -0.56667    0.37436    -1.51  0.133
## Year1998        0.06856    0.36476     0.19  0.851
## Year1999        0.22046    0.23374     0.94  0.348
## Year2000       -0.21328    0.38031    -0.56  0.576
## Year2001        0.24571    0.24998     0.98  0.328
## Year2002       -0.08379    0.28675    -0.29  0.771
## Year2003       -0.15039    0.28091    -0.54  0.593
## Year2004        0.01589    0.27897     0.06  0.955
## Year2005       -0.00769    0.23788    -0.03  0.974
## Year2006       -0.06540    0.21115    -0.31  0.757
## Year2007       -0.03276    0.27372    -0.12  0.905
## Year2008       -0.40199    0.23781    -1.69  0.094 .
## Year2009       -0.13437    0.25202    -0.53  0.595
## Year2010        0.02790    0.27768     0.10  0.920
## Year2011        0.01533    0.38543     0.04  0.968
## Year2012        0.35563    0.23833     1.49  0.138
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.512
## Multiple R-squared:  0.151, Adjusted R-squared:  0.0217
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 119 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.315  0.897  0.939  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      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] ""

```

```

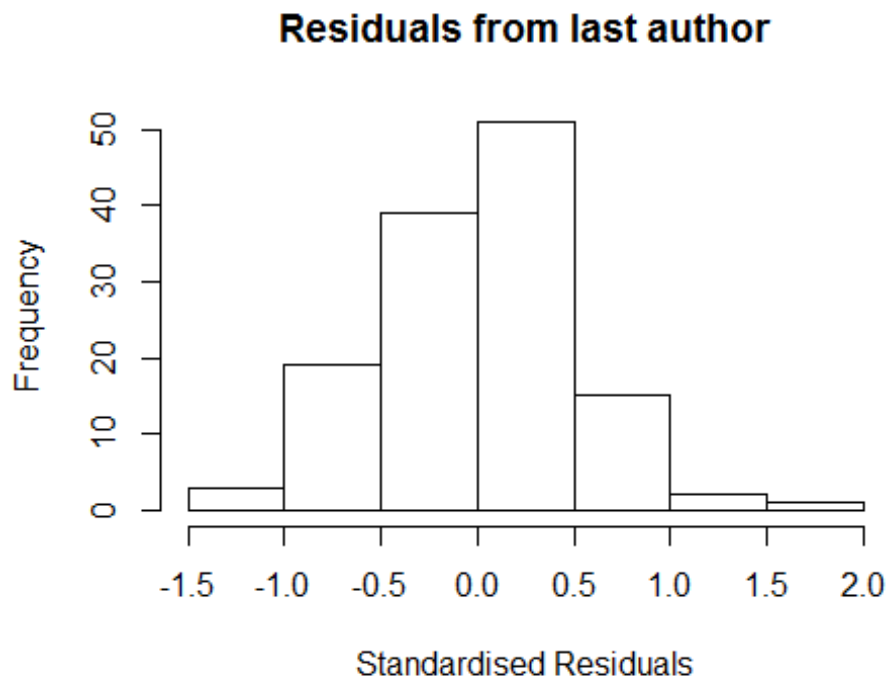
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1406"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    2    2    1    8    8    4    7    3   12   12   17   12   12   13
## 2012
##    11
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    2    0    1    6    6    4    6    3    9   10   17   10   11   10
## 2012
##     9
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    2    0    1    6    5    3    5    2    9    9   16   10   10   10
## 2012
##     8
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.91295063024394"
## [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 = 13, p-value = 0.2
## 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: 2.35215804504935"

## Warning in wilcox.test.default(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] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 4.200e+00  1          2.049
## LastAuthorFemale  3.834e+00  1          1.958
## UniqueAuthors    1.021e+26  3        21619.016
## Year              4.694e+26 14          8.965
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
```

```

##      Min      1Q   Median      3Q      Max
## -1.07492 -0.25310  0.00405  0.28387  1.14958
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    7.41e-01   7.76e-08  9.55e+06  <2e-16 ***
## FirstAuthorFemale1 -8.12e-02   1.39e-01 -5.80e-01   0.5621
## LastAuthorFemale1 -5.22e-02   1.18e-01 -4.40e-01   0.6604
## UniqueAuthors2    2.80e-01   1.31e-01  2.14e+00   0.0355 *
## UniqueAuthors3    1.61e-01   1.44e-01  1.12e+00   0.2675
## UniqueAuthors4    3.92e-01   2.15e-01  1.82e+00   0.0728 .
## Year1998         5.06e-01   1.75e-01  2.90e+00   0.0049 **
## Year2000        -4.26e-01   1.31e-01 -3.26e+00   0.0017 **
## Year2001         5.72e-01   2.11e-01  2.72e+00   0.0081 **
## Year2002         1.62e-01   5.36e-01  3.00e-01   0.7631
## Year2003        -2.22e-01   2.04e-01 -1.09e+00   0.2806
## Year2004         4.33e-01   1.78e-01  2.43e+00   0.0173 *
## Year2005        -8.64e-02   5.28e-01 -1.60e-01   0.8704
## Year2006        -1.82e-03   2.15e-01 -1.00e-02   0.9933
## Year2007         4.51e-02   1.50e-01  3.00e-01   0.7643
## Year2008         6.54e-01   1.47e-01  4.43e+00   3e-05 ***
## Year2009         7.77e-02   2.25e-01  3.40e-01   0.7313
## Year2010         5.37e-02   2.34e-01  2.30e-01   0.8193
## Year2011         4.17e-01   1.82e-01  2.29e+00   0.0246 *
## Year2012        -7.51e-03   2.18e-01 -3.00e-02   0.9727
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.427
## Multiple R-squared:  0.343, Adjusted R-squared:  0.181
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 83 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.448  0.870  0.947  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      1.03e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats

```

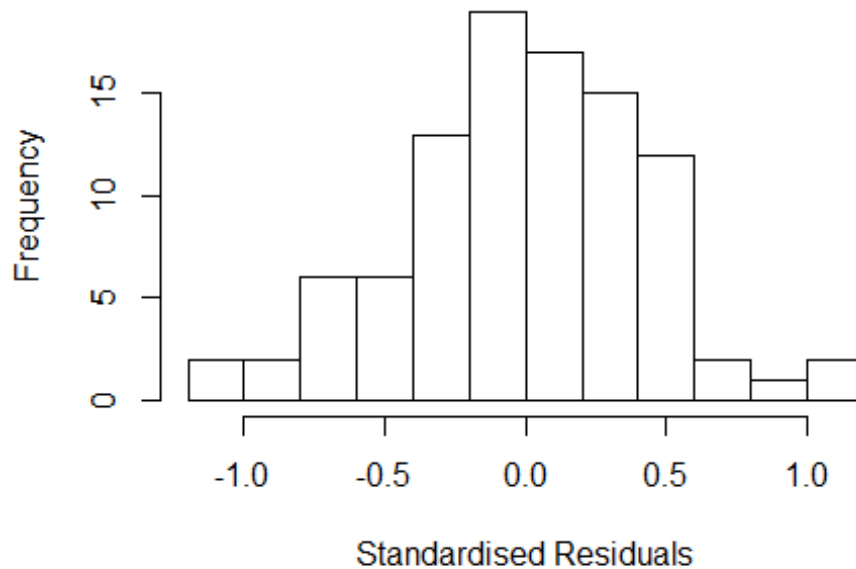


```
## "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"

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

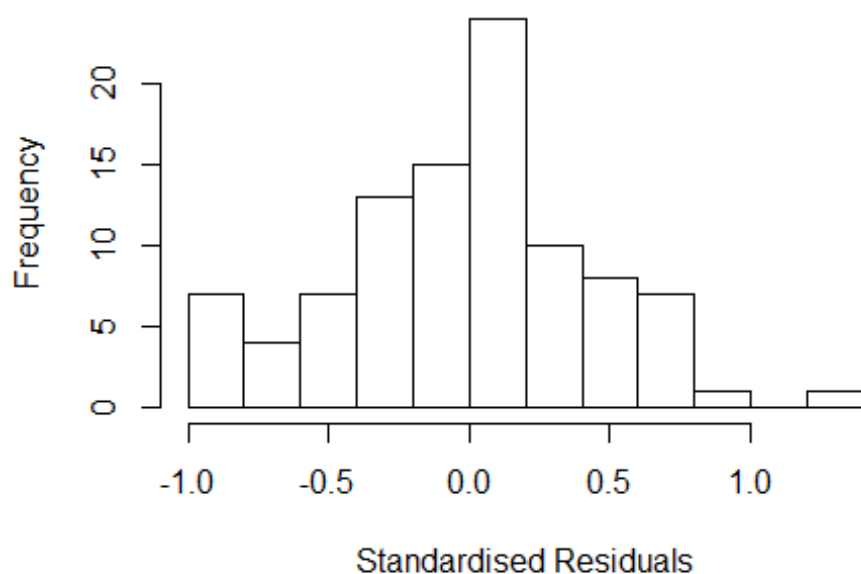
## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

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
```

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
## -0.9865 -0.2858 0.0178 0.2356 1.2368
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7410 0.0000 Inf < 2e-16 ***
## FirstAuthorFemale1 -0.1200 0.1323 -0.91 0.36708
## LastAuthorFemale1 -0.1000 0.1253 -0.80 0.42702
## Year1998 0.5065 0.1746 2.90 0.00479 **
## Year2000 -0.1460 0.0000 -Inf < 2e-16 ***
## Year2001 0.6831 0.1673 4.08 0.00010 ***
## Year2002 0.3887 0.4789 0.81 0.41945
## Year2003 -0.0619 0.2058 -0.30 0.76433
## Year2004 0.6444 0.1781 3.62 0.00052 ***
## Year2005 0.0970 0.7461 0.13 0.89688
## Year2006 0.1912 0.1686 1.13 0.25993
## Year2007 0.2670 0.1385 1.93 0.05739 .
```

```

## Year2008          0.8928      0.1082      8.25  2.6e-12 ***
## Year2009          0.3089      0.2097      1.47  0.14462
## Year2010          0.2455      0.2217      1.11  0.27149
## Year2011          0.6609      0.1239      5.33  8.8e-07 ***
## Year2012          0.2510      0.1781      1.41  0.16255
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.431
## Multiple R-squared:  0.294, Adjusted R-squared:  0.153
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 86 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.391  0.845   0.952   0.899   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      1.03e-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 2.386  1          1.545
## Year              2.386 14          1.032

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -0.9853 -0.3020  0.0468  0.2467  1.2290
##

```

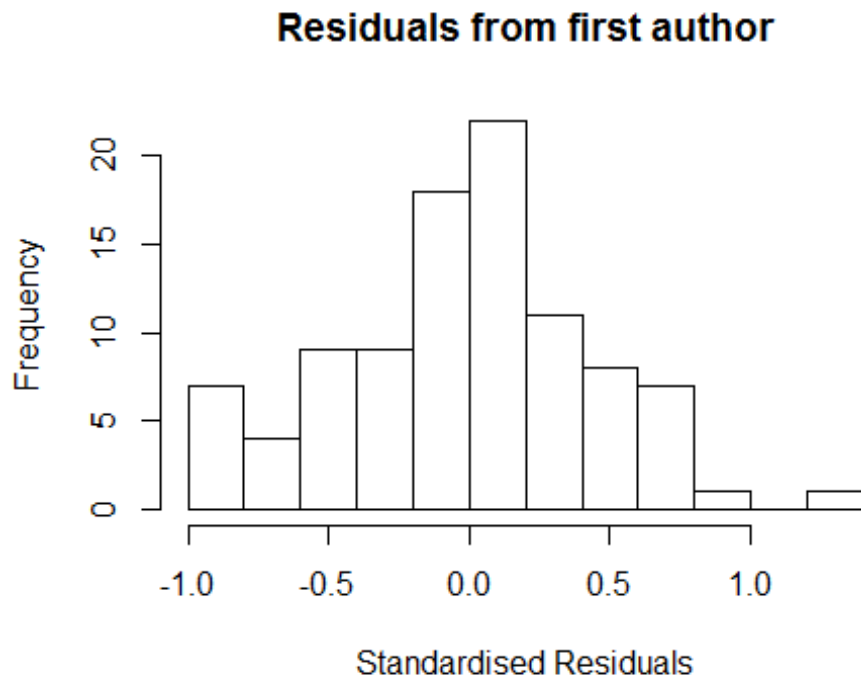
```

## Coefficients:
##               Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    7.41e-01   6.99e-08  1.06e+07 < 2e-16 ***
## FirstAuthorFemale1 -1.57e-01  1.16e-01 -1.35e+00  0.18075
## Year1998        5.06e-01  1.74e-01  2.91e+00  0.00473 **
## Year2000       -1.46e-01  5.34e-08 -2.73e+06 < 2e-16 ***
## Year2001        6.86e-01  1.63e-01  4.21e+00  6.6e-05 ***
## Year2002        4.02e-01  4.69e-01  8.60e-01  0.39343
## Year2003       -1.02e-01  2.14e-01 -4.70e-01  0.63671
## Year2004        6.23e-01  1.54e-01  4.03e+00  0.00013 ***
## Year2005        6.54e-02  8.37e-01  8.00e-02  0.93794
## Year2006        1.99e-01  1.68e-01  1.18e+00  0.24027
## Year2007        2.15e-01  1.06e-01  2.04e+00  0.04507 *
## Year2008        8.69e-01  1.08e-01  8.03e+00  6.6e-12 ***
## Year2009        2.71e-01  2.07e-01  1.31e+00  0.19401
## Year2010        2.44e-01  2.34e-01  1.05e+00  0.29906
## Year2011        6.04e-01  1.18e-01  5.13e+00  2.0e-06 ***
## Year2012        2.28e-01  1.90e-01  1.20e+00  0.23284
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.436
## Multiple R-squared:  0.287, Adjusted R-squared:  0.155
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 91 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.407  0.861  0.959  0.906  0.988  0.999
## Algorithmic parameters:
##           tuning.chi           bb           tuning.psi           refine.tol
##           1.55e+00           5.00e-01           4.69e+00           1.00e-07
##           rel.tol           solve.tol           eps.outlier           eps.x
##           1.00e-07           1.00e-07           1.03e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01           5.00e-01
##   nResample    max.it    best.r.s    k.fast.s    k.max maxit.scale
##           500         50         2         1        1000         200
##   trace.lev    mts    compute.rd
##           0        1000         0
##           psi           subsampling           cov
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"

## 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 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
## -1.0154 -0.2405  0.0166  0.2335  1.2596
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.7410    0.0000   Inf < 2e-16 ***
## LastAuthorFemale1 -0.1491    0.1091  -1.37  0.17562
## Year1998        0.5065    0.1749   2.90  0.00486 **
## Year2000       -0.1460    0.0000  -Inf < 2e-16 ***
## Year2001        0.6762    0.1825   3.71  0.00038 ***
```

```

## Year2002          0.3401      0.4534      0.75  0.45534
## Year2003         -0.1055      0.2165     -0.49  0.62728
## Year2004          0.6142      0.1937      3.17  0.00215 **
## Year2005          0.0616      0.8997      0.07  0.94561
## Year2006          0.1684      0.1612      1.04  0.29924
## Year2007          0.2669      0.1402      1.90  0.06046 .
## Year2008          0.8504      0.1050      8.10  4.8e-12 ***
## Year2009          0.3100      0.2221      1.40  0.16654
## Year2010          0.1731      0.1868      0.93  0.35669
## Year2011          0.6356      0.1235      5.15  1.8e-06 ***
## Year2012          0.2017      0.1528      1.32  0.19060
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.424
## Multiple R-squared:  0.292, Adjusted R-squared:  0.161
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 88 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.358  0.839   0.964   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      1.03e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 97"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 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
##    5  12    2    7    7   12    9    1    6    9   12   10   18   14   21
## 2011 2012

```

```

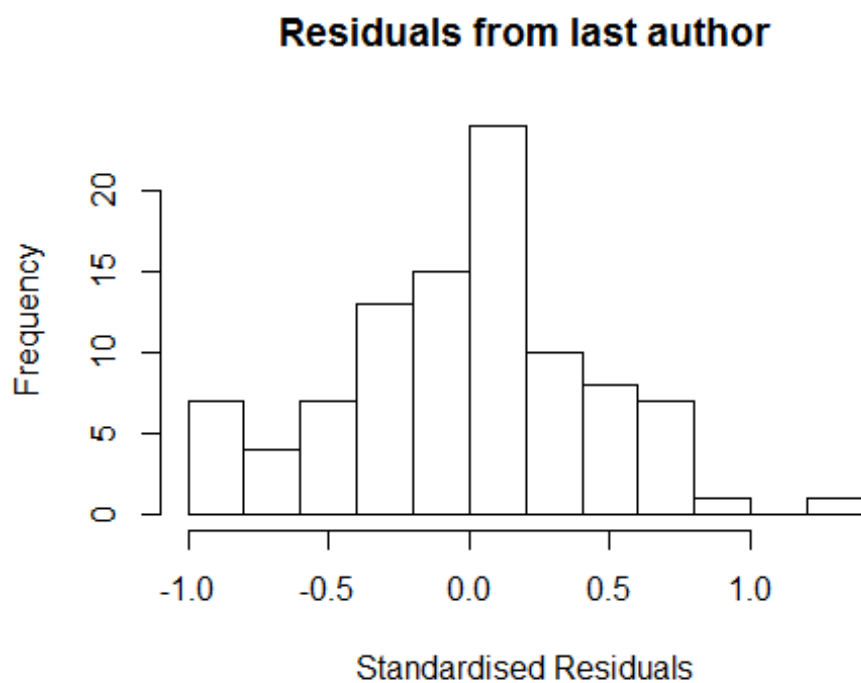
##    20    18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     5    12     2     6     6    11     8     1     5     8    12     9    16    13    18
## 2011 2012
##    17    14
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     3    12     2     6     6    11     8     1     5     8    11     8    14    11    17
## 2011 2012
##    16    13
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.90636858599387"
## [1] "Male first author team size 2018 geometric mean: 1.69838132956495"

## Warning in wilcox.test.default(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: 2.10743589934447"
## [1] "Male last 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

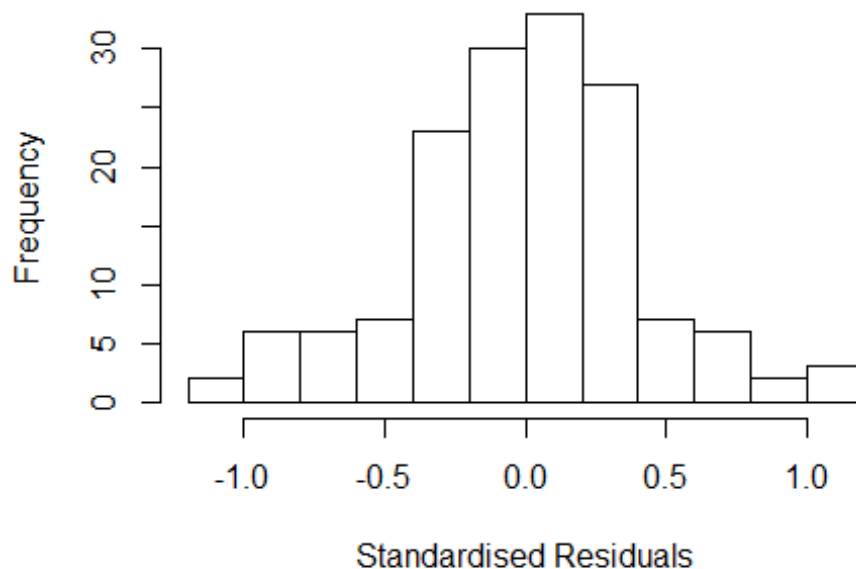
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 60, p-value = 0.07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
```

| | GVIF | Df | GVIF ^{1/(2*Df)} |
|----------------------|---------|----|--------------------------|
| ## FirstAuthorFemale | 5.377 | 1 | 2.319 |
| ## LastAuthorFemale | 4.094 | 1 | 2.023 |
| ## UniqueAuthors | 29.876 | 3 | 1.762 |
| ## Year | 135.341 | 16 | 1.166 |

Residuals from first and last author and team size



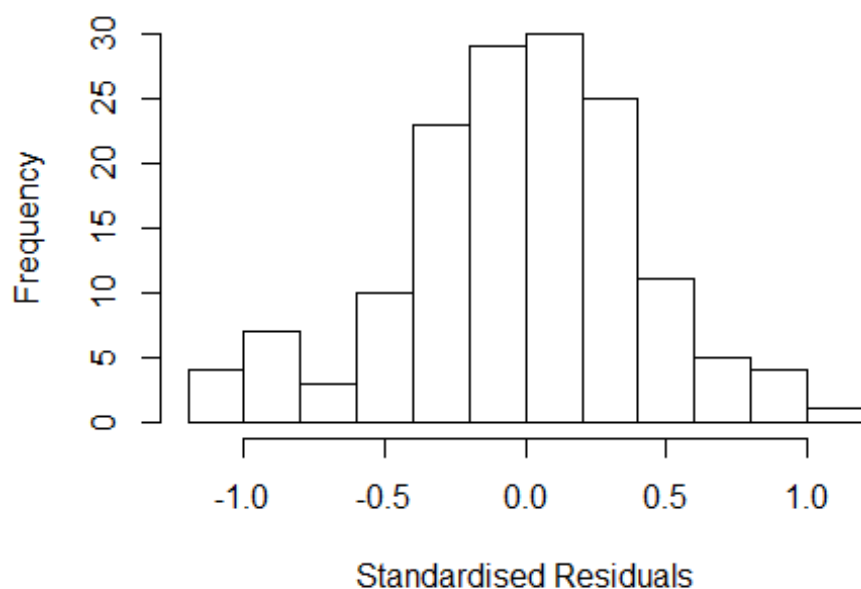
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1712 -0.2440 0.0119 0.2179 1.1223
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.409857 0.129290 10.90 < 2e-16 ***
## FirstAuthorFemale1 0.027905 0.121187 0.23 0.81825
## LastAuthorFemale1 -0.061080 0.104590 -0.58 0.56024
## UniqueAuthors2 -0.016921 0.096617 -0.18 0.86125
## UniqueAuthors3 0.156071 0.140157 1.11 0.26753
## UniqueAuthors4 -0.078296 0.212170 -0.37 0.71271
## Year1997 -0.430883 0.200150 -2.15 0.03318 *
## Year1998 -0.674305 0.186162 -3.62 0.00042 ***
## Year1999 -0.118158 0.136356 -0.87 0.38779
## Year2000 -0.988362 0.192757 -5.13 1.0e-06 ***
```

```

## Year2001      -0.323605    0.179614    -1.80    0.07391 .
## Year2002      -0.001013    0.171909    -0.01    0.99531
## Year2003      -0.694936    0.153832    -4.52    1.4e-05 ***
## Year2004      -0.393256    0.122733    -3.20    0.00170 **
## Year2005      -0.635399    0.141589    -4.49    1.6e-05 ***
## Year2006      -0.428796    0.158466    -2.71    0.00773 **
## Year2007      -0.221265    0.227008    -0.97    0.33152
## Year2008      -0.262720    0.151607    -1.73    0.08548 .
## Year2009       0.032532    0.219915     0.15    0.88263
## Year2010      -0.059996    0.174105    -0.34    0.73096
## Year2011       0.000448    0.161429     0.00    0.99779
## Year2012      -0.637172    0.362934    -1.76    0.08151 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.365
## Multiple R-squared:  0.332, Adjusted R-squared:  0.224
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 142 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.283  0.880  0.956  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      6.58e-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 4.798 1      2.190
## LastAuthorFemale  3.384 1      1.839
## Year              7.785 16      1.066

```

Residuals from first and last author



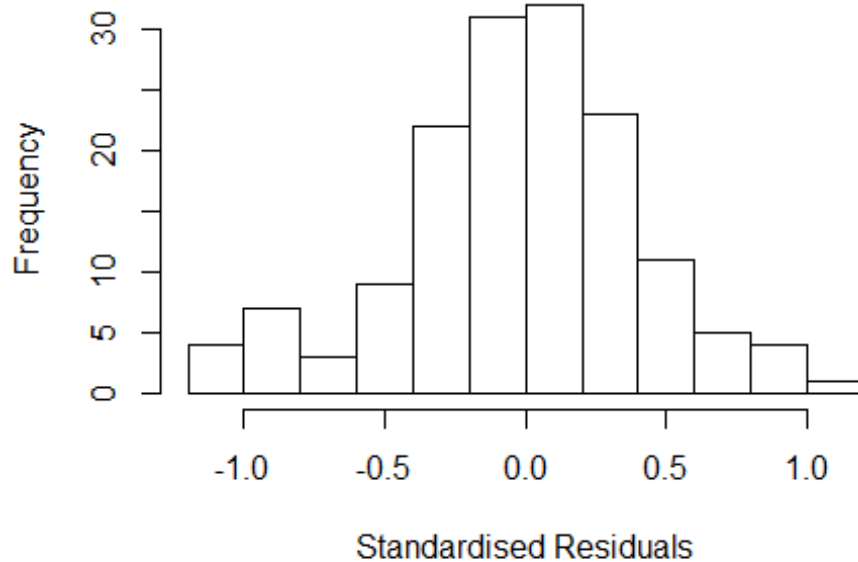
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.12580 -0.24811  0.00112  0.22048  1.05621
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.40557    0.13255   10.60 < 2e-16 ***
## FirstAuthorFemale1 -0.00421    0.11950   -0.04  0.97197
## LastAuthorFemale1 -0.02266    0.10414   -0.22  0.82810
## Year1997        -0.37773    0.19067   -1.98  0.04964 *
## Year1998        -0.59514    0.24299   -2.45  0.01562 *
## Year1999        -0.11237    0.12976   -0.87  0.38809
## Year2000        -0.98281    0.19476   -5.05  1.4e-06 ***
## Year2001        -0.28935    0.17099   -1.69  0.09295 .
## Year2002         0.02145    0.16349    0.13  0.89580
## Year2003        -0.70757    0.13255   -5.34  3.9e-07 ***
## Year2004        -0.39833    0.11694   -3.41  0.00087 ***
## Year2005        -0.63696    0.14087   -4.52  1.3e-05 ***
```

```

## Year2006          -0.41822    0.15079   -2.77  0.00634 **
## Year2007          -0.19492    0.23791   -0.82  0.41409
## Year2008          -0.25291    0.15521   -1.63  0.10557
## Year2009           0.06171    0.25360    0.24  0.80811
## Year2010          -0.03771    0.17959   -0.21  0.83400
## Year2011           0.00247    0.16697    0.01  0.98820
## Year2012          -0.56678    0.29317   -1.93  0.05533 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.363
## Multiple R-squared:  0.311, Adjusted R-squared:  0.218
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 144 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.315  0.877  0.955  0.883  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      6.58e-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 5.121 1      2.263
## Year              5.121 16      1.052

```

Residuals from first author



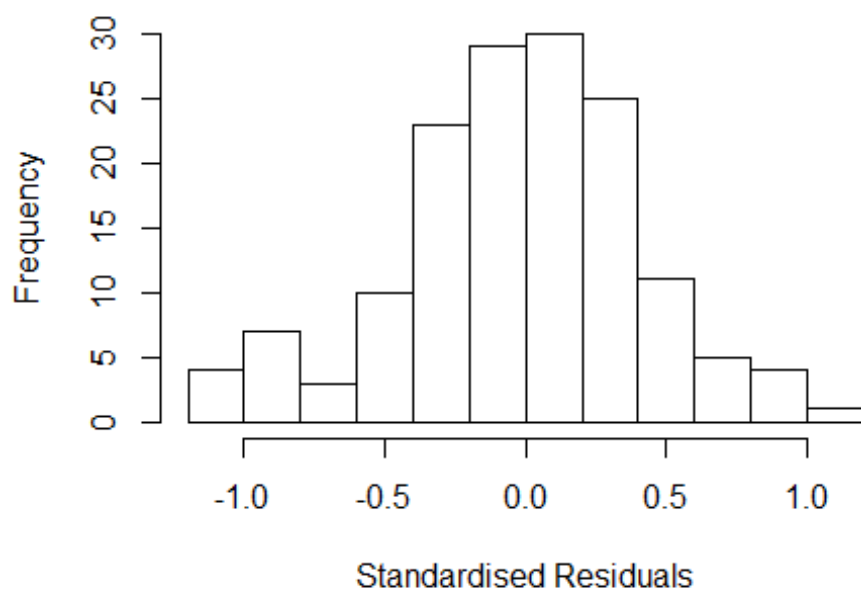
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.136656 -0.241244 0.000026 0.223632 1.070266
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.39706 0.12204 11.45 < 2e-16 ***
## FirstAuthorFemale1 -0.01432 0.12281 -0.12 0.90735
## Year1997 -0.37029 0.18547 -2.00 0.04790 *
## Year1998 -0.59290 0.24813 -2.39 0.01827 *
## Year1999 -0.10233 0.11474 -0.89 0.37406
## Year2000 -0.97211 0.18189 -5.34 3.8e-07 ***
## Year2001 -0.28596 0.17044 -1.68 0.09572 .
## Year2002 0.02658 0.15793 0.17 0.86662
## Year2003 -0.69906 0.12204 -5.73 6.4e-08 ***
## Year2004 -0.39503 0.11371 -3.47 0.00069 ***
## Year2005 -0.63120 0.13431 -4.70 6.4e-06 ***
## Year2006 -0.41451 0.14890 -2.78 0.00615 **
```

```

## Year2007          -0.18626    0.22762   -0.82  0.41466
## Year2008          -0.24608    0.14903   -1.65  0.10103
## Year2009           0.07289    0.24614    0.30  0.76760
## Year2010          -0.02786    0.17259   -0.16  0.87200
## Year2011           0.00717    0.16452    0.04  0.96531
## Year2012          -0.57232    0.29735   -1.92  0.05638 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.357
## Multiple R-squared:  0.316, Adjusted R-squared:  0.229
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 142 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.290  0.874  0.956  0.878  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      6.58e-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 3.635 1          1.907
## Year            3.635 16          1.041

```

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
## -1.12958 -0.24974  0.00152  0.22127  1.06169
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.40356    0.11814   11.88 < 2e-16 ***
## LastAuthorFemale1 -0.02388    0.10962   -0.22  0.8279
## Year1997        -0.37576    0.18628   -2.02  0.0457 *
## Year1998        -0.59462    0.24263   -2.45  0.0155 *
## Year1999        -0.11227    0.13053   -0.86  0.3913
## Year2000        -0.98161    0.19316   -5.08  1.2e-06 ***
## Year2001        -0.28950    0.16943   -1.71  0.0898 .
## Year2002         0.02302    0.15912    0.14  0.8852
## Year2003        -0.70556    0.11814   -5.97  2.0e-08 ***
## Year2004        -0.39756    0.11455   -3.47  0.0007 ***
## Year2005        -0.63561    0.13311   -4.78  4.6e-06 ***
## Year2006        -0.41585    0.13360   -3.11  0.0023 **
```

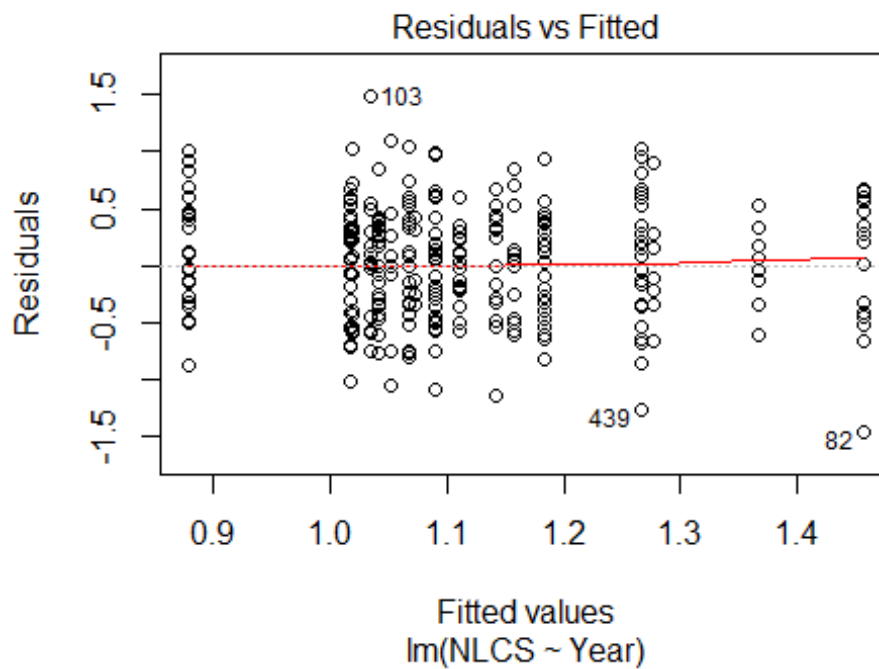
```

## Year2007          -0.19418      0.23476    -0.83    0.4096
## Year2008          -0.25009      0.14219    -1.76    0.0809 .
## Year2009           0.06422      0.23569     0.27    0.7857
## Year2010          -0.03466      0.17024    -0.20    0.8390
## Year2011           0.00427      0.15428     0.03    0.9780
## Year2012          -0.57025      0.28741    -1.98    0.0493 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.36
## Multiple R-squared:  0.313, Adjusted R-squared:  0.226
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 144 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.304  0.875  0.954  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      6.58e-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: 152"
## [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
##   10   6   8   10   16   22   16   14   15   26   25   24   35   34   36
## 2011 2012
##   27   29
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    8    5    7    8   13   15   14   14   13   25   17   21   29   28   27
## 2011 2012

```



```
## 22 24
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 7 4 7 8 13 14 13 13 12 23 17 20 28 27 25
## 2011 2012
## 19 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 = 18, df = 16, p-value = 0.3
```



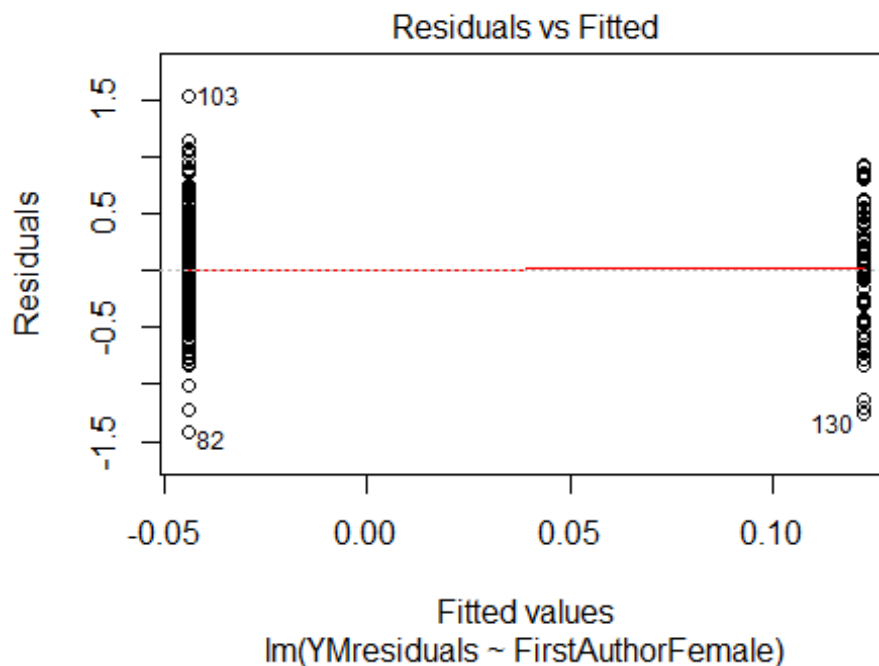
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.58, df = 1, p-value = 0.4

## [1] "Female first author team size 2018 geometric mean: 1.87614254491239"
## [1] "Male first author team size 2018 geometric mean: 1.782602457966"

## Warning in wilcox.test.default(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 = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.18998516052019"
## [1] "Male last author team size 2018 geometric mean: 1.52301909300492"

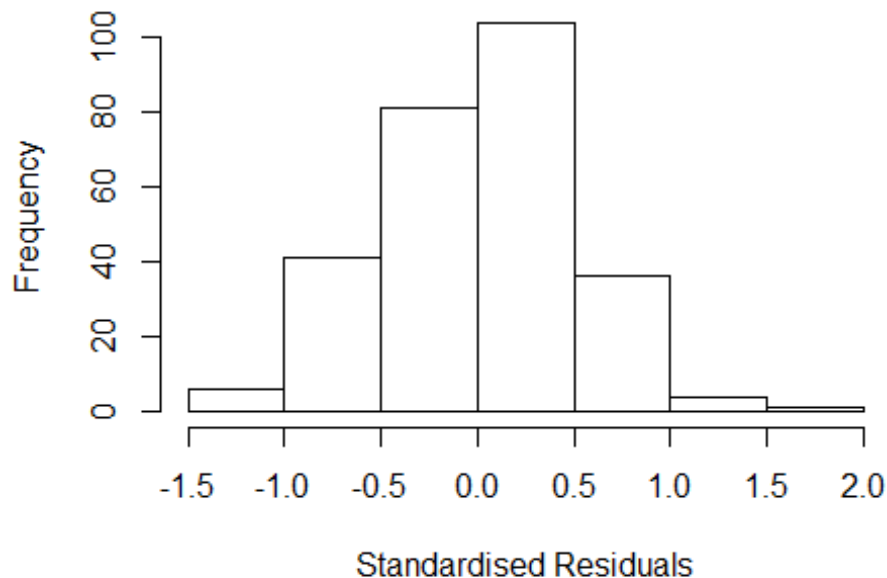
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 200, p-value = 0.05
## 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 | 2.319 | 1 | 1.523 |
| LastAuthorFemale | 2.674 | 1 | 1.635 |
| UniqueAuthors | 17.794 | 4 | 1.433 |
| Year | 20.207 | 16 | 1.098 |

Residuals from first and last author and team size



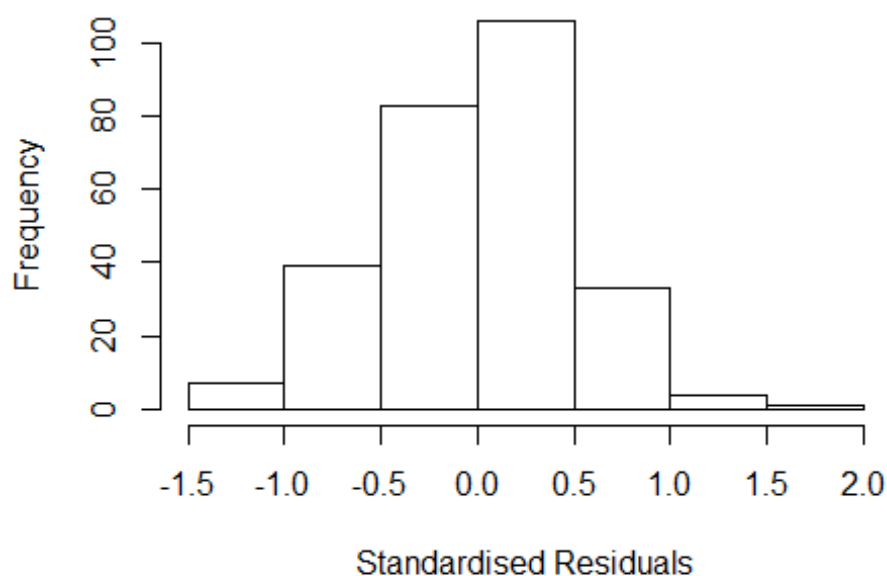
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4031 -0.3815 0.0299 0.3488 1.5599
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.99816 0.29916 3.34 0.00098 ***
## FirstAuthorFemale1 0.20857 0.09126 2.29 0.02313 *
## LastAuthorFemale1 -0.05198 0.09714 -0.54 0.59303
## UniqueAuthors2 0.04585 0.07269 0.63 0.52878
## UniqueAuthors3 0.11582 0.10437 1.11 0.26819
## UniqueAuthors4 0.19528 0.20779 0.94 0.34823
## UniqueAuthors5 0.01730 0.10399 0.17 0.86803
## Year1997 0.10346 0.33369 0.31 0.75679
## Year1998 0.22636 0.35337 0.64 0.52239
## Year1999 0.31133 0.33225 0.94 0.34965
```

```

## Year2000      -0.08653    0.32575   -0.27  0.79074
## Year2001      0.40494    0.35033    1.16  0.24883
## Year2002     -0.08288    0.33615   -0.25  0.80544
## Year2003      0.12801    0.33096    0.39  0.69925
## Year2004      0.07913    0.34620    0.23  0.81939
## Year2005     -0.03648    0.30801   -0.12  0.90582
## Year2006      0.03948    0.30647    0.13  0.89761
## Year2007     -0.04429    0.31975   -0.14  0.88994
## Year2008     -0.04298    0.31442   -0.14  0.89138
## Year2009     -0.00792    0.31708   -0.02  0.98010
## Year2010     -0.18509    0.31618   -0.59  0.55880
## Year2011      0.17367    0.31483    0.55  0.58168
## Year2012      0.20406    0.32877    0.62  0.53537
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.536
## Multiple R-squared:  0.105, Adjusted R-squared:  0.0267
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 248 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.378  0.884  0.952  0.917  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      3.66e-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 1.582 1      1.258
## LastAuthorFemale  1.730 1      1.315
## Year              1.391 16      1.010

```

Residuals from first and last author



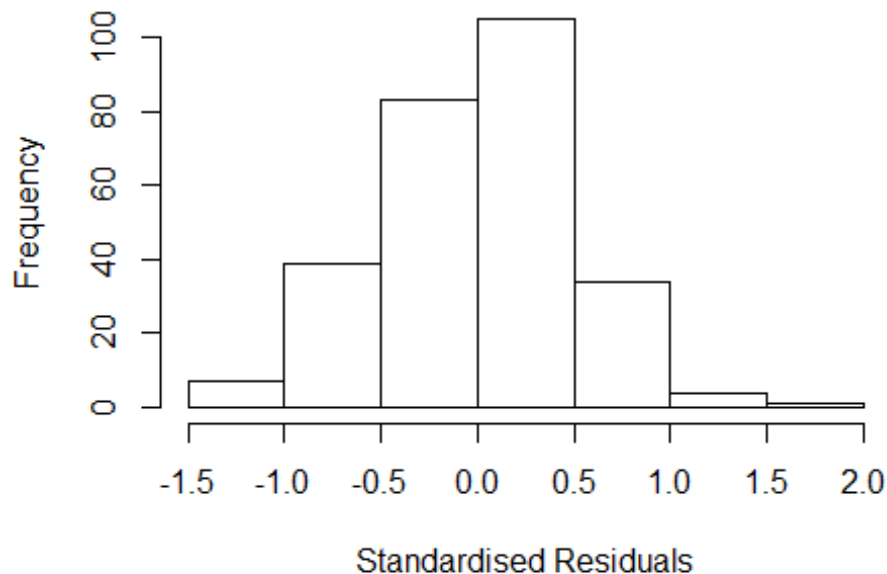
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4412 -0.3887 0.0478 0.3408 1.5782
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.03586 0.28135 3.68 0.00028 ***
## FirstAuthorFemale1 0.21732 0.09278 2.34 0.01994 *
## LastAuthorFemale1 -0.05784 0.09844 -0.59 0.55739
## Year1997 0.13738 0.31348 0.44 0.66159
## Year1998 0.19649 0.34012 0.58 0.56398
## Year1999 0.29209 0.31852 0.92 0.36001
## Year2000 -0.09788 0.31217 -0.31 0.75413
## Year2001 0.40534 0.33176 1.22 0.22292
## Year2002 -0.09308 0.32037 -0.29 0.77165
## Year2003 0.11207 0.30996 0.36 0.71797
## Year2004 0.07925 0.32724 0.24 0.80884
## Year2005 -0.05399 0.29233 -0.18 0.85361
```

```

## Year2006          0.04233      0.28808      0.15  0.88330
## Year2007          -0.04157      0.30759     -0.14  0.89261
## Year2008          -0.05300      0.29638     -0.18  0.85822
## Year2009          -0.00714      0.29890     -0.02  0.98095
## Year2010          -0.17251      0.30304     -0.57  0.56968
## Year2011           0.17642      0.29929      0.59  0.55609
## Year2012           0.20434      0.31451      0.65  0.51647
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.539
## Multiple R-squared:  0.0991, Adjusted R-squared:  0.0353
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 251 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.371  0.883   0.953   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      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 1.125 1      1.061
## Year              1.125 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
## -1.4274 -0.3834 0.0475 0.3482 1.5820
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.03076 0.27862 3.70 0.00026 ***
## FirstAuthorFemale1 0.18414 0.07852 2.35 0.01979 *
## Year1997 0.12836 0.31602 0.41 0.68496
## Year1998 0.19838 0.33916 0.58 0.55913
## Year1999 0.29814 0.31196 0.96 0.34013
## Year2000 -0.08918 0.30863 -0.29 0.77285
## Year2001 0.39664 0.33091 1.20 0.23179
## Year2002 -0.09176 0.31761 -0.29 0.77289
## Year2003 0.11539 0.30753 0.38 0.70782
## Year2004 0.07539 0.32325 0.23 0.81577
## Year2005 -0.04523 0.28917 -0.16 0.87584
## Year2006 0.03845 0.28605 0.13 0.89317
```

```

## Year2007          -0.03489    0.30613   -0.11  0.90934
## Year2008          -0.05098    0.29405   -0.17  0.86249
## Year2009          -0.00992    0.29642   -0.03  0.97333
## Year2010          -0.17122    0.30124   -0.57  0.57028
## Year2011           0.17141    0.29648    0.58  0.56366
## Year2012           0.20379    0.31253    0.65  0.51495
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.542
## Multiple R-squared:  0.0975, Adjusted R-squared:  0.0373
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 253 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.375  0.885   0.955   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      3.66e-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 1.184 1          1.088
## Year              1.184 16          1.005
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ 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 -0.4011  0.0394  0.3427  1.5613

```



```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.05032    0.28506   3.68 0.00028 ***
## LastAuthorFemale1 0.08326    0.08270   1.01 0.31496
## Year1997        0.08892    0.32720   0.27 0.78602
## Year1998        0.19455    0.34005   0.57 0.56773
## Year1999        0.31010    0.30733   1.01 0.31393
## Year2000       -0.06205    0.31900  -0.19 0.84594
## Year2001        0.40083    0.33661   1.19 0.23484
## Year2002       -0.09058    0.32676  -0.28 0.78185
## Year2003        0.12256    0.31828   0.39 0.70051
## Year2004        0.06327    0.32045   0.20 0.84365
## Year2005       -0.03402    0.29786  -0.11 0.90915
## Year2006        0.03806    0.29307   0.13 0.89678
## Year2007       -0.02267    0.31514  -0.07 0.94270
## Year2008       -0.04702    0.30042  -0.16 0.87575
## Year2009        0.00669    0.30183   0.02 0.98232
## Year2010       -0.16164    0.30959  -0.52 0.60204
## Year2011        0.18456    0.30141   0.61 0.54088
## Year2012        0.23935    0.32150   0.74 0.45727
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.553
## Multiple R-squared:  0.0789, Adjusted R-squared:  0.0175
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 253 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.405  0.896  0.953  0.921  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      3.66e-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 1409"
## [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    4    1    2    5    4    2    9    6    3    5    7    8
## 2012
##    8
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    1    0    4    1    2    3    3    0    8    6    2    5    7    6
## 2012
##    6
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    1    0    4    1    2    2    3    0    8    6    1    5    7    5
## 2012
##    6
## [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: 2.0800838230519"

## 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 = 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: 2.0800838230519"

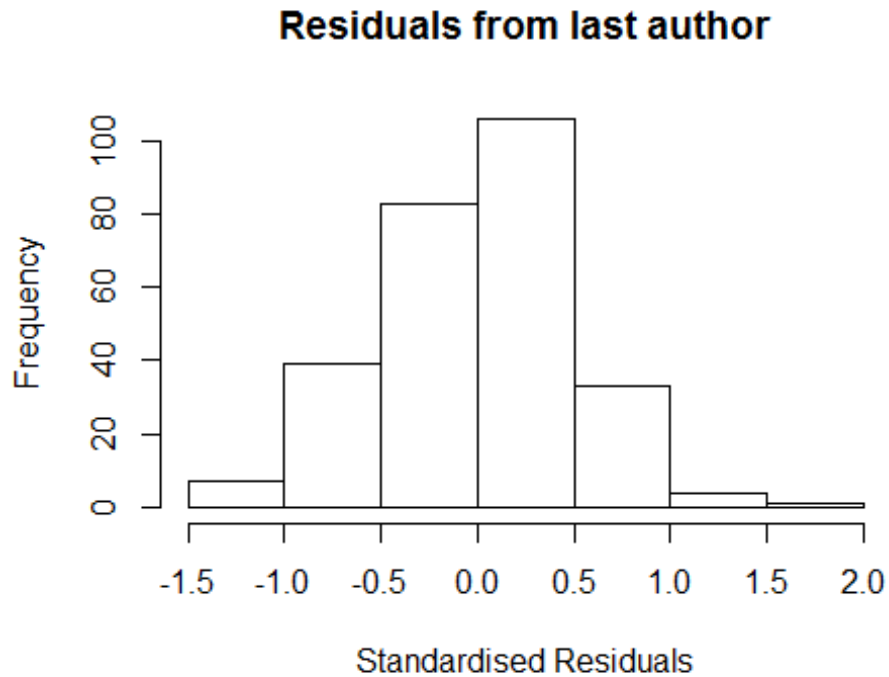
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 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"

```

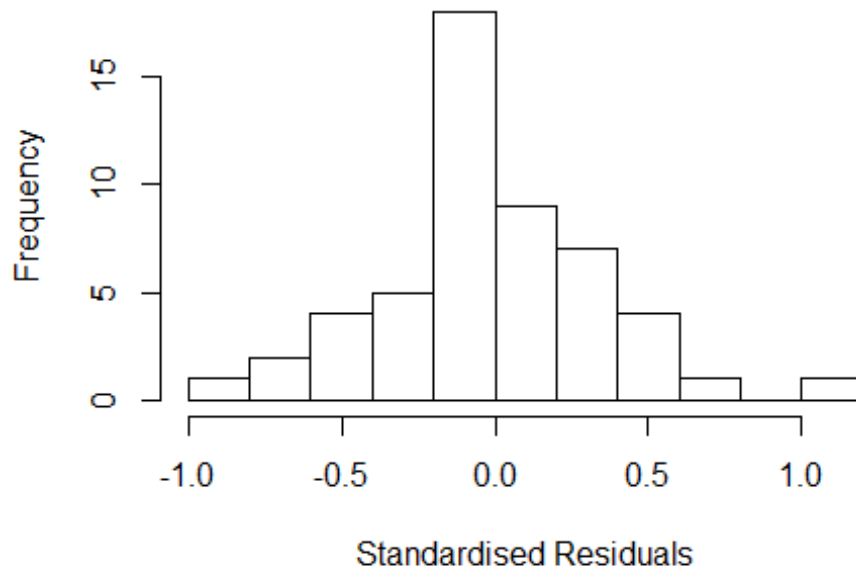
```
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

## 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 | 3 | NaN |
| ## Year | NaN | 13 | NaN |

Residuals from first and last author and team size



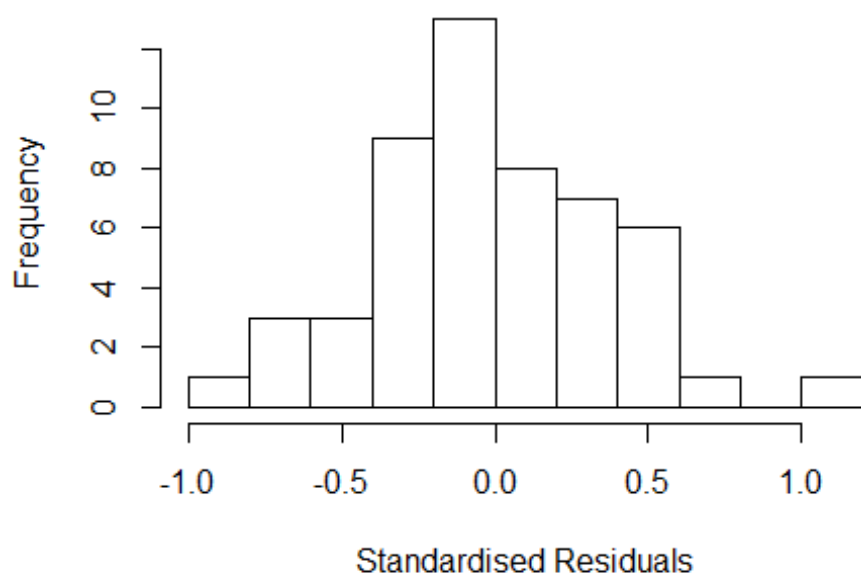
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -8.26e-01 -1.76e-01 -1.67e-16 2.07e-01 1.05e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.81800 0.00000 Inf < 2e-16 ***
## FirstAuthorFemale1 -0.06678 0.15536 -0.43 0.67012
## LastAuthorFemale1 0.10449 0.14947 0.70 0.48939
## UniqueAuthors2 0.00669 0.21161 0.03 0.97495
## UniqueAuthors3 0.12894 0.20411 0.63 0.53194
## UniqueAuthors5 -1.02748 0.20157 -5.10 1.4e-05 ***
## Year1998 -0.64600 0.00000 -Inf < 2e-16 ***
## Year2000 -1.03220 0.31324 -3.30 0.00236 **
## Year2001 -1.34000 0.00000 -Inf < 2e-16 ***
## Year2002 -0.87093 0.22807 -3.82 0.00056 ***
```

```

## Year2003      -0.96069    0.21323   -4.51  7.9e-05 ***
## Year2004      -1.19859    0.24914   -4.81  3.2e-05 ***
## Year2006      -0.78924    0.28640   -2.76  0.00946 **
## Year2007      -0.78690    0.24422   -3.22  0.00286 **
## Year2008       0.38008    0.24339    1.56  0.12792
## Year2009      -0.61101    0.20163   -3.03  0.00472 **
## Year2010      -0.72374    0.19773   -3.66  0.00087 ***
## Year2011      -0.72921    0.24756   -2.95  0.00587 **
## Year2012      -0.93352    0.22339   -4.18  0.00020 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.45
## Multiple R-squared:  0.426, Adjusted R-squared:  0.113
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 44 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.563  0.928  0.969   0.937   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 2: First author gender, Last author gender, Year as
factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 5.980e+15 1 7.733e+07
## LastAuthorFemale 1.455e+01 1 3.814e+00
## Year -1.077e+16 13 NaN

```

Residuals from first and last author



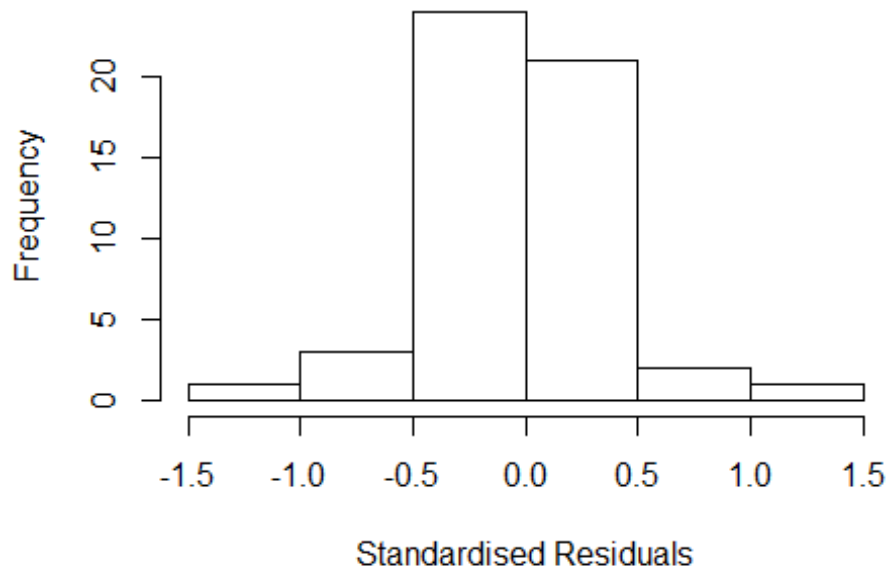
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -8.73e-01 -2.56e-01 -1.33e-15 2.34e-01 1.02e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.82e+00 2.59e-08 7.02e+07 < 2e-16 ***
## FirstAuthorFemale1 -1.40e-01 1.66e-01 -8.50e-01 0.40209
## LastAuthorFemale1 2.04e-01 1.66e-01 1.23e+00 0.22838
## Year1998 -6.46e-01 2.96e-08 -2.18e+07 < 2e-16 ***
## Year2000 -1.04e+00 3.13e-01 -3.32e+00 0.00206 **
## Year2001 -1.34e+00 3.00e-08 -4.46e+07 < 2e-16 ***
## Year2002 -7.66e-01 1.91e-01 -4.02e+00 0.00029 ***
## Year2003 -9.54e-01 2.62e-02 -3.64e+01 < 2e-16 ***
## Year2004 -1.24e+00 2.01e-01 -6.15e+00 4.3e-07 ***
## Year2006 -7.85e-01 1.53e-01 -5.14e+00 9.7e-06 ***
## Year2007 -7.91e-01 2.06e-01 -3.84e+00 0.00047 ***
## Year2008 4.60e-01 1.66e-01 2.78e+00 0.00858 **
```

```

## Year2009          -5.52e-01    1.92e-01 -2.88e+00    0.00671 **
## Year2010          -8.05e-01    2.12e-01 -3.80e+00    0.00054 ***
## Year2011          -7.54e-01    2.13e-01 -3.54e+00    0.00113 **
## Year2012          -9.19e-01    2.23e-01 -4.12e+00    0.00021 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.44
## Multiple R-squared:  0.335, Adjusted R-squared:  0.0581
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.568  0.887  0.950  0.915  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      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 5.079e+16 1      2.254e+08
## Year              5.079e+16 13      4.391e+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
## -1.07e+00 -2.07e-01 -9.44e-16 2.30e-01 1.09e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.82e+00 2.80e-08 6.50e+07 < 2e-16 ***
## FirstAuthorFemale1 -3.60e-02 1.51e-01 -2.40e-01 0.81211
## Year1998 -6.46e-01 3.09e-08 -2.09e+07 < 2e-16 ***
## Year2000 -1.00e+00 3.32e-01 -3.01e+00 0.00464 **
## Year2001 -1.34e+00 3.83e-08 -3.50e+07 < 2e-16 ***
## Year2002 -8.18e-01 1.51e-01 -5.43e+00 3.7e-06 ***
## Year2003 -9.54e-01 2.62e-02 -3.64e+01 < 2e-16 ***
## Year2004 -1.13e+00 1.73e-01 -6.57e+00 1.1e-07 ***
## Year2006 -7.52e-01 1.46e-01 -5.16e+00 8.7e-06 ***
## Year2007 -7.04e-01 1.68e-01 -4.20e+00 0.00016 ***
## Year2008 3.56e-01 1.51e-01 2.36e+00 0.02339 *
## Year2009 -4.90e-01 1.71e-01 -2.87e+00 0.00671 **
```



```

## Year2010          -7.13e-01   1.92e-01 -3.72e+00   0.00066 ***
## Year2011          -6.10e-01   1.62e-01 -3.76e+00   0.00059 ***
## Year2012          -8.88e-01   2.30e-01 -3.86e+00   0.00043 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.42
## Multiple R-squared:  0.322, Adjusted R-squared:  0.0661
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 44 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.482  0.897  0.968  0.914  0.984  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.92e-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 -28411 1      NaN
## Year              -28411 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
## -0.967551 -0.228727 -0.000722  0.211406  1.001560
##
## Coefficients:
##      Estimate Std. Error  t value Pr(>|t|)

```

```

## (Intercept)      1.82e+00    1.51e-08    1.20e+08    < 2e-16 ***
## LastAuthorFemale1 1.25e-01    1.57e-01    8.00e-01    0.42883
## Year1998          -6.46e-01    2.23e-08   -2.89e+07    < 2e-16 ***
## Year2000          -1.05e+00    3.08e-01   -3.43e+00    0.00151 **
## Year2001          -1.34e+00    1.13e-08   -1.19e+08    < 2e-16 ***
## Year2002          -8.37e-01    1.17e-01   -7.17e+00    1.7e-08 ***
## Year2003          -9.54e-01    2.62e-02   -3.64e+01    < 2e-16 ***
## Year2004          -1.22e+00    2.27e-01   -5.40e+00    4.1e-06 ***
## Year2006          -8.26e-01    1.71e-01   -4.83e+00    2.4e-05 ***
## Year2007          -8.05e-01    2.04e-01   -3.94e+00    0.00034 ***
## Year2008           3.20e-01    1.80e-08    1.78e+07    < 2e-16 ***
## Year2009          -5.92e-01    1.92e-01   -3.08e+00    0.00384 **
## Year2010          -8.50e-01    2.31e-01   -3.68e+00    0.00074 ***
## Year2011          -7.08e-01    1.77e-01   -3.99e+00    0.00030 ***
## Year2012          -9.39e-01    2.26e-01   -4.15e+00    0.00019 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.42
## Multiple R-squared:  0.328, Adjusted R-squared:  0.0739
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 41 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.549  0.865  0.954  0.910  0.987  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           1.92e-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 1410"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"

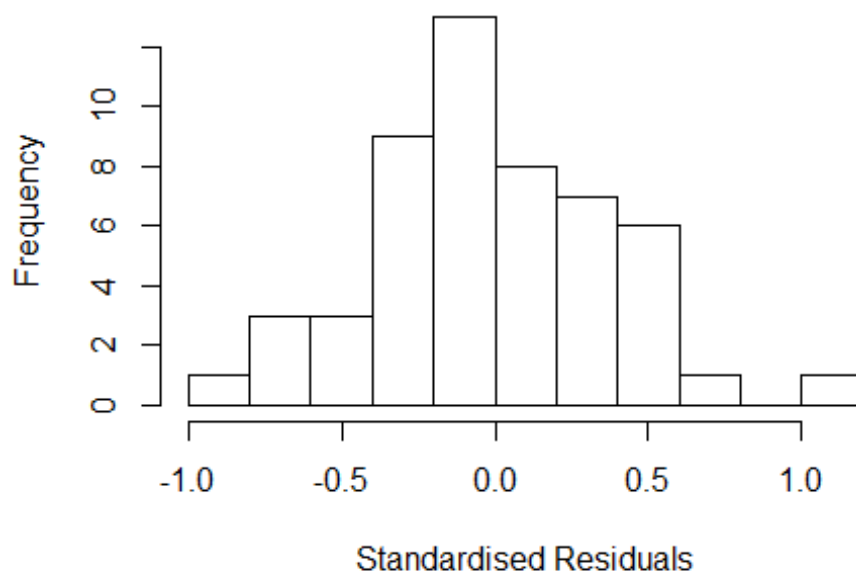
```

```
##
## 1996 1997 1998 2001 2002 2003 2004 2005 2007 2008 2009 2010 2011 2012
##    1    3    2    2    1    4    2    2    3    4    2    4    2    3
##
## 1996 1997 1998 2001 2002 2003 2004 2005 2007 2008 2009 2010 2011 2012
##    1    3    2    1    1    4    2    2    3    4    1    4    2    2
##
## 1996 1997 1998 2001 2002 2003 2004 2005 2007 2008 2009 2010 2011 2012
##    1    3    2    0    1    4    1    1    3    4    1    4    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: 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

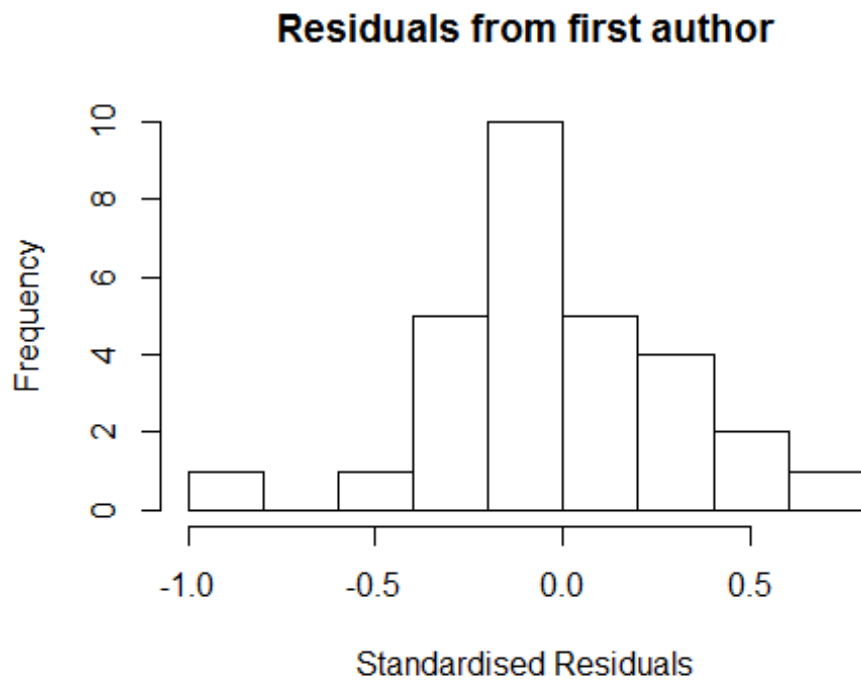


```
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1             NaN
## Year           NaN 12             NaN

## [1] "Regression 4: Last author gender, Year as factors"

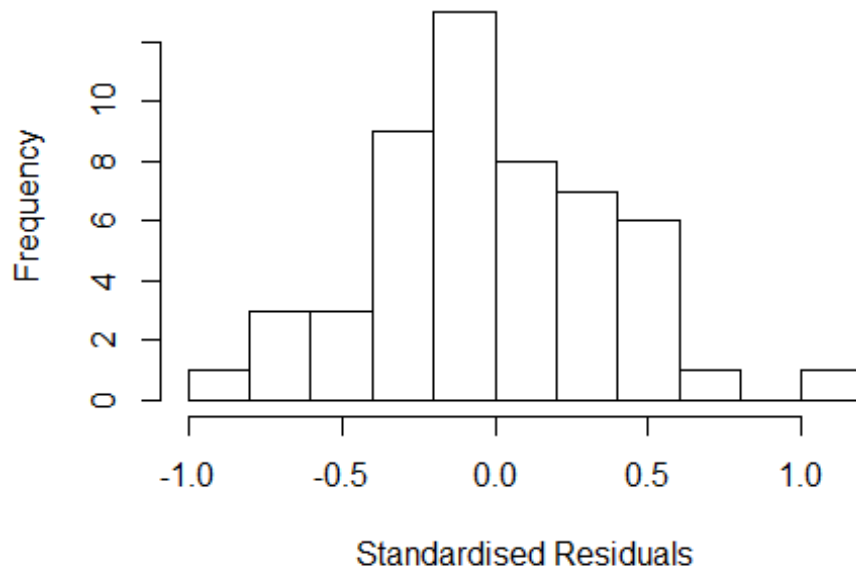
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

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



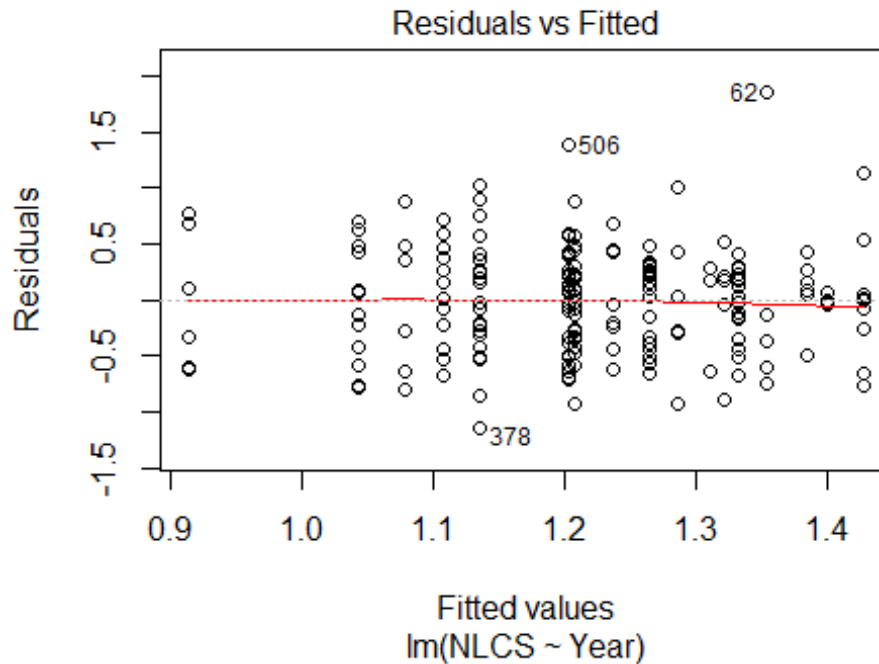
```
##              GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale NaN 1             NaN
## Year           NaN 12             NaN
```

Residuals from last author



```
## [1] "Sample size for the above analysis: 29"
## [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
## 18 12 13 17 12 19 12 17 22 15 15 22 34 31 44
## 2011 2012
## 45 42
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 8 4 6 5 6 4 6 13 13 7 5 9 22 21 24
## 2011 2012
## 24 26
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 7 4 6 4 5 3 5 11 11 6 5 9 19 19 20
## 2011 2012
## 22 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 = 34, df = 16, p-value = 0.006
```



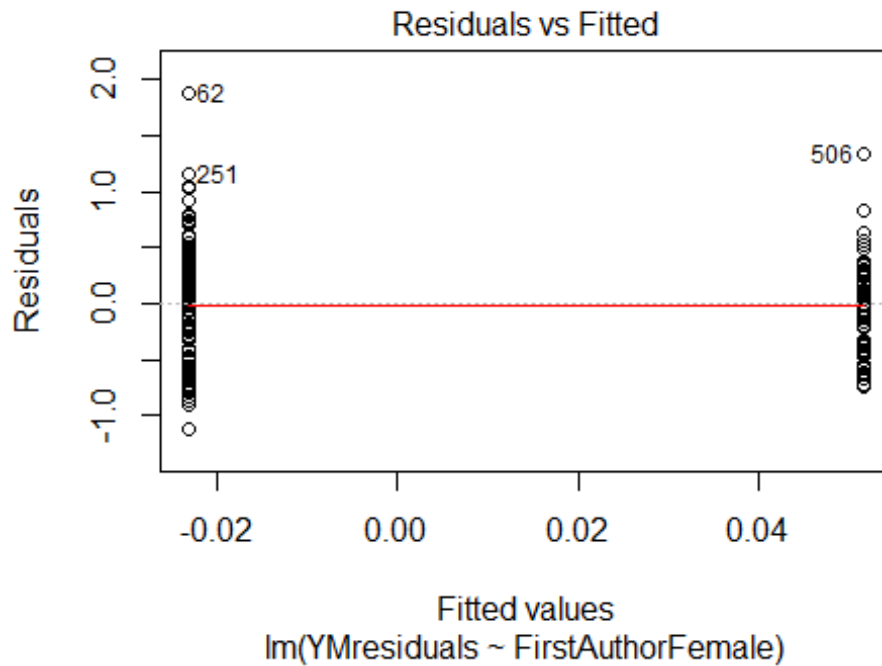
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.3, df = 1, p-value = 0.3

## [1] "Female first author team size 2018 geometric mean: 4.49400889557082"
## [1] "Male first author team size 2018 geometric mean: 3.53518946489351"

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

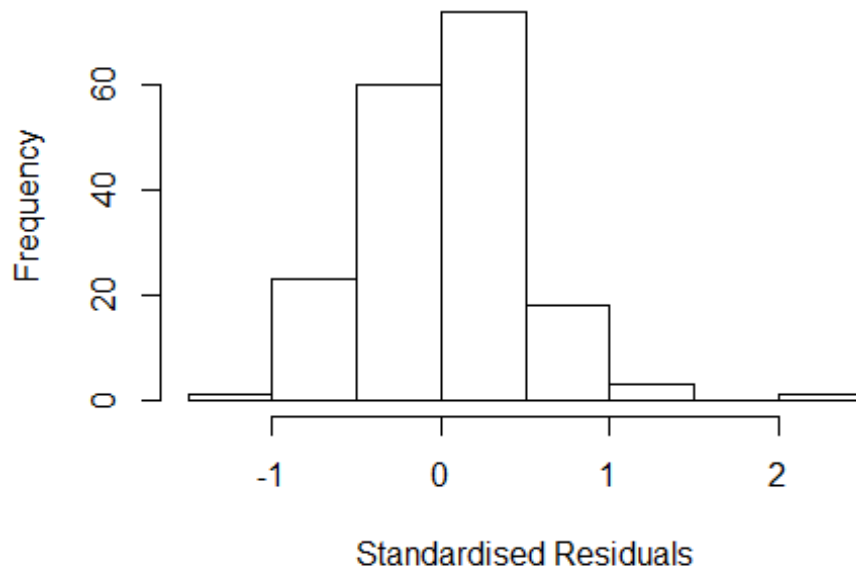
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 70, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.96694938546153"
## [1] "Male last author team size 2018 geometric mean: 3.45517105161634"

## Warning in wilcox.test.default(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 = 0.3
## 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 4.153 1      2.038
## LastAuthorFemale  2.068 1      1.438
## UniqueAuthors    20.499 4      1.459
## Year              81.533 16     1.147
```

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
## -1.220 -0.284 0.046 0.263 2.357
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8069 0.1805 4.47 1.5e-05 ***
## FirstAuthorFemale1 0.1078 0.0964 1.12 0.26517
## LastAuthorFemale1 0.0327 0.1102 0.30 0.76682
## UniqueAuthors2 0.3970 0.1128 3.52 0.00057 ***
## UniqueAuthors3 0.3738 0.1184 3.16 0.00191 **
## UniqueAuthors4 0.3651 0.1259 2.90 0.00426 **
## UniqueAuthors5 0.3900 0.1321 2.95 0.00364 **
## Year1997 0.1974 0.3150 0.63 0.53170
## Year1998 -0.2967 0.3334 -0.89 0.37481
## Year1999 -0.3523 0.2226 -1.58 0.11556
```

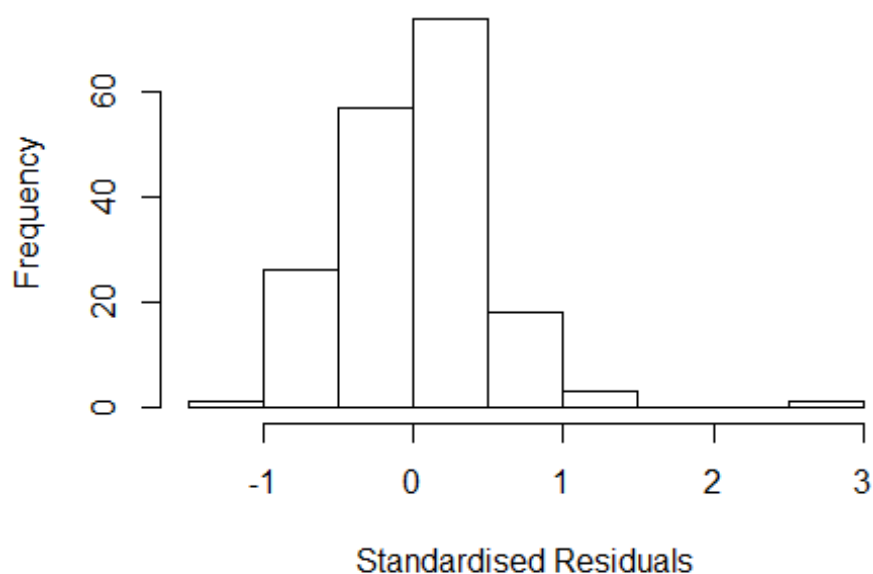


```

## Year2000          -0.0505      0.3747   -0.13  0.89305
## Year2001           0.2210      0.1886    1.17  0.24303
## Year2002          -0.2500      0.4682   -0.53  0.59418
## Year2003          -0.0185      0.2635   -0.07  0.94402
## Year2004          -0.0142      0.2944   -0.05  0.96156
## Year2005           0.2029      0.2097    0.97  0.33464
## Year2006           0.2201      0.2586    0.85  0.39598
## Year2007           0.1209      0.2610    0.46  0.64385
## Year2008           0.0405      0.2037    0.20  0.84267
## Year2009           0.1766      0.1940    0.91  0.36417
## Year2010           0.0231      0.2141    0.11  0.91430
## Year2011           0.0388      0.2005    0.19  0.84680
## Year2012          -0.0127      0.2191   -0.06  0.95372
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.416
## Multiple R-squared:  0.151, Adjusted R-squared:  0.0317
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## observation 19 is an outlier with |weight| = 0 ( < 0.00056);
## 14 weights are ~= 1. The remaining 165 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.300  0.863  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          5.56e-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 2.923 1          1.710
## LastAuthorFemale 1.819 1          1.349
## Year             5.061 16          1.052

```

Residuals from first and last author



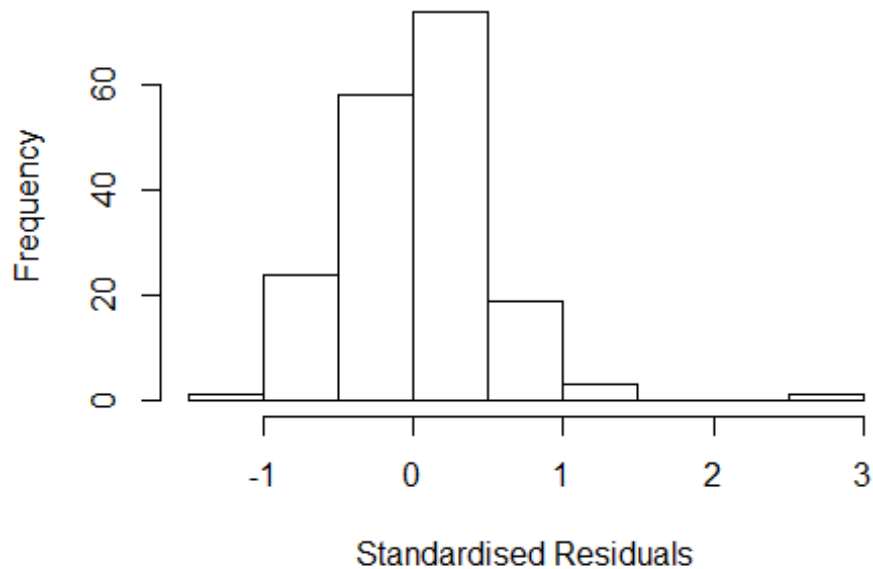
```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 62 0033207774 3.209 1999      1500      2      2.537
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1761 -0.3072  0.0466  0.2831  2.5366
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1318    0.1878   6.03 1.1e-08 ***
## FirstAuthorFemale1  0.1419    0.0891   1.59  0.113
## LastAuthorFemale1  0.0363    0.1100   0.33  0.742
## Year1997          0.1284    0.2641   0.49  0.627
## Year1998         -0.3371    0.3530  -0.95  0.341
## Year1999         -0.4594    0.2170  -2.12  0.036 *
## Year2000         -0.0706    0.4640  -0.15  0.879
## Year2001          0.2718    0.1902   1.43  0.155
## Year2002         -0.1951    0.4254  -0.46  0.647
## Year2003         -0.0246    0.2712  -0.09  0.928
## Year2004         -0.0542    0.2716  -0.20  0.842
## Year2005          0.1837    0.2372   0.77  0.440
```

```

## Year2006          0.2374      0.2905      0.82      0.415
## Year2007          0.1792      0.2571      0.70      0.487
## Year2008          0.0807      0.2112      0.38      0.703
## Year2009          0.1926      0.2056      0.94      0.350
## Year2010          0.0443      0.2258      0.20      0.845
## Year2011          0.0480      0.2145      0.22      0.823
## Year2012          0.0198      0.2185      0.09      0.928
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.451
## Multiple R-squared:  0.1,   Adjusted R-squared:  -0.000427
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## observation 19 is an outlier with |weight| = 0 ( < 0.00056);
## 12 weights are ~= 1. The remaining 167 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.392  0.878   0.955   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          5.56e-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 2.994 1           1.730
## Year              2.994 16           1.035

```

Residuals from first author

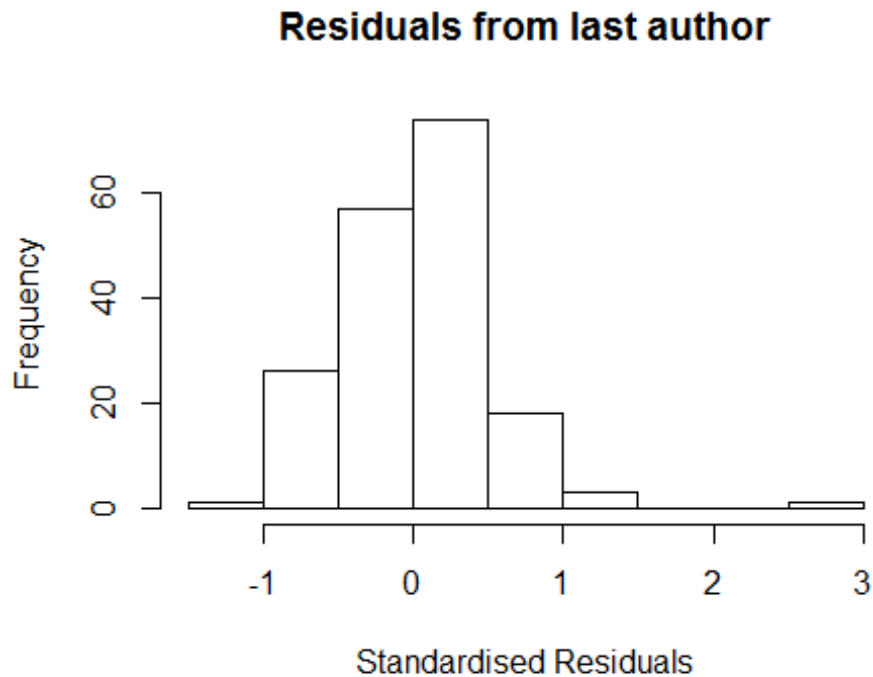


```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 62 0033207774 3.209 1999      1500      2      2.537
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.176 -0.312  0.043  0.283  2.528
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1332     0.1927   5.88 2.3e-08 ***
## FirstAuthorFemale1 0.1466     0.0896   1.64  0.104
## Year1997          0.1245     0.2681   0.46  0.643
## Year1998         -0.3421     0.3561  -0.96  0.338
## Year1999         -0.4517     0.2142  -2.11  0.037 *
## Year2000         -0.0549     0.4634  -0.12  0.906
## Year2001          0.2704     0.1951   1.39  0.168
## Year2002         -0.1956     0.4366  -0.45  0.655
## Year2003         -0.0150     0.2710  -0.06  0.956
## Year2004         -0.0508     0.2746  -0.19  0.853
## Year2005          0.1806     0.2416   0.75  0.456
## Year2006          0.2464     0.2876   0.86  0.393
```

```

## Year2007          0.1786      0.2579      0.69      0.489
## Year2008          0.0808      0.2155      0.38      0.708
## Year2009          0.1925      0.2096      0.92      0.360
## Year2010          0.0430      0.2295      0.19      0.852
## Year2011          0.0533      0.2145      0.25      0.804
## Year2012          0.0253      0.2188      0.12      0.908
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.449
## Multiple R-squared:  0.0998, Adjusted R-squared:  0.00537
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## observation 19 is an outlier with |weight| = 0 ( < 0.00056);
## 13 weights are ~= 1. The remaining 166 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.396  0.877   0.953   0.910   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      5.56e-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 1.934 1      1.391
## Year              1.934 16      1.021

```



```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 62 0033207774 3.209 1999    1500      2      2.537
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2202 -0.3419  0.0621  0.2909  2.4527
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1446     0.1717   6.67 3.9e-10 ***
## LastAuthorFemale1  0.0665     0.1117   0.60  0.553
## Year1997          0.1957     0.2593   0.75  0.451
## Year1998         -0.2623     0.3517  -0.75  0.457
## Year1999         -0.3884     0.2057  -1.89  0.061 .
## Year2000          0.0435     0.4412   0.10  0.922
## Year2001          0.2590     0.1744   1.49  0.139
## Year2002         -0.2043     0.4460  -0.46  0.647
## Year2003         -0.0215     0.2553  -0.08  0.933
## Year2004         -0.0637     0.2549  -0.25  0.803
## Year2005          0.2201     0.2320   0.95  0.344
## Year2006          0.2160     0.2811   0.77  0.443
```

```

## Year2007          0.2248      0.2523      0.89      0.374
## Year2008          0.1292      0.1925      0.67      0.503
## Year2009          0.2462      0.1827      1.35      0.180
## Year2010          0.0756      0.2102      0.36      0.720
## Year2011          0.0649      0.2001      0.32      0.746
## Year2012          0.0373      0.2034      0.18      0.855
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.453
## Multiple R-squared:  0.084, Adjusted R-squared:  -0.0121
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## observation 19 is an outlier with |weight| = 0 ( < 0.00056);
## 11 weights are ~= 1. The remaining 168 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.315  0.883   0.950   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      5.56e-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: 180"
## [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]"
##
## 2005 2007 2008 2010 2012
##    1    2    2    1    2
##
## 2005 2007 2008 2010 2012
##    1    1    2    0    2
##
## 2005 2007 2008 2010 2012

```

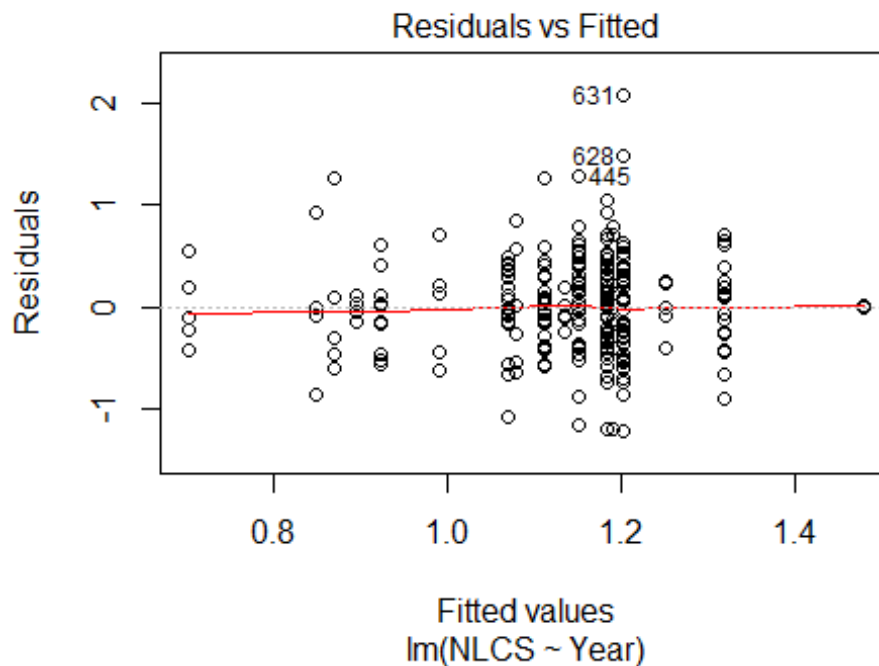
```

##      1      1      2      0      0
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4"
## [1] "Male first author team size 2018 geometric mean: 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: 4"
## [1] "Male last 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] "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 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
##    15    14    16     9    18    14    13    13    14    13    27    35    35    40    53
## 2011 2012
##    63    66
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     5     4     6     2     8     5     5     5     6     5    12    20    23    27    38
## 2011 2012
##    47    46
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     4     4     4     0     7     4     5     5     5     4    10    18    19    22    27
## 2011 2012
##    33    37

```



```
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 38, df = 16, p-value = 0.002
```



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

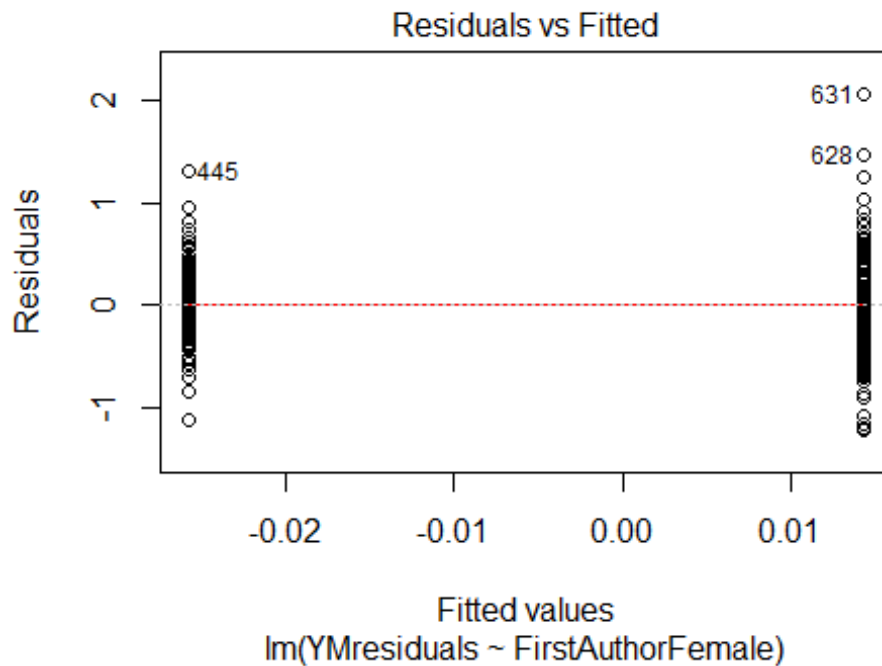
## [1] "Female first author team size 2018 geometric mean: 3.75116612261711"
## [1] "Male first author team size 2018 geometric mean: 4.06399843346462"

## Warning in wilcox.test.default(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 = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
```

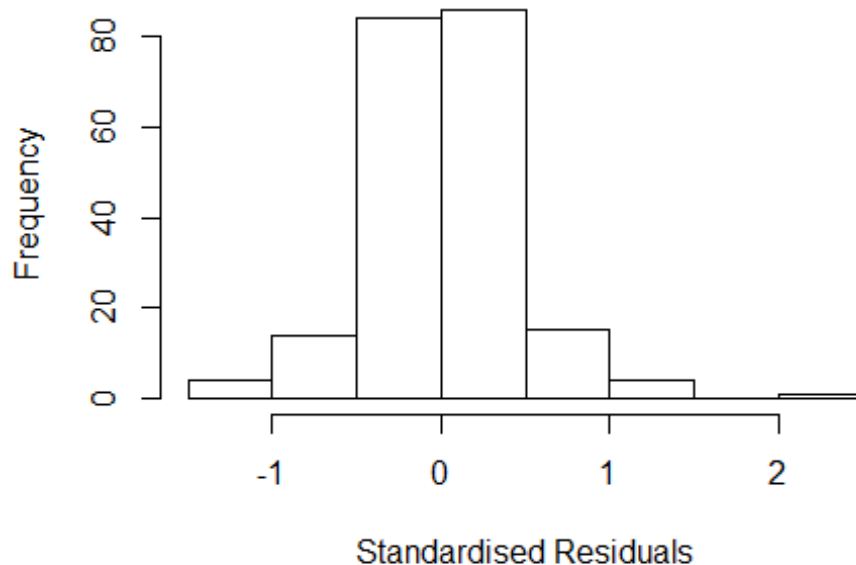
```
## [1] "Female last author team size 2018 geometric mean: 11"
## [1] "Male last author team size 2018 geometric mean: 3.64184276171929"

## Warning in wilcox.test.default(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 = 0.1
## 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 1.530  1      1.237
## LastAuthorFemale  1.218  1      1.104
## UniqueAuthors    3.430  4      1.167
## Year              4.703 15      1.053
```

Residuals from first and last author and team size



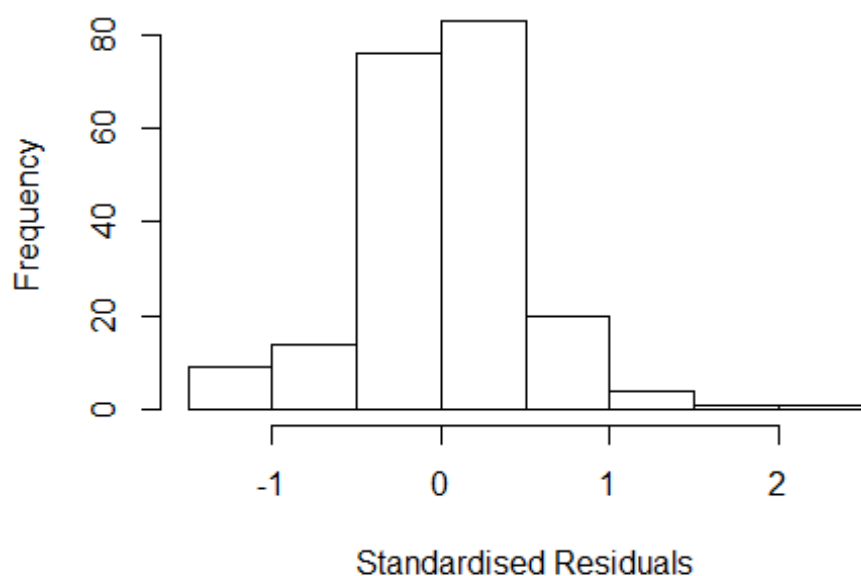
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1562 -0.2724 0.0125 0.2557 2.0722
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.1574 0.3275 0.48 0.632
## FirstAuthorFemale1 -0.1103 0.0664 -1.66 0.098 .
## LastAuthorFemale1 -0.0272 0.1039 -0.26 0.794
## UniqueAuthors2 0.5889 0.1473 4.00 9.2e-05 ***
## UniqueAuthors3 0.7101 0.1484 4.79 3.5e-06 ***
## UniqueAuthors4 0.7539 0.1486 5.07 9.4e-07 ***
## UniqueAuthors5 0.9104 0.1370 6.64 3.3e-10 ***
## Year1997 0.3213 0.3594 0.89 0.372
## Year1998 0.3609 0.3504 1.03 0.304
## Year2000 0.4472 0.3487 1.28 0.201
```

```

## Year2001          0.0296      0.3610      0.08      0.935
## Year2002          0.2448      0.3798      0.64      0.520
## Year2003          0.4640      0.3090      1.50      0.135
## Year2004          0.3455      0.3204      1.08      0.282
## Year2005         -0.0577      0.2897     -0.20      0.842
## Year2006          0.0859      0.3024      0.28      0.777
## Year2007          0.2783      0.3027      0.92      0.359
## Year2008          0.4174      0.3090      1.35      0.178
## Year2009          0.2600      0.2947      0.88      0.379
## Year2010          0.1988      0.3147      0.63      0.528
## Year2011          0.2458      0.3000      0.82      0.414
## Year2012          0.2936      0.3006      0.98      0.330
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.415
## Multiple R-squared:  0.271, Adjusted R-squared:  0.189
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## observation 208 is an outlier with |weight| = 0 ( < 0.00048);
## 18 weights are ~= 1. The remaining 189 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.177  0.895   0.959   0.908   0.985   0.999
## Algorithmic parameters:
##           tuning.chi          bb          tuning.psi          refine.tol
##           1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol          solve.tol          eps.outlier          eps.x
##           1.00e-07          1.00e-07          4.81e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01          5.00e-01
## nResample    max.it    best.r.s    k.fast.s    k.max maxit.scale
##           500         50         2         1         1000         200
## trace.lev    mts    compute.rd
##           0         1000         0
##           psi          subsampling          cov
##           "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.797 1          1.341
## LastAuthorFemale  1.319 1          1.149
## Year              2.343 15          1.029

```

Residuals from first and last author



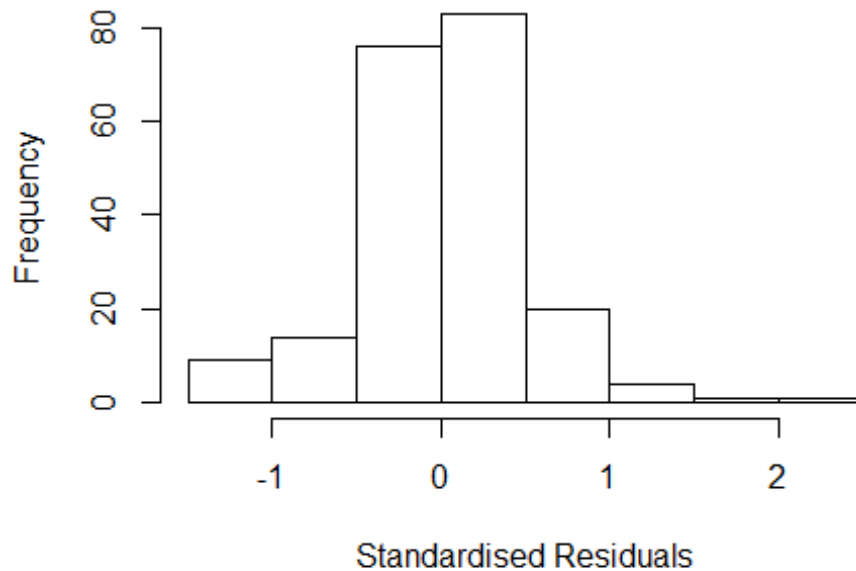
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3066 -0.3081 0.0206 0.2800 2.1193
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.878919 0.359534 2.44 0.015 *
## FirstAuthorFemale1 -0.079891 0.076462 -1.04 0.297
## LastAuthorFemale1 0.000548 0.122682 0.00 0.996
## Year1997 -0.068758 0.532567 -0.13 0.897
## Year1998 0.072594 0.443968 0.16 0.870
## Year2000 0.427712 0.442133 0.97 0.335
## Year2001 -0.062410 0.395027 -0.16 0.875
## Year2002 0.145939 0.425255 0.34 0.732
## Year2003 0.421465 0.367548 1.15 0.253
## Year2004 0.265752 0.366535 0.73 0.469
## Year2005 0.026438 0.350336 0.08 0.940
## Year2006 0.010995 0.368287 0.03 0.976
```

```

## Year2007          0.256696    0.365576    0.70    0.483
## Year2008          0.447570    0.362600    1.23    0.219
## Year2009          0.242986    0.357507    0.68    0.498
## Year2010          0.230517    0.385397    0.60    0.550
## Year2011          0.269615    0.354150    0.76    0.447
## Year2012          0.278825    0.365683    0.76    0.447
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.438
## Multiple R-squared:  0.0752, Adjusted R-squared:  -0.00752
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## observation 208 is an outlier with |weight| = 0 ( < 0.00048);
## 16 weights are ~= 1. The remaining 191 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.199  0.886  0.954  0.898  0.988  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      4.81e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.746 1      1.321
## Year              1.746 15      1.019

```

Residuals from first author



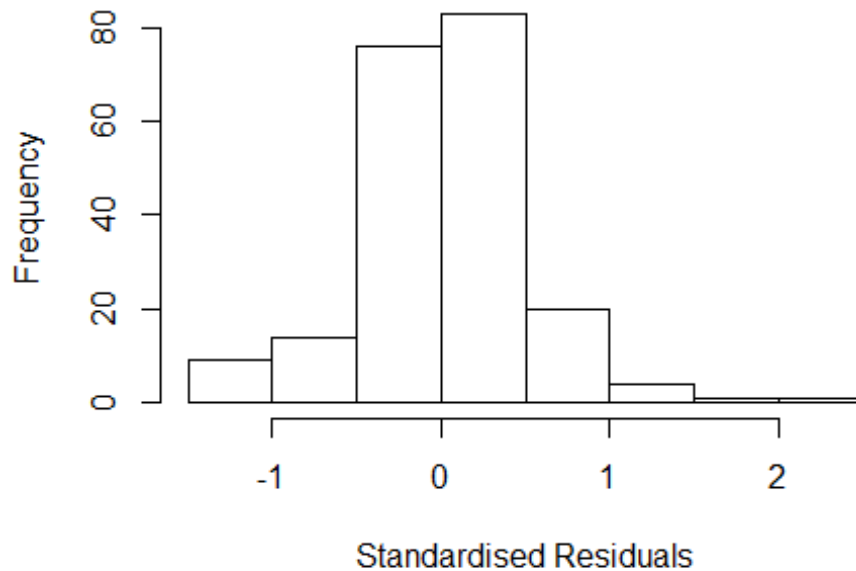
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3014 -0.3122 0.0216 0.2804 2.1197
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.89156 0.36235 2.46 0.015 *
## FirstAuthorFemale1 -0.08014 0.07557 -1.06 0.290
## Year1997 -0.07857 0.53149 -0.15 0.883
## Year1998 0.06140 0.44630 0.14 0.891
## Year2000 0.40980 0.45530 0.90 0.369
## Year2001 -0.07447 0.39752 -0.19 0.852
## Year2002 0.13350 0.42488 0.31 0.754
## Year2003 0.40891 0.37023 1.10 0.271
## Year2004 0.25331 0.36954 0.69 0.494
## Year2005 0.01392 0.35380 0.04 0.969
## Year2006 -0.00134 0.37200 0.00 0.997
## Year2007 0.24213 0.36836 0.66 0.512
```

```

## Year2008          0.43472    0.36579    1.19    0.236
## Year2009          0.23200    0.36102    0.64    0.521
## Year2010          0.21449    0.38602    0.56    0.579
## Year2011          0.25792    0.35814    0.72    0.472
## Year2012          0.26572    0.36736    0.72    0.470
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.449
## Multiple R-squared:  0.0732, Adjusted R-squared:  -0.00446
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## observation 208 is an outlier with |weight| = 0 ( < 0.00048);
## 17 weights are ~= 1. The remaining 190 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.223  0.890  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      4.81e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.361 1      1.167
## Year              1.361 15      1.010

```


Residuals from last author



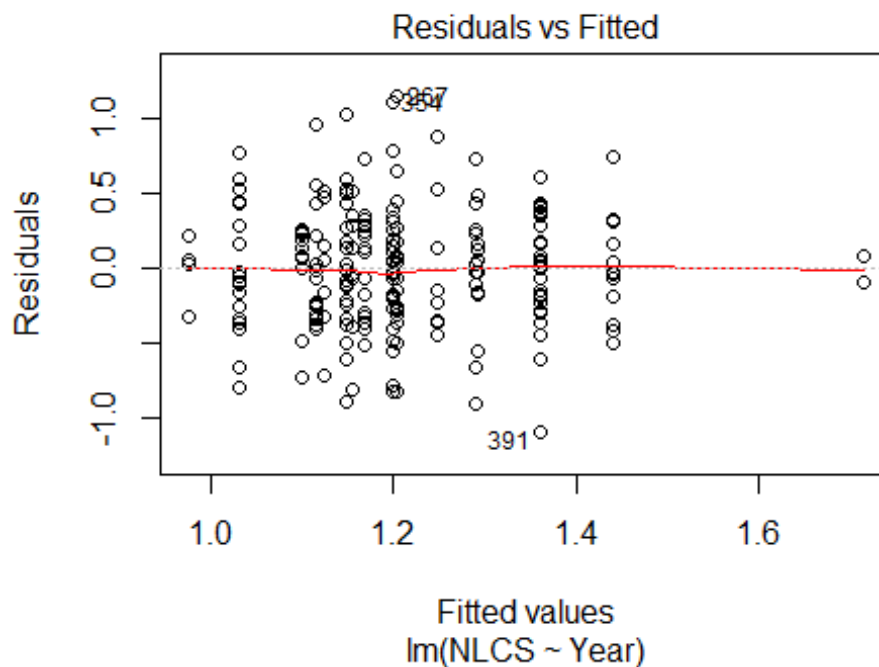
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2334 -0.2969 0.0368 0.2816 2.1518
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.82162 0.33337 2.46 0.015 *
## LastAuthorFemale1 0.00577 0.12445 0.05 0.963
## Year1997 -0.00933 0.51296 -0.02 0.986
## Year1998 0.08047 0.44268 0.18 0.856
## Year2000 0.41182 0.43080 0.96 0.340
## Year2001 -0.02067 0.36588 -0.06 0.955
## Year2002 0.16261 0.40677 0.40 0.690
## Year2003 0.43321 0.35210 1.23 0.220
## Year2004 0.30653 0.34378 0.89 0.374
## Year2005 0.04397 0.33534 0.13 0.896
## Year2006 0.03303 0.35463 0.09 0.926
## Year2007 0.28126 0.34940 0.81 0.422
```

```

## Year2008          0.46539      0.34832      1.34      0.183
## Year2009          0.27595      0.33907      0.81      0.417
## Year2010          0.27408      0.36189      0.76      0.450
## Year2011          0.28716      0.34087      0.84      0.401
## Year2012          0.30354      0.35020      0.87      0.387
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.446
## Multiple R-squared:  0.0675, Adjusted R-squared:  -0.0106
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## observation 208 is an outlier with |weight| = 0 ( < 0.00048);
## 13 weights are ~= 1. The remaining 194 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.196  0.896  0.953   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      4.81e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  208"
## [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
##   12   8  16  13  10   8  10  18  14  20   9  12  20  23  23
## 2011 2012
##   34  33
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   7   5  12   9   2   4   8  10  11  14   8   9  18  18  21
## 2011 2012

```

```
## 28 30
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 7 5 11 8 2 2 6 9 8 12 6 7 14 15 19
## 2011 2012
## 23 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 = 9, df = 16, p-value = 0.9
```



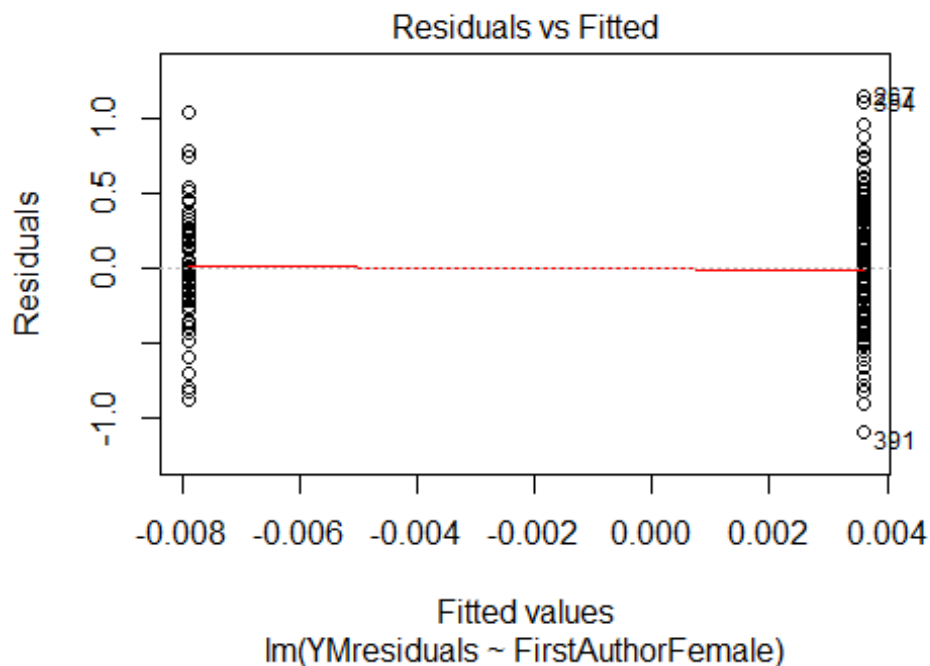
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.083, df = 1, p-value = 0.8

## [1] "Female first author team size 2018 geometric mean: 4.48140474655716"
## [1] "Male first author team size 2018 geometric mean: 4.16593874788318"

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

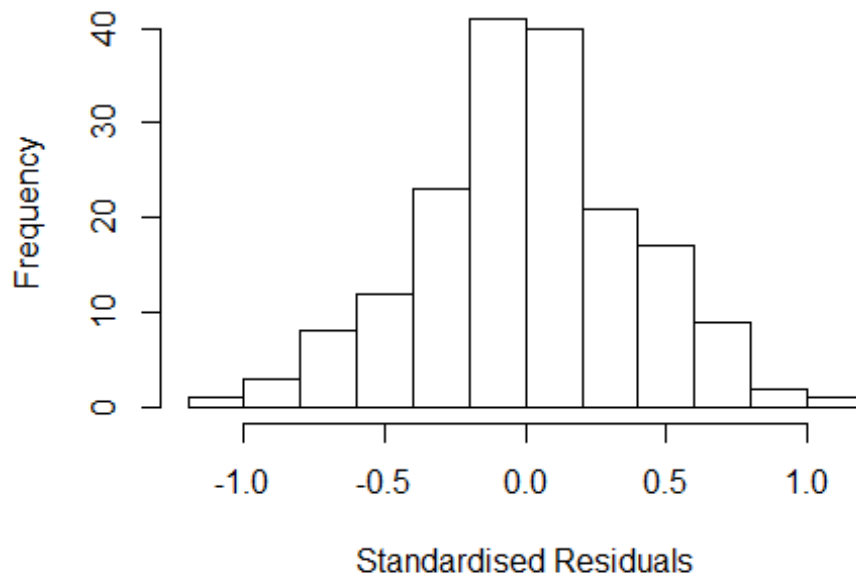
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 14, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.01329793496458"
## [1] "Male last author team size 2018 geometric mean: 3.9943569808963"

## Warning in wilcox.test.default(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 1.447 1      1.203
## LastAuthorFemale  1.616 1      1.271
## UniqueAuthors    7.010 4      1.276
## Year             12.697 16     1.083
```

Residuals from first and last author and team size



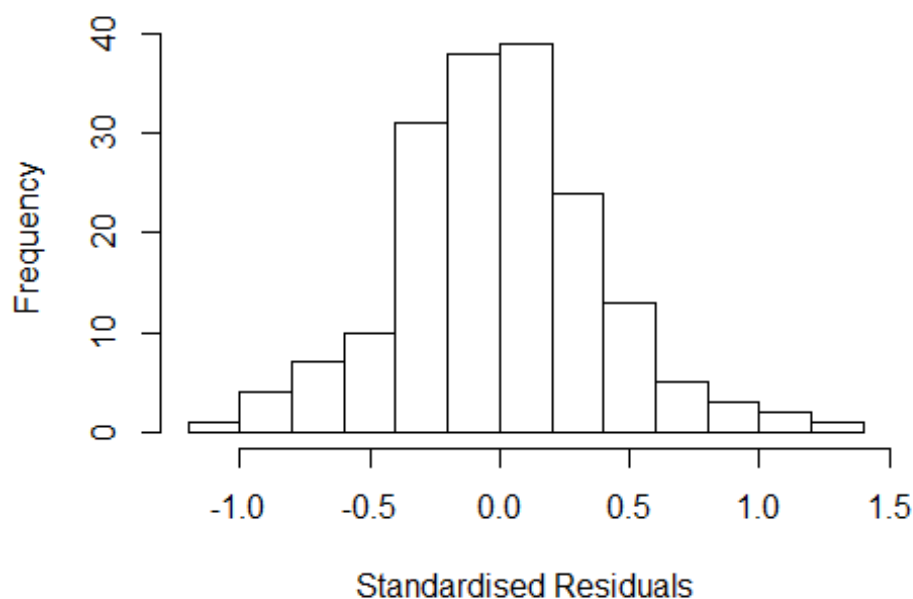
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.05924 -0.22022  0.00694  0.22732  1.08226
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.488288   0.345218   4.31 2.9e-05 ***
## FirstAuthorFemale1 -0.034384   0.068048  -0.51  0.6141
## LastAuthorFemale1 -0.081270   0.083417  -0.97  0.3314
## UniqueAuthors2    -0.332828   0.293073  -1.14  0.2579
## UniqueAuthors3    -0.292937   0.294830  -0.99  0.3220
## UniqueAuthors4    -0.364496   0.301284  -1.21  0.2282
## UniqueAuthors5    -0.014151   0.317161  -0.04  0.9645
## Year1997           0.039811   0.257674   0.15  0.8774
## Year1998           0.221626   0.195219   1.14  0.2580
## Year1999           0.254659   0.241510   1.05  0.2933
```

```

## Year2000          0.592067    0.197201    3.00    0.0031 **
## Year2001          -0.391052    0.467790   -0.84    0.4045
## Year2002          -0.075958    0.342222   -0.22    0.8246
## Year2003          -0.064902    0.220908   -0.29    0.7693
## Year2004           0.027299    0.224678    0.12    0.9035
## Year2005          -0.151715    0.208343   -0.73    0.4676
## Year2006           0.097440    0.241083    0.40    0.6866
## Year2007           0.189950    0.204518    0.93    0.3545
## Year2008          -0.295132    0.231630   -1.27    0.2045
## Year2009          -0.108260    0.227555   -0.48    0.6349
## Year2010          -0.000067    0.218410    0.00    0.9998
## Year2011           0.005602    0.202908    0.03    0.9780
## Year2012           0.173777    0.203991    0.85    0.3956
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.356
## Multiple R-squared:  0.21,   Adjusted R-squared:  0.0975
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 170 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.336  0.869  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           5.62e-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 1.462 1         1.209
## LastAuthorFemale  1.613 1         1.270
## Year              2.026 16         1.022

```

Residuals from first and last author



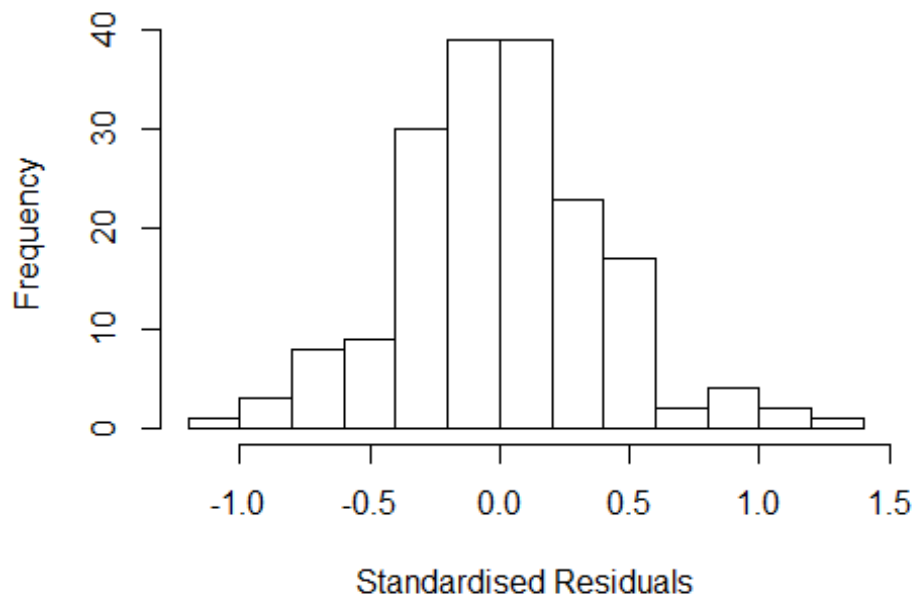
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1498 -0.2344 -0.0132  0.2398  1.2333
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.16464    0.17713     6.58 6.6e-10 ***
## FirstAuthorFemale1 -0.03638    0.06929    -0.53  0.6003
## LastAuthorFemale1 -0.10779    0.07930    -1.36  0.1760
## Year1997          0.04263    0.24203     0.18  0.8604
## Year1998          0.24848    0.19424     1.28  0.2027
## Year1999          0.26476    0.22352     1.18  0.2380
## Year2000          0.56805    0.19112     2.97  0.0034 **
## Year2001         -0.21995    0.29730    -0.74  0.4605
## Year2002          0.04437    0.31282     0.14  0.8874
## Year2003         -0.00729    0.21128    -0.03  0.9725
## Year2004          0.03288    0.22403     0.15  0.8835
## Year2005         -0.06052    0.22039    -0.27  0.7840
```

```

## Year2006          0.10214      0.23678      0.43      0.6668
## Year2007          0.23861      0.19540      1.22      0.2238
## Year2008         -0.18206      0.22261     -0.82      0.4147
## Year2009         -0.04095      0.22027     -0.19      0.8527
## Year2010          0.03921      0.20926      0.19      0.8516
## Year2011          0.03407      0.20093      0.17      0.8656
## Year2012          0.25520      0.18924      1.35      0.1794
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.365
## Multiple R-squared:  0.147, Adjusted R-squared:  0.0506
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 159 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.230  0.872  0.945  0.890  0.984  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      5.62e-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 1.354 1          1.163
## Year              1.354 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
## -1.13105 -0.22357 -0.00403 0.24358 1.24155
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1426 0.1725 6.63 5e-10 ***
## FirstAuthorFemale1 -0.0246 0.0684 -0.36 0.7195
## Year1997 0.0469 0.2530 0.19 0.8530
## Year1998 0.2376 0.1966 1.21 0.2286
## Year1999 0.2849 0.2212 1.29 0.1997
## Year2000 0.5842 0.1885 3.10 0.0023 **
## Year2001 -0.2038 0.2900 -0.70 0.4832
## Year2002 0.0621 0.3060 0.20 0.8394
## Year2003 -0.0052 0.2110 -0.02 0.9804
## Year2004 0.0314 0.2230 0.14 0.8882
## Year2005 -0.0537 0.2170 -0.25 0.8048
## Year2006 0.0941 0.2438 0.39 0.7001
```

```

## Year2007          0.2570      0.1911      1.35      0.1805
## Year2008         -0.1792      0.2213     -0.81      0.4195
## Year2009         -0.0272      0.2113     -0.13      0.8978
## Year2010          0.0242      0.2074      0.12      0.9075
## Year2011          0.0160      0.1970      0.08      0.9355
## Year2012          0.2584      0.1883      1.37      0.1719
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.365
## Multiple R-squared:  0.137, Adjusted R-squared:  0.0454
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 160 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.224  0.868  0.948  0.890  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      5.62e-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 1.492  1          1.222
## Year              1.492 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
## -1.1409 -0.2325 -0.0152  0.2340  1.2477

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1632    0.1810   6.43 1.4e-09 ***
## LastAuthorFemale1 -0.1016    0.0774  -1.31  0.1910
## Year1997          0.0351    0.2408   0.15  0.8843
## Year1998          0.2380    0.1930   1.23  0.2192
## Year1999          0.2608    0.2267   1.15  0.2517
## Year2000          0.5513    0.1920   2.87  0.0046 **
## Year2001         -0.2367    0.2799  -0.85  0.3990
## Year2002          0.0310    0.3069   0.10  0.9197
## Year2003         -0.0210    0.2069  -0.10  0.9192
## Year2004          0.0125    0.2199   0.06  0.9549
## Year2005         -0.0699    0.2195  -0.32  0.7506
## Year2006          0.0883    0.2392   0.37  0.7125
## Year2007          0.2290    0.1946   1.18  0.2411
## Year2008         -0.2009    0.2194  -0.92  0.3613
## Year2009         -0.0539    0.2201  -0.24  0.8070
## Year2010          0.0314    0.2120   0.15  0.8823
## Year2011          0.0207    0.1974   0.10  0.9166
## Year2012          0.2478    0.1901   1.30  0.1943
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.363
## Multiple R-squared:  0.146, Adjusted R-squared:  0.0553
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 161 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.214  0.872  0.950  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      5.62e-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: 178"
## [1] ""

```

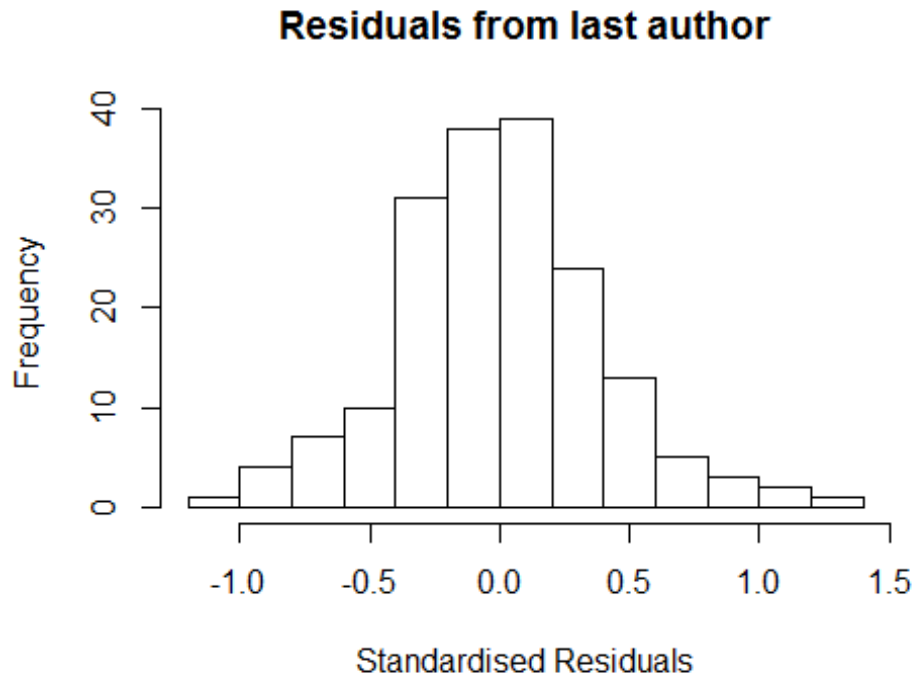
```

## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1504"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 2000 2006 2007 2008 2012
##    1    1    1    1    1
##
## 2000 2006 2007 2008 2012
##    0    0    1    1    1
##
## 2000 2006 2007 2008 2012
##    0    0    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: 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 1505"
## [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
##    3    7    5    6    4    6    6    2    4    6    7    5    10    10    16
## 2012
##    9
##
## 1996 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    0    0    3    2    1    3    4    2    3    5    3    4    7    7    7
## 2012
##    7
##
## 1996 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    0    0    3    2    1    1    4    2    3    5    3    2    6    7    6
## 2012
##    6
## [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: 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] "Female last author team size 2018 geometric mean: 3"
## [1] "Male last author team size 2018 geometric mean: 5.47722557505166"
##
## 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] "Regression 1: First author gender, last author gender, team size,
Year as factors"

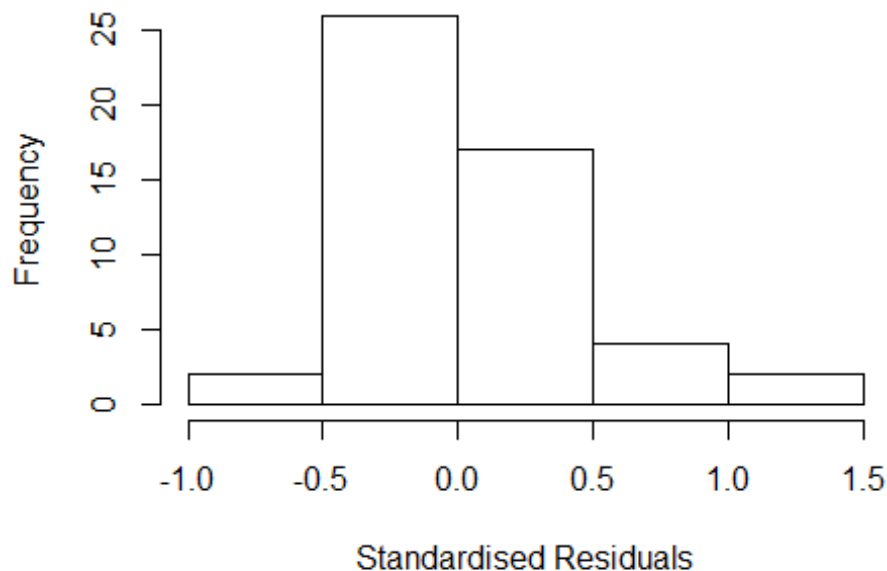
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead
```



```
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##
## GVIF Df GVIF^(1/(2*Df))
```

```
## FirstAuthorFemale 10.450 1 3.233
## LastAuthorFemale 3.358 1 1.833
## Year 76.191 13 1.181
```

Residuals from first and last author



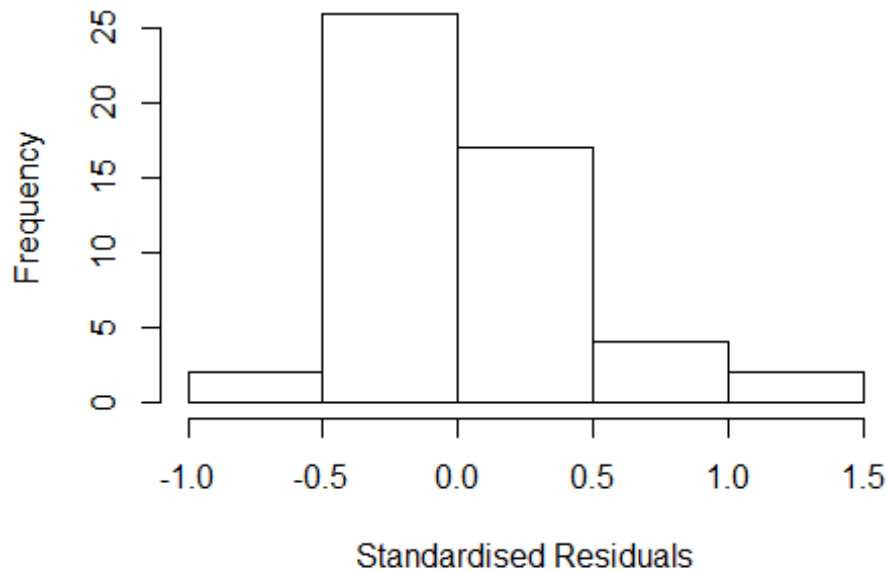
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.6181 -0.1567 -0.0217 0.2110 1.4063
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3048 0.1227 10.64 1.6e-12 ***
## FirstAuthorFemale1 0.1923 0.1952 0.99 0.33118
## LastAuthorFemale1 0.0405 0.2163 0.19 0.85264
## Year2000 -0.6788 0.1797 -3.78 0.00059 ***
## Year2001 -0.1638 0.1227 -1.34 0.19024
## Year2002 -0.0788 0.1227 -0.64 0.52457
## Year2003 -0.1665 0.1741 -0.96 0.34561
## Year2004 0.1188 0.2904 0.41 0.68488
```

```

## Year2005          -0.1030      0.2576   -0.40  0.69184
## Year2006          -0.1765      0.2280   -0.77  0.44396
## Year2007          -0.4068      0.3375   -1.21  0.23609
## Year2008          -0.3275      0.2770   -1.18  0.24500
## Year2009          -0.3541      0.2251   -1.57  0.12473
## Year2010           0.0327      0.1697    0.19  0.84847
## Year2011          -0.3586      0.4177   -0.86  0.39644
## Year2012           0.2043      0.1893    1.08  0.28775
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.372
## Multiple R-squared:  0.325, Adjusted R-squared:  0.0358
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 43 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.123  0.928  0.959   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      1.96e-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"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 24.45 1      4.944
## Year              24.45 13      1.131

```

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
## -0.6196 -0.1595 -0.0229 0.2098 1.4281
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3176 0.1069 12.32 1.8e-14 ***
## FirstAuthorFemale1 0.2026 0.1799 1.13 0.26754
## Year2000 -0.6916 0.1699 -4.07 0.00024 ***
## Year2001 -0.1766 0.1069 -1.65 0.10723
## Year2002 -0.0916 0.1069 -0.86 0.39714
## Year2003 -0.1869 0.1813 -1.03 0.30940
## Year2004 0.0958 0.2968 0.32 0.74869
## Year2005 -0.1217 0.2666 -0.46 0.65090
## Year2006 -0.1983 0.2376 -0.83 0.40932
## Year2007 -0.4245 0.3481 -1.22 0.23059
## Year2008 -0.3454 0.2946 -1.17 0.24874
## Year2009 -0.3888 0.2164 -1.80 0.08075 .
```

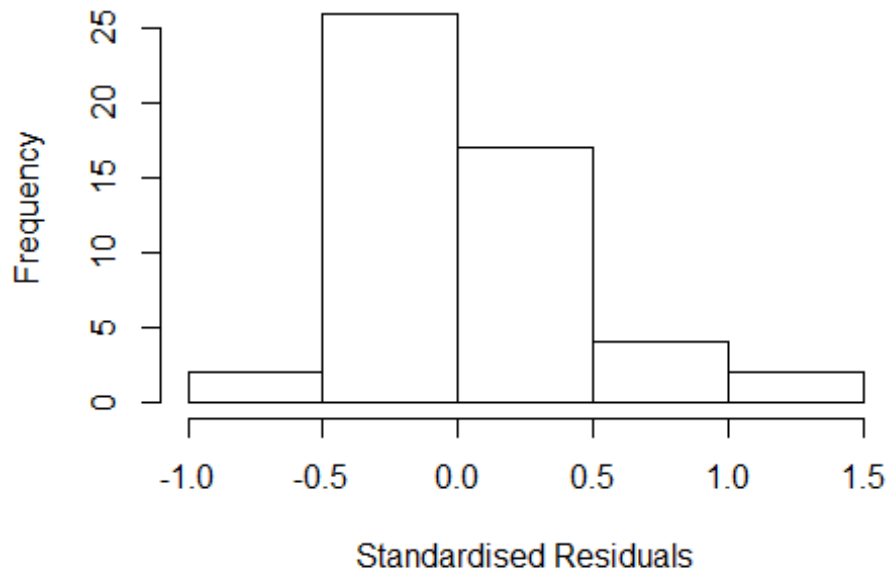


```

## Year2010          0.0308      0.1574      0.20  0.84623
## Year2011         -0.3877      0.3948     -0.98  0.33258
## Year2012          0.1929      0.1792      1.08  0.28893
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.355
## Multiple R-squared:  0.349, Adjusted R-squared:  0.0958
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 45 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0681 0.9220 0.9580 0.8960 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      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"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 4.626 1          2.151
## Year            4.626 13          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
## -0.7161 -0.1765 -0.0591 0.2070 1.2729
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.317698 0.117074 11.26 2.4e-13 ***
## LastAuthorFemale1 -0.000978 0.226447 0.00 0.99658
## Year2000 -0.691698 0.176040 -3.93 0.00037 ***
## Year2001 -0.176698 0.117074 -1.51 0.13995
## Year2002 -0.091698 0.117074 -0.78 0.43860
## Year2003 -0.035190 0.121675 -0.29 0.77408
## Year2004 0.298302 0.252153 1.18 0.24456
## Year2005 -0.045095 0.211432 -0.21 0.83231
## Year2006 -0.063373 0.155212 -0.41 0.68547
## Year2007 -0.289393 0.413124 -0.70 0.48812
## Year2008 -0.244198 0.384026 -0.64 0.52887
## Year2009 -0.233638 0.209202 -1.12 0.27148
```

```

## Year2010          0.036390    0.172235    0.21  0.83386
## Year2011         -0.228608    0.320418   -0.71  0.48016
## Year2012          0.191610    0.185489    1.03  0.30849
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.37
## Multiple R-squared:  0.268, Adjusted R-squared:  -0.0169
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 47 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.212  0.896  0.966  0.907  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      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 1506"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1999 2000 2002 2004 2005 2006 2007 2008 2009 2010 2012
##    1    1    2    1    2    7    5    1    4    3    1
##
## 1999 2000 2002 2004 2005 2006 2007 2008 2009 2010 2012
##    1    1    0    1    0    6    2    1    3    3    1
##
## 1999 2000 2002 2004 2005 2006 2007 2008 2009 2010 2012
##    1    1    0    1    0    6    2    1    3    3    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: 19"
## [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 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    4    2    4    3    4    3    8    4   11    5    7    6    7    4
##
## 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    0    1    0    1    2    0    4    1    6    3    6    4    6    4
##
## 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    0    1    0    1    2    0    3    1    5    3    5    4    6    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: 3.46410161513775"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.078e+14  1      1.038e+07
## LastAuthorFemale  5.401e+02  1      2.324e+01
## UniqueAuthors    1.652e+19  4      2.525e+02
## Year              1.054e+33 11      3.170e+01

## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 60 40649105816 1.551 2008    1507      5      2.895
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##      Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##      k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -5.41e-01 -1.36e-01 -3.33e-16  1.04e-01  2.90e+00

```

```

##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.96e+00   1.57e-01   1.25e+01  2.7e-10 ***
## FirstAuthorFemale1 -1.55e+00   2.02e-01  -7.69e+00  4.3e-07 ***
## LastAuthorFemale1  1.74e-02   1.69e-01   1.00e-01   0.9194
## UniqueAuthors2    -4.52e-01   1.57e-01  -2.88e+00   0.0100 *
## UniqueAuthors3    -1.29e+00   1.99e-01  -6.50e+00  4.2e-06 ***
## UniqueAuthors4    -1.48e+00   2.33e-01  -6.35e+00  5.5e-06 ***
## UniqueAuthors5    -9.05e-01   2.67e-01  -3.38e+00   0.0033 **
## Year2000           7.88e-01   2.17e-08   3.64e+07  < 2e-16 ***
## Year2002           1.16e+00   2.02e-01   5.75e+00  1.9e-05 ***
## Year2003          -5.59e-01   1.80e-01  -3.10e+00   0.0062 **
## Year2005          -4.39e-01   2.66e-01  -1.65e+00   0.1161
## Year2006           1.18e+00   1.98e-01   5.95e+00  1.3e-05 ***
## Year2007           2.87e-01   1.08e-01   2.66e+00   0.0159 *
## Year2008          -4.59e-01   8.53e-02  -5.38e+00  4.1e-05 ***
## Year2009           7.52e-01   2.95e-01   2.54e+00   0.0203 *
## Year2010           5.27e-01   1.69e-01   3.13e+00   0.0058 **
## Year2011           6.59e-01   3.13e-01   2.10e+00   0.0500 *
## Year2012           2.13e-01   1.77e-01   1.21e+00   0.2435
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.407
## Multiple R-squared:  0.706, Adjusted R-squared:  0.428
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## observation 15 is an outlier with |weight| = 0 ( < 0.0028);
## 9 weights are ~ = 1. The remaining 26 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.659  0.922  0.984   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           2.78e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01             5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500           50           2           1           1000           200
## trace.lev      mts      compute.rd
##           0           1000           0
##           psi             subsampling             cov
##           "bisquare"             "nonsingular"             ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)

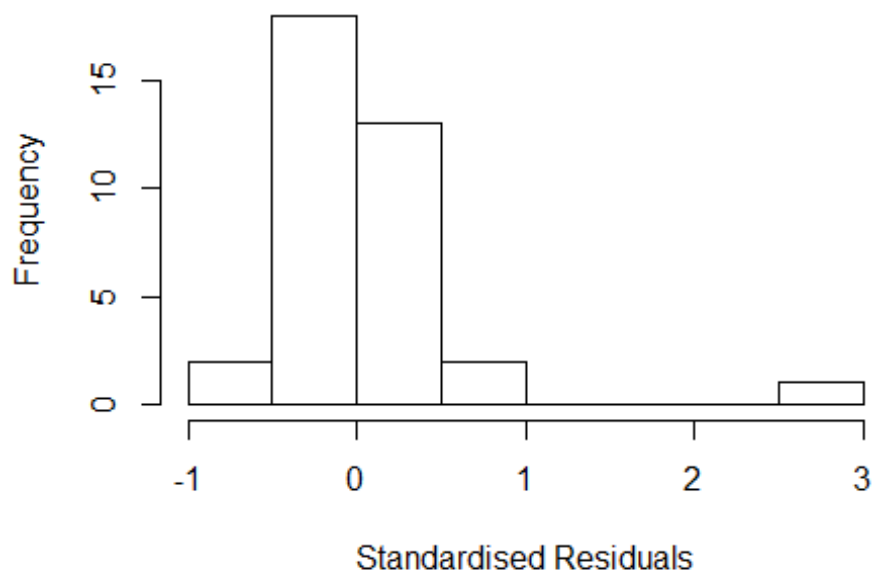
```

```
## [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 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 + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.7936 -0.3300  0.0133  0.2735  0.7686
##
```

```

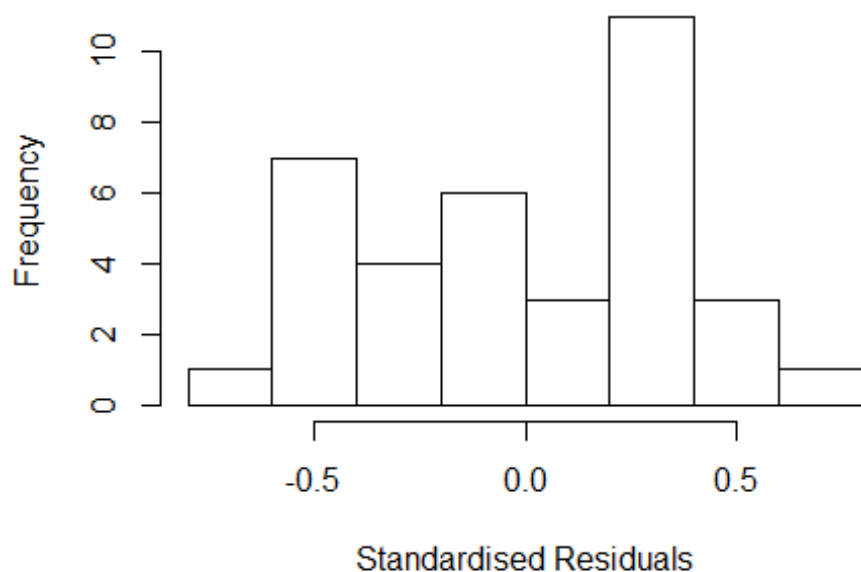
## Coefficients:
##               Estimate Std. Error   t value Pr(>|t|)
## (Intercept)    1.51e+00   2.01e-08  7.49e+07  <2e-16 ***
## FirstAuthorFemale1 -4.00e-01  1.89e-01 -2.12e+00  0.0457 *
## LastAuthorFemale1 -8.08e-02  1.90e-01 -4.30e-01  0.6745
## Year2000         7.88e-01  7.88e-16  1.00e+15  <2e-16 ***
## Year2002         9.35e-03  1.89e-01  5.00e-02  0.9610
## Year2003        -6.34e-01  3.30e-01 -1.92e+00  0.0673 .
## Year2005        -2.76e-01  2.63e-01 -1.05e+00  0.3060
## Year2006         3.37e-01  0.00e+00      Inf  <2e-16 ***
## Year2007         2.03e-01  1.79e-01  1.13e+00  0.2686
## Year2008        -1.07e-02  2.09e-01 -5.00e-02  0.9596
## Year2009         3.28e-02  2.25e-01  1.50e-01  0.8856
## Year2010        -4.57e-01  1.49e-01 -3.07e+00  0.0056 **
## Year2011        -2.58e-02  2.19e-01 -1.20e-01  0.9073
## Year2012        -5.16e-01  2.09e-01 -2.47e+00  0.0217 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.521
## Multiple R-squared:  0.485, Adjusted R-squared:  0.181
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 29 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.800  0.932  0.964  0.944  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           2.78e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01           5.00e-01
##   nResample    max.it    best.r.s    k.fast.s    k.max maxit.scale
##           500         50         2         1         1000         200
##   trace.lev    mts    compute.rd
##           0         1000         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 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
## -0.8320 -0.2896  0.0248  0.2933  0.7367
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.50800    0.00000      Inf < 2e-16 ***
## FirstAuthorFemale1 -0.42584    0.17709    -2.40  0.02463 *
## Year2000         0.78800    0.00000      Inf < 2e-16 ***
## Year2002         0.03484    0.17709     0.20  0.84575
## Year2003        -0.66208    0.34854    -1.90  0.07010 .
## Year2005        -0.27625    0.26199    -1.05  0.30263
## Year2006         0.33700    0.00000      Inf < 2e-16 ***
## Year2007         0.16105    0.20351     0.79  0.43684
```



```

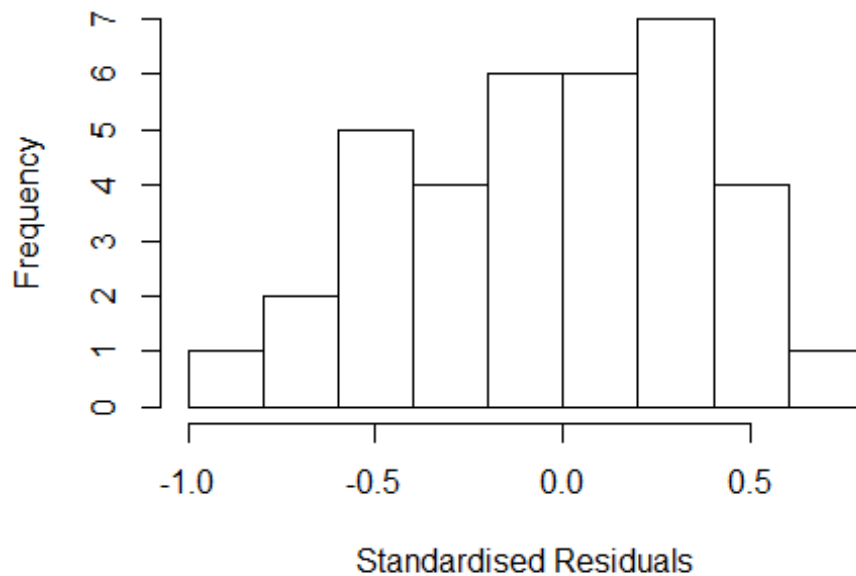
## Year2008          -0.00291    0.21563   -0.01  0.98934
## Year2009          0.00646    0.23287    0.03  0.97812
## Year2010         -0.47180    0.13196   -3.58  0.00160 **
## Year2011         -0.07474    0.22953   -0.33  0.74764
## Year2012         -0.57166    0.12653   -4.52  0.00015 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.528
## Multiple R-squared:  0.483, Adjusted R-squared:  0.214
## Convergence in 9 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 28 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.786  0.927  0.960  0.943  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          2.78e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##  nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##  trace.lev      mts      compute.rd
##      0          1000          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
## -7.52e-01 -3.52e-01 -5.55e-16  3.29e-01  8.00e-01
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   1.51e+00   1.82e-08  8.28e+07  <2e-16 ***
## LastAuthorFemale1 -1.63e-01   1.88e-01 -8.70e-01   0.394
## Year2000        7.88e-01   0.00e+00      Inf  <2e-16 ***
## Year2002       -3.91e-01   2.39e-08 -1.63e+07  <2e-16 ***
## Year2003       -7.93e-01   4.95e-01 -1.60e+00   0.123
## Year2005       -2.77e-01   2.60e-01 -1.07e+00   0.297
## Year2006        3.37e-01   1.29e-08  2.62e+07  <2e-16 ***
## Year2007        2.45e-01   1.86e-01  1.32e+00   0.201
```

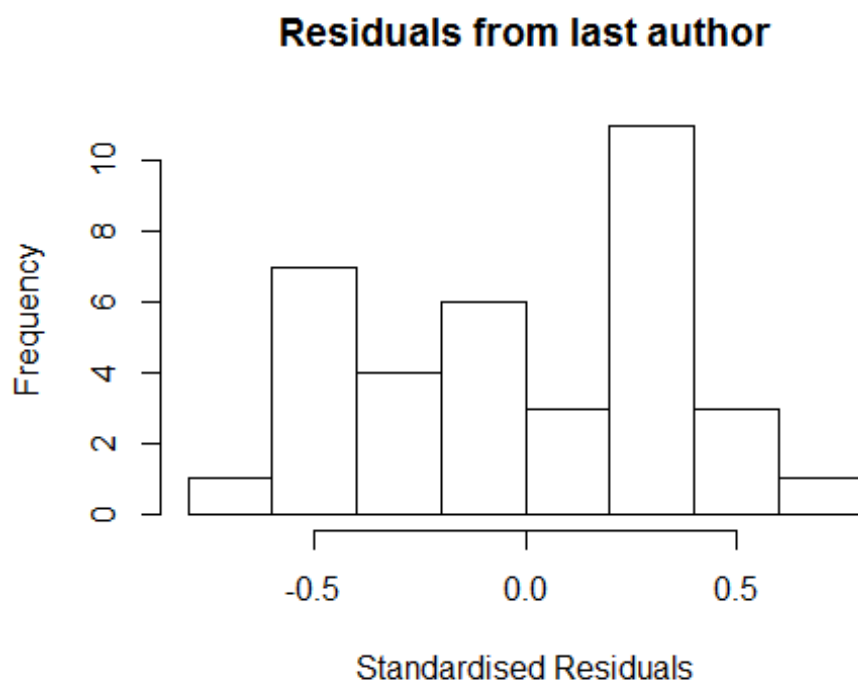
```

## Year2008          -1.40e-01   1.15e-01 -1.21e+00   0.237
## Year2009          -2.57e-02   2.13e-01 -1.20e-01   0.905
## Year2010          -5.25e-01   2.36e-01 -2.22e+00   0.036 *
## Year2011           2.57e-02   2.21e-01  1.20e-01   0.908
## Year2012          -5.58e-01   2.23e-01 -2.50e+00   0.020 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.543
## Multiple R-squared:  0.429, Adjusted R-squared:  0.13
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 30 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.812  0.909  0.960  0.945  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      2.78e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 36"
## [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
##    6    5    4    1    7    3    3    5    3   11    8    8    9    9    7
## 2011 2012
##    6   15
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    2    2    0    3    1    1    2    2    6    4    5    8    5    5
## 2011 2012
##    5   10

```

```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    2    2    0    3    1    1    2    1    5    4    3    7    4    4
## 2011 2012
##    4    8
## [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: 3.93597934253086"

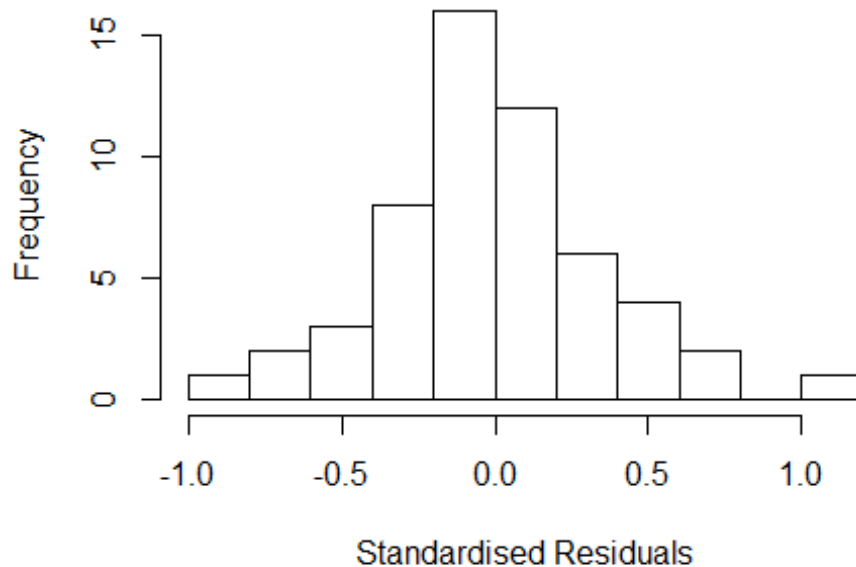
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## 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] "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"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.094e+01 1          3.308
## LastAuthorFemale  6.773e+00 1          2.602
```

| | | | |
|------------------|------------|----|-----|
| ## UniqueAuthors | -1.619e+33 | 4 | NaN |
| ## Year | -9.496e+33 | 15 | NaN |

Residuals from first and last author and team size



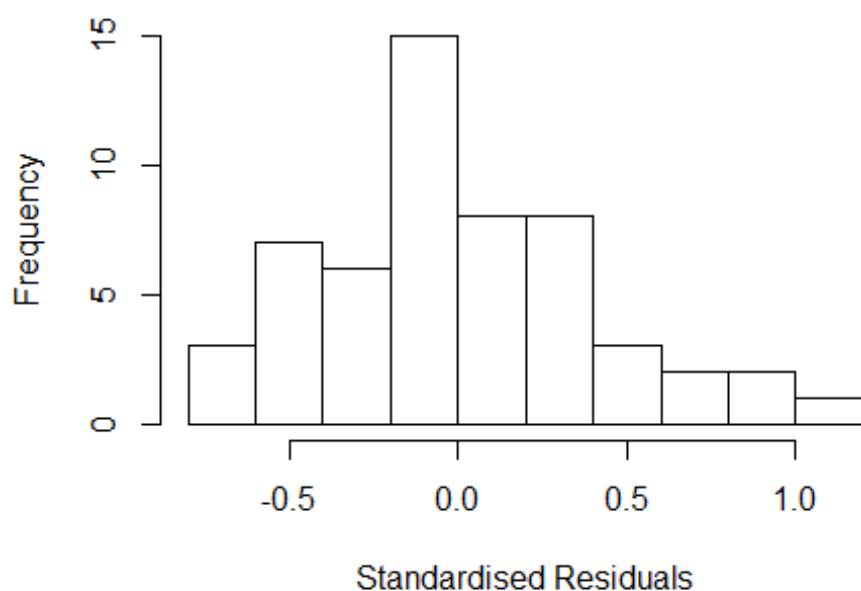
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -8.19e-01 -2.00e-01 -7.22e-16 1.64e-01 1.13e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.04538 0.31317 0.14 0.88566
## FirstAuthorFemale1 -0.31642 0.20019 -1.58 0.12350
## LastAuthorFemale1 0.24845 0.20282 1.22 0.22925
## UniqueAuthors2 0.74770 0.26020 2.87 0.00705 **
## UniqueAuthors3 1.08989 0.20601 5.29 7.8e-06 ***
## UniqueAuthors4 1.13544 0.29008 3.91 0.00043 ***
## UniqueAuthors5 1.22233 0.22955 5.32 7.1e-06 ***
```

```

## Year1997          0.28710      0.26648      1.08  0.28913
## Year1998          0.49177      0.23457      2.10  0.04379 *
## Year2000          0.00319      0.36951      0.01  0.99317
## Year2001          0.09992      0.19279      0.52  0.60772
## Year2002         -0.17072      0.25550     -0.67  0.50868
## Year2003          0.04751      0.25545      0.19  0.85361
## Year2004          0.84662      0.31317      2.70  0.01076 *
## Year2005          0.00433      0.32146      0.01  0.98934
## Year2006          0.25280      0.41406      0.61  0.54568
## Year2007         -0.26902      0.32685     -0.82  0.41638
## Year2008          0.04721      0.28426      0.17  0.86910
## Year2009          0.05901      0.22466      0.26  0.79443
## Year2010         -0.29691      0.29406     -1.01  0.32000
## Year2011          0.02908      0.28550      0.10  0.91948
## Year2012          0.03628      0.44372      0.08  0.93533
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.403
## Multiple R-squared:  0.437, Adjusted R-squared:  0.0795
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 45 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.409  0.908  0.971  0.922  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      1.82e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##   nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##   trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 9.830e+00 1          3.135
## LastAuthorFemale  6.919e+00 1          2.630
## Year              -4.600e-12 15          NaN

```

Residuals from first and last author



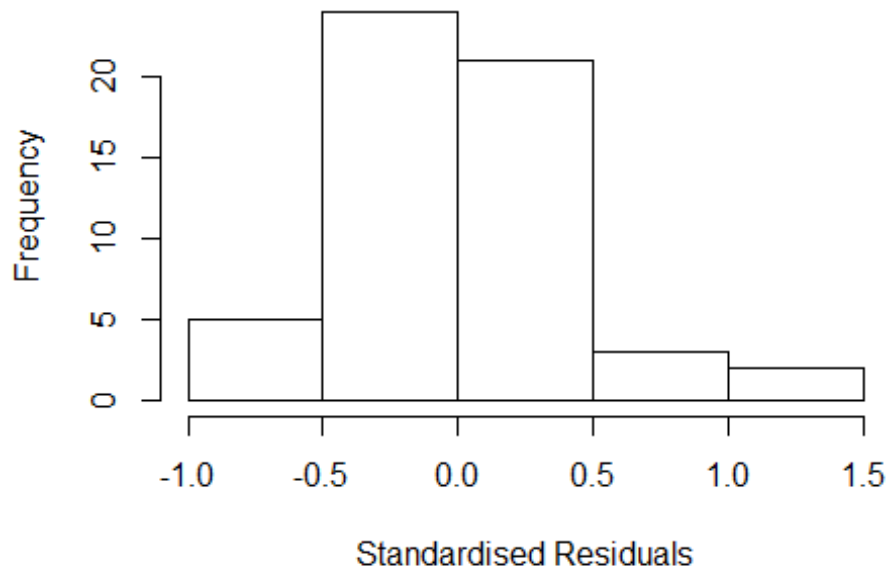
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.79170 -0.25422 -0.00461 0.22679 1.07712
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8476 0.1642 5.16 8.5e-06 ***
## FirstAuthorFemale1 -0.1935 0.2335 -0.83 0.413
## LastAuthorFemale1 0.1931 0.3597 0.54 0.595
## Year1997 0.4699 0.4539 1.04 0.307
## Year1998 0.0634 0.4546 0.14 0.890
## Year2000 -0.0559 0.8696 -0.06 0.949
## Year2001 0.0454 0.1642 0.28 0.784
## Year2002 0.2494 0.1642 1.52 0.137
## Year2003 0.4014 0.1738 2.31 0.027 *
## Year2004 0.0444 0.1642 0.27 0.788
## Year2005 0.3152 0.2561 1.23 0.226
## Year2006 0.0875 0.5121 0.17 0.865
```

```

## Year2007          0.0666      0.2710      0.25      0.807
## Year2008          0.2585      0.2684      0.96      0.342
## Year2009          0.1916      0.1859      1.03      0.309
## Year2010         -0.1107      0.2257     -0.49      0.627
## Year2011          0.3243      0.2164      1.50      0.142
## Year2012          0.3898      0.2602      1.50      0.143
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.426
## Multiple R-squared:  0.212, Adjusted R-squared:  -0.151
## Convergence in 30 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 46 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.502  0.893  0.951  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      1.82e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -6.432e-13 1      NaN
## Year              -6.432e-13 15      NaN

```


Residuals from first author

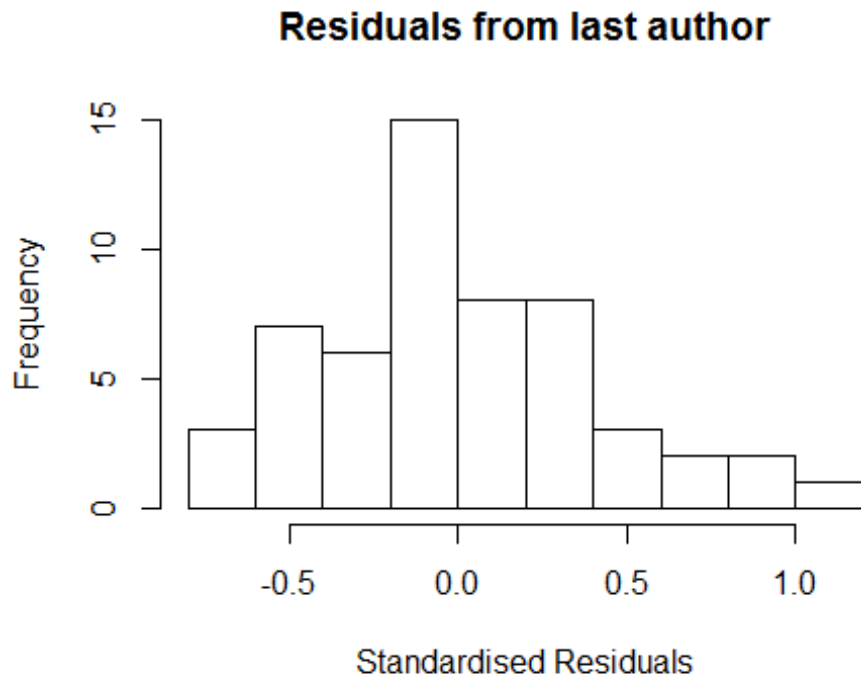


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -8.07e-01 -2.30e-01 -5.20e-17 2.27e-01 1.29e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8634 0.1541 5.60 2e-06 ***
## FirstAuthorFemale1 -0.2538 0.1826 -1.39 0.173
## Year1997 0.4541 0.4589 0.99 0.329
## Year1998 0.0476 0.4596 0.10 0.918
## Year2000 -0.1765 1.0123 -0.17 0.862
## Year2001 0.0296 0.1541 0.19 0.849
## Year2002 0.2336 0.1541 1.52 0.138
## Year2003 0.3856 0.1642 2.35 0.024 *
## Year2004 0.0286 0.1541 0.19 0.854
## Year2005 0.3709 0.2078 1.78 0.082 .
## Year2006 0.0685 0.5445 0.13 0.901
## Year2007 0.0511 0.2669 0.19 0.849
```

```

## Year2008          0.3080      0.2240      1.38      0.177
## Year2009          0.2392      0.1515      1.58      0.123
## Year2010         -0.0960      0.2045     -0.47      0.641
## Year2011          0.3876      0.1559      2.49      0.017 *
## Year2012          0.4281      0.2811      1.52      0.136
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.423
## Multiple R-squared:  0.218, Adjusted R-squared:  -0.111
## Convergence in 30 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 47 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.331  0.893  0.957  0.904  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      1.82e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##  nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##  trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 3.451 1          1.858
## Year             3.451 15          1.042

```



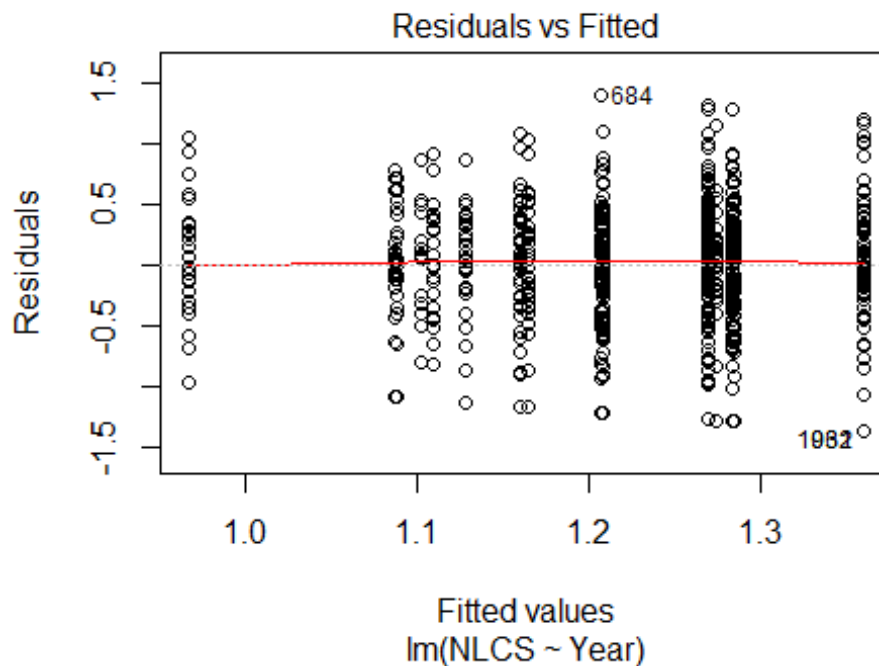
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8101 -0.3065 -0.0204 0.2742 0.9236
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.80158 0.17824 4.50 6.3e-05 ***
## LastAuthorFemale1 0.31438 0.21329 1.47 0.149
## Year1997 0.51592 0.42288 1.22 0.230
## Year1998 0.10942 0.42339 0.26 0.797
## Year2000 0.00855 0.53625 0.02 0.987
## Year2001 0.09142 0.17824 0.51 0.611
## Year2002 0.29542 0.17824 1.66 0.106
## Year2003 0.44742 0.18707 2.39 0.022 *
## Year2004 0.09042 0.17824 0.51 0.615
## Year2005 0.21611 0.24383 0.89 0.381
## Year2006 0.12834 0.37035 0.35 0.731
## Year2007 0.10584 0.27811 0.38 0.706
```

```

## Year2008          0.17047      0.26266      0.65      0.520
## Year2009          0.15881      0.19823      0.80      0.428
## Year2010         -0.16106      0.27539     -0.58      0.562
## Year2011          0.24154      0.21131      1.14      0.260
## Year2012          0.39579      0.26357      1.50      0.141
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.536
## Multiple R-squared:  0.186, Adjusted R-squared:  -0.157
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 48 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.747  0.926  0.961   0.941   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      1.82e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 55"
## [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
##   35   42   50   39   36   55   50   50   47   46   56   72   81   97  120
## 2011 2012
##  121  139
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   18   28   35   28   19   17   38   28   28   28   36   46   55   68   90
## 2011 2012
##   85  106

```

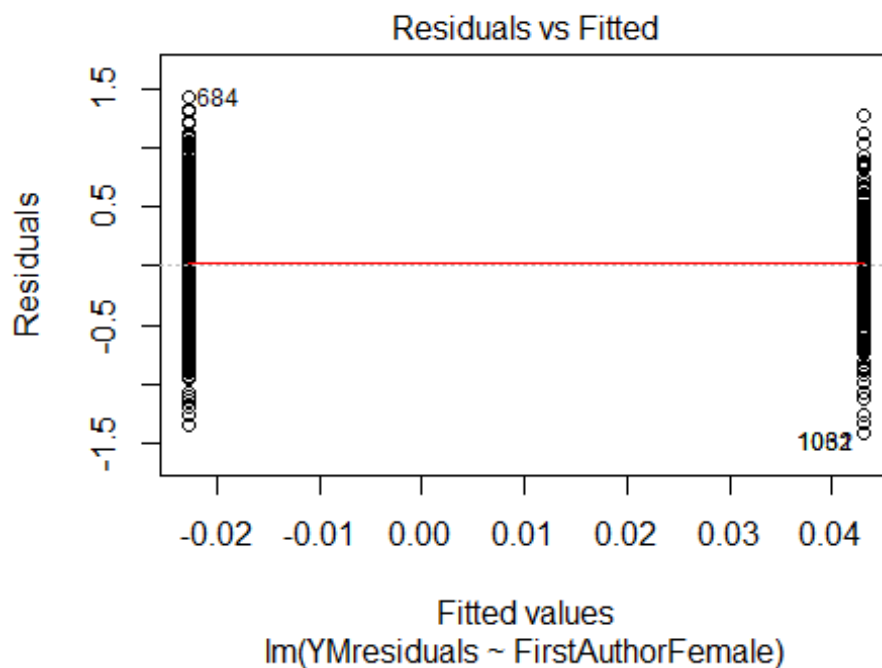
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 18 26 30 24 16 15 33 24 25 24 32 41 41 54 73
## 2011 2012
## 72 87
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS 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 = 8.3, df = 1, p-value = 0.004
## [1] "Female first author team size 2018 geometric mean: 4.29291366013518"
## [1] "Male first author team size 2018 geometric mean: 3.46313099781811"
## Warning in wilcox.test.default(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] "Female last author team size 2018 geometric mean: 5.23032591155292"
## [1] "Male last author team size 2018 geometric mean: 3.34076123746072"

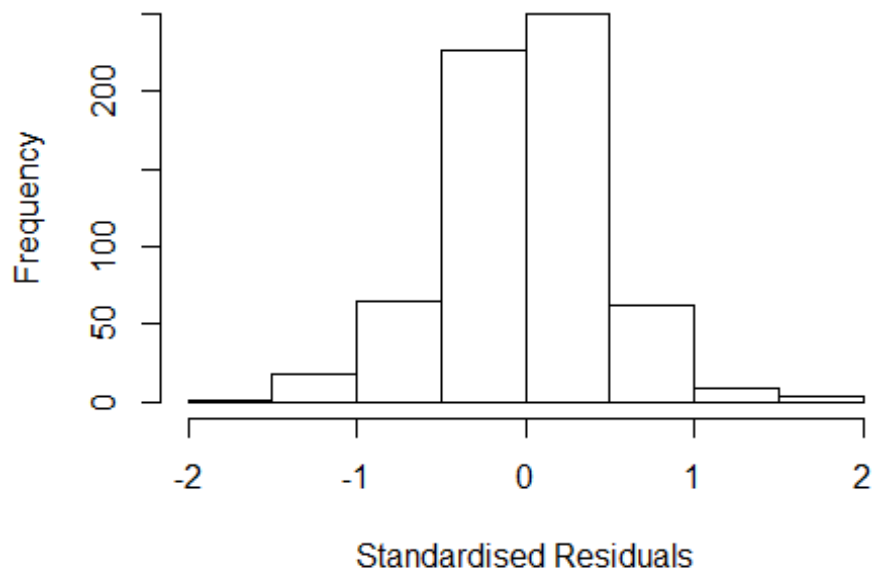
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 360, 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 | 1.112 | 1 | 1.055 |
| LastAuthorFemale | 1.353 | 1 | 1.163 |
| UniqueAuthors | 2.028 | 4 | 1.092 |
| Year | 2.616 | 16 | 1.031 |

Residuals from first and last author and team size



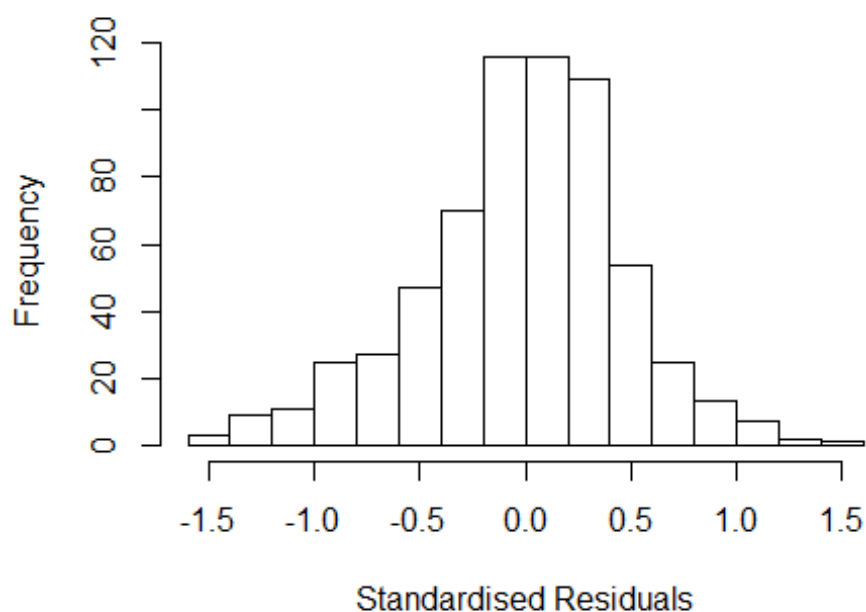
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.51825 -0.26339  0.00751  0.28058  1.88858
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.72071    0.18222   3.96 8.5e-05 ***
## FirstAuthorFemale1 0.07858    0.03685   2.13 0.03336 *
## LastAuthorFemale1 -0.17579    0.04861  -3.62 0.00032 ***
## UniqueAuthors2    0.44559    0.16561   2.69 0.00733 **
## UniqueAuthors3    0.51365    0.16057   3.20 0.00145 **
## UniqueAuthors4    0.45486    0.16331   2.79 0.00551 **
## UniqueAuthors5    0.51805    0.16171   3.20 0.00143 **
## Year1997         -0.01159    0.14753  -0.08 0.93743
## Year1998          0.02819    0.13757   0.20 0.83772
## Year1999         -0.12804    0.16364  -0.78 0.43425
```

```

## Year2000      -0.01679    0.17885   -0.09  0.92525
## Year2001      -0.02814    0.16914   -0.17  0.86792
## Year2002      -0.00430    0.15170   -0.03  0.97738
## Year2003       0.04060    0.14284    0.28  0.77634
## Year2004       0.06208    0.14395    0.43  0.66640
## Year2005       0.07828    0.14096    0.56  0.57889
## Year2006      -0.00329    0.13797   -0.02  0.98097
## Year2007       0.04878    0.14177    0.34  0.73090
## Year2008       0.10995    0.13574    0.81  0.41824
## Year2009       0.20531    0.14064    1.46  0.14485
## Year2010       0.14041    0.13064    1.07  0.28287
## Year2011       0.02243    0.13308    0.17  0.86623
## Year2012       0.10714    0.12600    0.85  0.39547
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.414
## Multiple R-squared:  0.125, Adjusted R-squared:  0.094
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 61 weights are ~= 1. The remaining 574 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0026 0.8590 0.9510 0.8810 0.9840 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00      5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      1.57e-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 1.132 1      1.064
## LastAuthorFemale  1.238 1      1.113
## Year              1.400 16      1.011

```


Residuals from first and last author



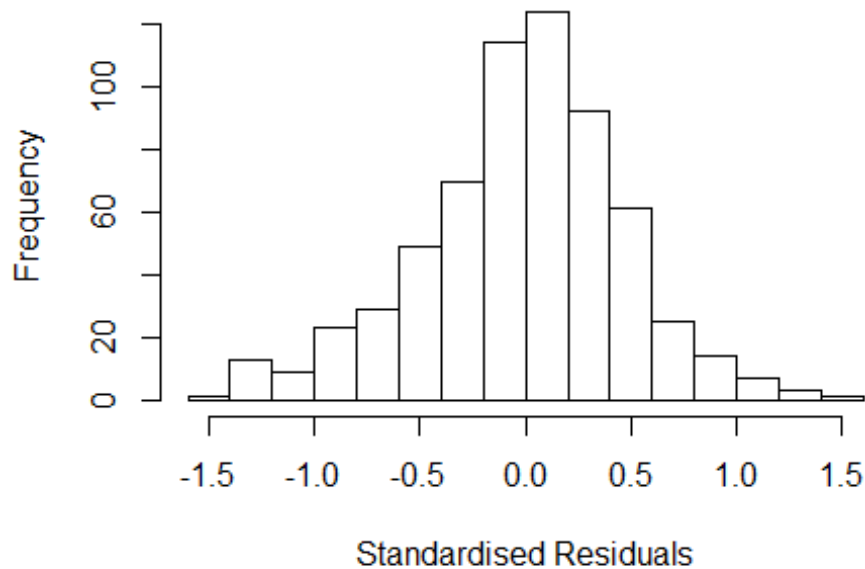
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.4981 -0.2905  0.0155  0.2935  1.4338
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.16068    0.12722   9.12 < 2e-16 ***
## FirstAuthorFemale1  0.09782    0.03828   2.56  0.01085 *
## LastAuthorFemale1 -0.17979    0.04738  -3.79  0.00016 ***
## Year1997        -0.04373    0.15532  -0.28  0.77838
## Year1998         0.06168    0.14507   0.43  0.67084
## Year1999        -0.16497    0.17872  -0.92  0.35635
## Year2000        -0.03154    0.20420  -0.15  0.87732
## Year2001        -0.06027    0.17726  -0.34  0.73398
## Year2002         0.00783    0.16111   0.05  0.96125
## Year2003         0.03206    0.15986   0.20  0.84113
## Year2004         0.10125    0.15212   0.67  0.50590
## Year2005         0.10665    0.15044   0.71  0.47866
```

```

## Year2006          0.01154      0.15403      0.07  0.94031
## Year2007          0.07193      0.15141      0.48  0.63492
## Year2008          0.15457      0.14340      1.08  0.28150
## Year2009          0.23965      0.14566      1.65  0.10044
## Year2010          0.18559      0.13780      1.35  0.17853
## Year2011          0.04622      0.14026      0.33  0.74186
## Year2012          0.13706      0.13546      1.01  0.31203
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.43
## Multiple R-squared:  0.0671, Adjusted R-squared:  0.0398
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 63 weights are ~= 1. The remaining 572 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.199  0.854  0.949  0.886  0.986  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      1.57e-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 1.153 1          1.074
## Year              1.153 16          1.004

```

Residuals from first author



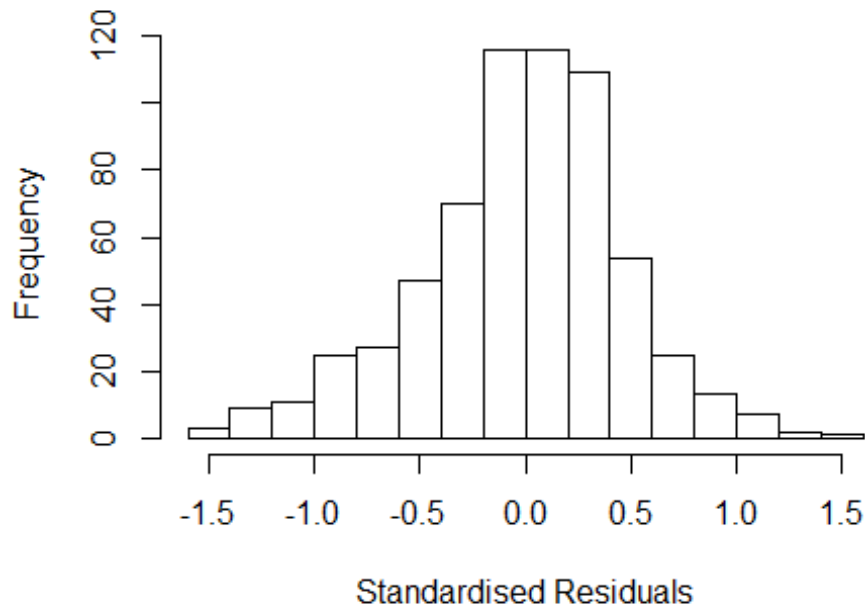
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4471 -0.2956 0.0129 0.2844 1.4540
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0999 0.1225 8.98 <2e-16 ***
## FirstAuthorFemale1 0.0857 0.0386 2.22 0.027 *
## Year1997 -0.0245 0.1562 -0.16 0.875
## Year1998 0.0822 0.1444 0.57 0.570
## Year1999 -0.1426 0.1760 -0.81 0.418
## Year2000 -0.0144 0.2042 -0.07 0.944
## Year2001 -0.0262 0.1734 -0.15 0.880
## Year2002 0.0485 0.1558 0.31 0.756
## Year2003 0.0552 0.1576 0.35 0.726
## Year2004 0.1447 0.1497 0.97 0.334
## Year2005 0.1540 0.1480 1.04 0.298
## Year2006 0.0521 0.1498 0.35 0.728
```

```

## Year2007          0.1339      0.1474      0.91      0.364
## Year2008          0.1978      0.1387      1.43      0.154
## Year2009          0.2615      0.1419      1.84      0.066 .
## Year2010          0.1945      0.1359      1.43      0.153
## Year2011          0.0866      0.1378      0.63      0.530
## Year2012          0.1657      0.1328      1.25      0.213
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.434
## Multiple R-squared:  0.046, Adjusted R-squared:  0.0197
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 59 weights are ~= 1. The remaining 576 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.240  0.861  0.950  0.886  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      1.57e-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 1.236 1      1.112
## Year              1.236 16      1.007

```

Residuals from last author



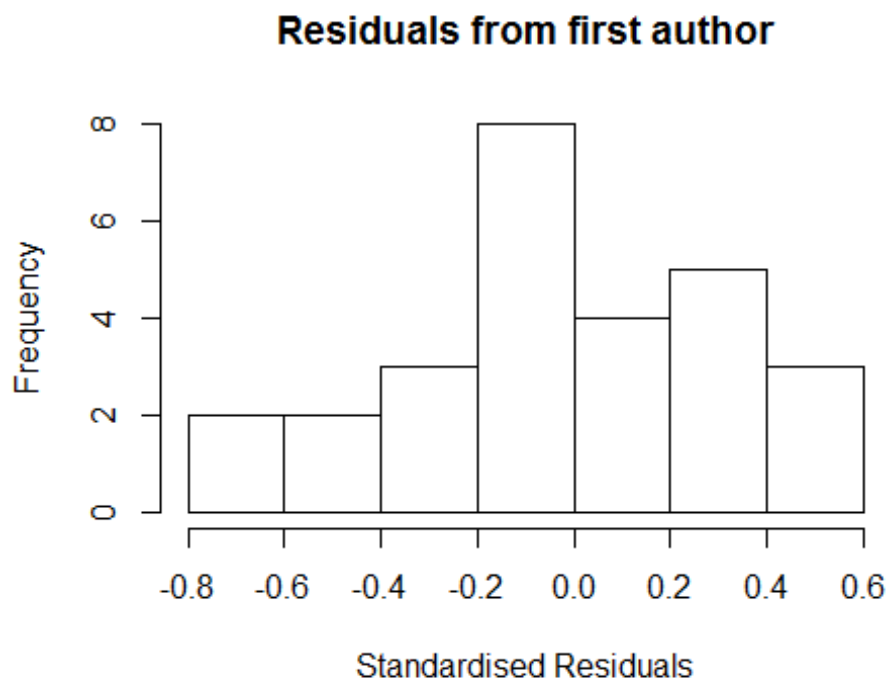
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4254 -0.2884 0.0202 0.2792 1.3977
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.16970 0.12007 9.74 < 2e-16 ***
## LastAuthorFemale1 -0.17038 0.04842 -3.52 0.00047 ***
## Year1997 -0.02074 0.15085 -0.14 0.89070
## Year1998 0.08595 0.13830 0.62 0.53451
## Year1999 -0.13371 0.17114 -0.78 0.43495
## Year2000 0.00197 0.20161 0.01 0.99222
## Year2001 -0.02259 0.17241 -0.13 0.89580
## Year2002 0.02702 0.15432 0.18 0.86108
## Year2003 0.07774 0.15477 0.50 0.61563
## Year2004 0.14345 0.14232 1.01 0.31388
## Year2005 0.12707 0.14285 0.89 0.37407
## Year2006 0.03858 0.14672 0.26 0.79266
```

```

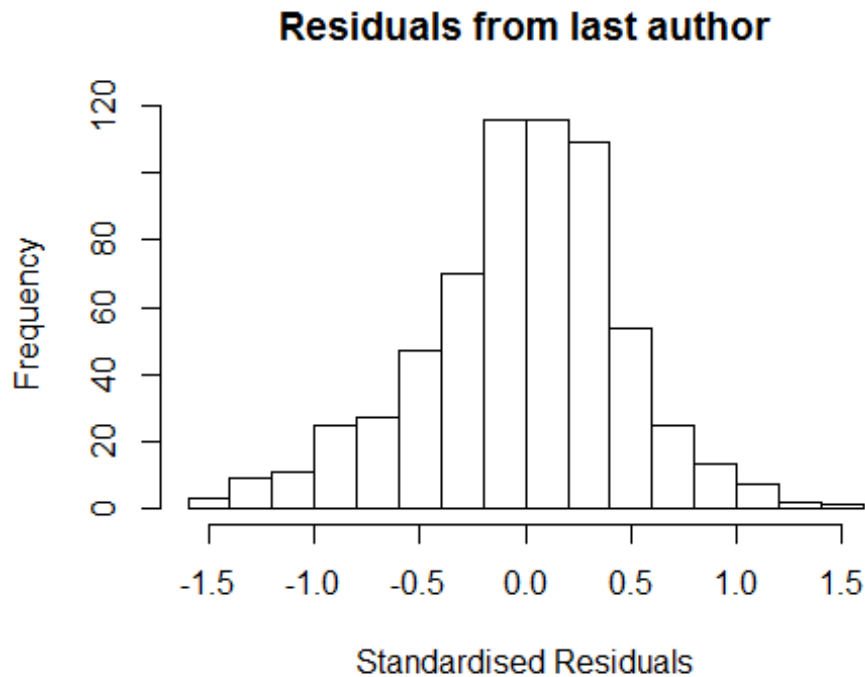
## Year2007      0.11043    0.14312    0.77  0.44065
## Year2008      0.18191    0.13683    1.33  0.18419
## Year2009      0.25573    0.13894    1.84  0.06616 .
## Year2010      0.20839    0.12967    1.61  0.10855
## Year2011      0.06888    0.13234    0.52  0.60293
## Year2012      0.15867    0.12700    1.25  0.21201
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.431
## Multiple R-squared:  0.0572, Adjusted R-squared:  0.0312
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 57 weights are ~= 1. The remaining 578 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.252  0.855  0.951  0.886  0.986  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.57e-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: 635"
## [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 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    2    3    4    1    2    5    3    6    9    3    6    4    3    1
##
## 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    0    2    2    1    0    0    2    1    2    5    1    5    4    2    1
##
## 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    0    2    2    1    0    0    2    1    2    5    1    5    3    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: 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"
##
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.376e+14  1      1.541e+07
## Year              2.376e+14 11      4.502e+00
```



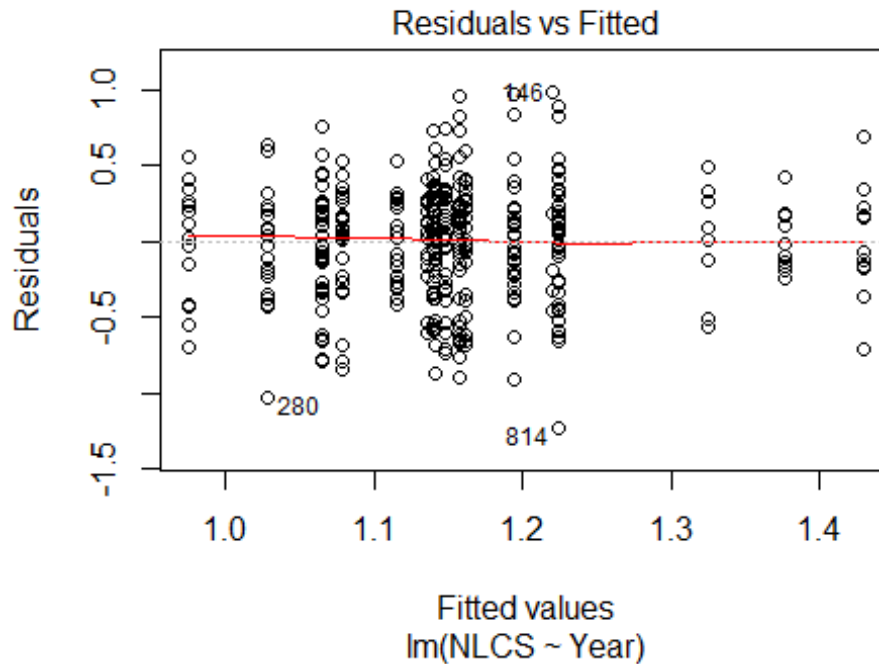
```
## [1] "Regression 4: Last author gender, Year as factors"
##
##           GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 3.333  1      1.826
## Year              3.333 11      1.056
```



```
## [1] "Sample size for the above analysis: 27"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1602"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    25   23   21   27   14   20   37   31   25   30   34   45   32   45   49
## 2011 2012
##    56   53
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    17   11    8   17    6   12   25   27   23   20   28   32   26   37   33
## 2011 2012
##    44   37
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    16   11    7   17    5   12   23   27   21   19   27   31   26   34   30
## 2011 2012
##    38   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 = 18, df = 16, p-value = 0.3
```



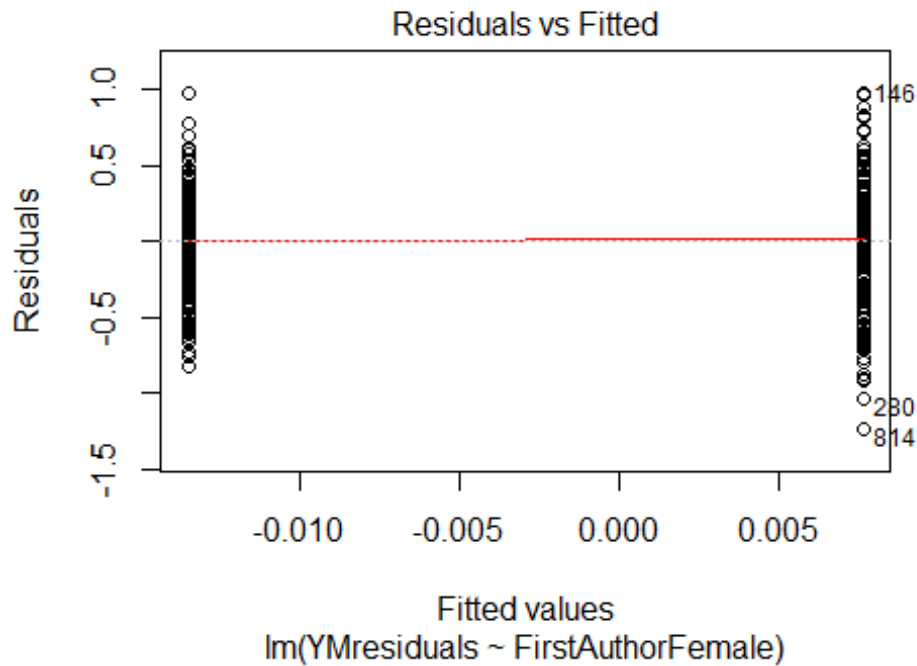
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.3, df = 1, p-value = 0.1

## [1] "Female first author team size 2018 geometric mean: 5.26143373051742"
## [1] "Male first author team size 2018 geometric mean: 3.04009615602979"

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

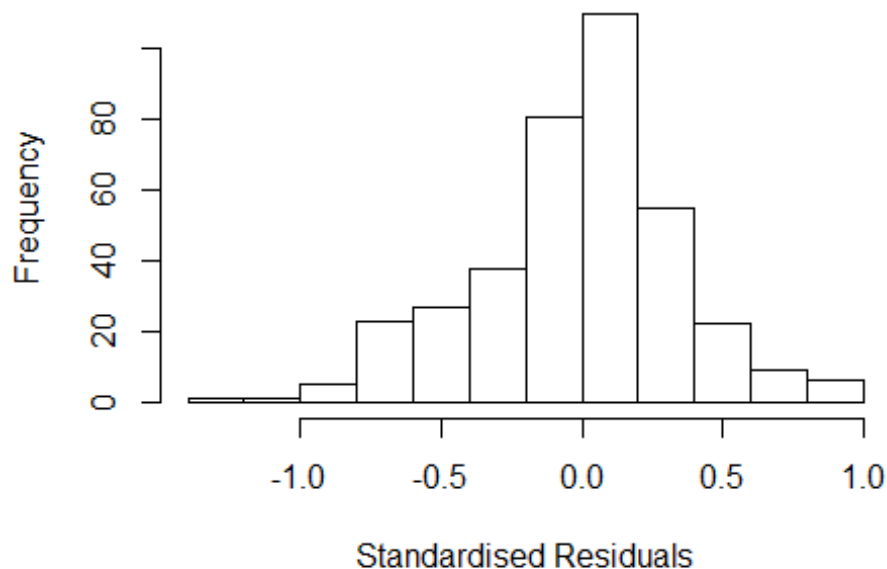
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 26, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6.16014057648205"
## [1] "Male last author team size 2018 geometric mean: 3.00899287331754"

## Warning in wilcox.test.default(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 = 0.06
## 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 1.280 1          1.132
## LastAuthorFemale  1.350 1          1.162
## UniqueAuthors    4.348 4          1.202
## Year              6.049 16         1.058
```

Residuals from first and last author and team size



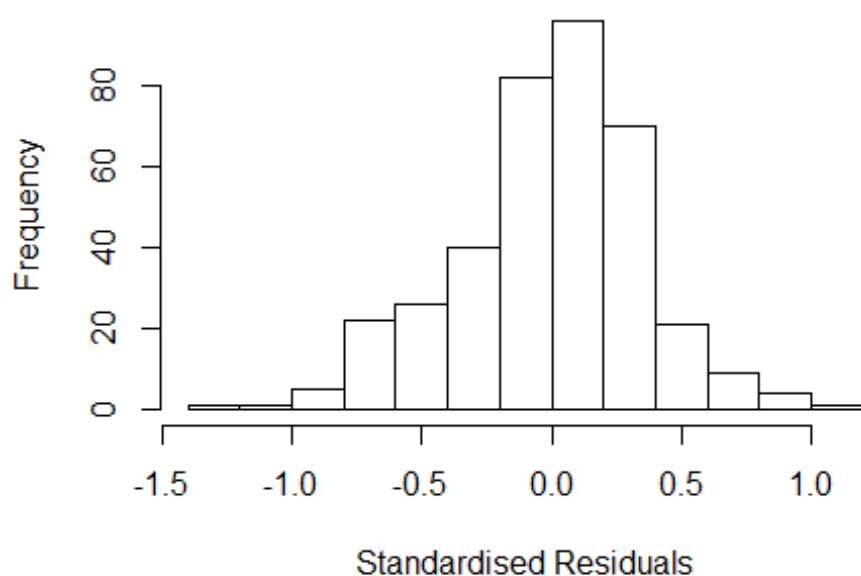
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2498 -0.1997 0.0286 0.1963 0.9624
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8721 0.1968 4.43 1.3e-05 ***
## FirstAuthorFemale1 -0.0328 0.0382 -0.86 0.391
## LastAuthorFemale1 0.0416 0.0580 0.72 0.474
## UniqueAuthors2 0.2216 0.1186 1.87 0.062 .
## UniqueAuthors3 0.1873 0.1216 1.54 0.124
## UniqueAuthors4 0.2176 0.1265 1.72 0.086 .
## UniqueAuthors5 0.2386 0.1241 1.92 0.055 .
## Year1997 0.3748 0.1837 2.04 0.042 *
## Year1998 0.3219 0.2406 1.34 0.182
## Year1999 0.1066 0.1521 0.70 0.484
```

```

## Year2000          0.1259      0.2698      0.47      0.641
## Year2001          0.2765      0.1527      1.81      0.071 .
## Year2002          0.0870      0.1525      0.57      0.569
## Year2003          0.1029      0.1503      0.68      0.494
## Year2004         -0.0122      0.1611     -0.08      0.940
## Year2005          0.0564      0.1521      0.37      0.711
## Year2006          0.0666      0.1477      0.45      0.652
## Year2007          0.1221      0.1568      0.78      0.437
## Year2008          0.0846      0.1596      0.53      0.597
## Year2009          0.1543      0.1480      1.04      0.298
## Year2010          0.0929      0.1487      0.62      0.532
## Year2011          0.0335      0.1456      0.23      0.818
## Year2012          0.1561      0.1511      1.03      0.302
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.312
## Multiple R-squared:  0.0745, Adjusted R-squared:  0.0172
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 33 weights are ~= 1. The remaining 345 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0724 0.8490 0.9520 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          2.65e-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 1.187 1          1.090
## LastAuthorFemale 1.205 1          1.098
## Year          1.423 16          1.011

```

Residuals from first and last author



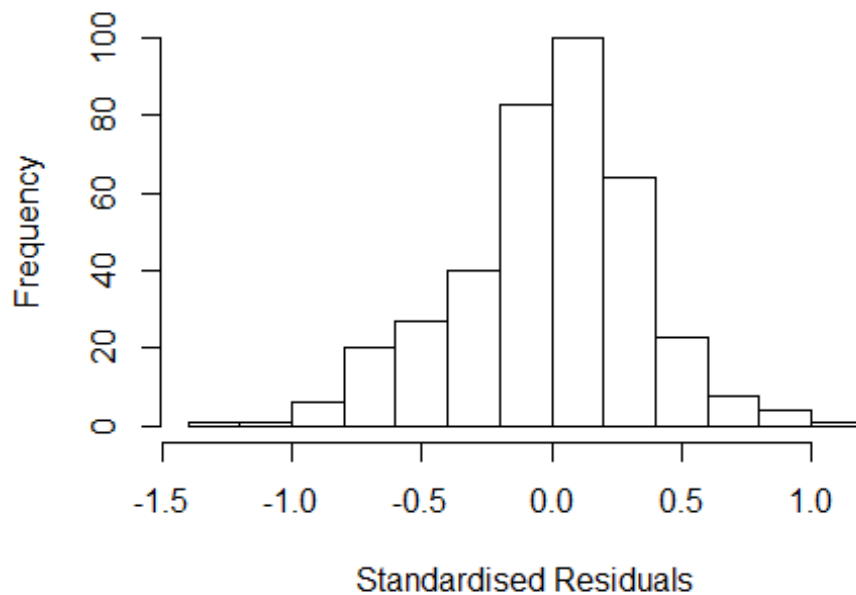
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2484 -0.1993  0.0291  0.2144  1.0491
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.05580    0.11776   8.97   <2e-16 ***
## FirstAuthorFemale1 -0.03166    0.03828  -0.83    0.409
## LastAuthorFemale1  0.02819    0.05741   0.49    0.624
## Year1997         0.38778    0.16054   2.42    0.016 *
## Year1998         0.33371    0.20126   1.66    0.098 .
## Year1999         0.10143    0.13243   0.77    0.444
## Year2000         0.07189    0.23452   0.31    0.759
## Year2001         0.31482    0.13060   2.41    0.016 *
## Year2002         0.09477    0.13724   0.69    0.490
## Year2003         0.12203    0.13371   0.91    0.362
## Year2004         0.00967    0.13901   0.07    0.945
## Year2005         0.08140    0.13229   0.62    0.539
```

```

## Year2006          0.09082      0.12690      0.72      0.475
## Year2007          0.15157      0.13549      1.12      0.264
## Year2008          0.11830      0.14236      0.83      0.407
## Year2009          0.18551      0.13203      1.41      0.161
## Year2010          0.12085      0.12949      0.93      0.351
## Year2011          0.07012      0.12628      0.56      0.579
## Year2012          0.19260      0.13322      1.45      0.149
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.317
## Multiple R-squared:  0.057, Adjusted R-squared:  0.00974
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 348 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0853 0.8540 0.9500 0.8850 0.9850 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      2.65e-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 1.197 1      1.094
## Year              1.197 16      1.006

```

Residuals from first author



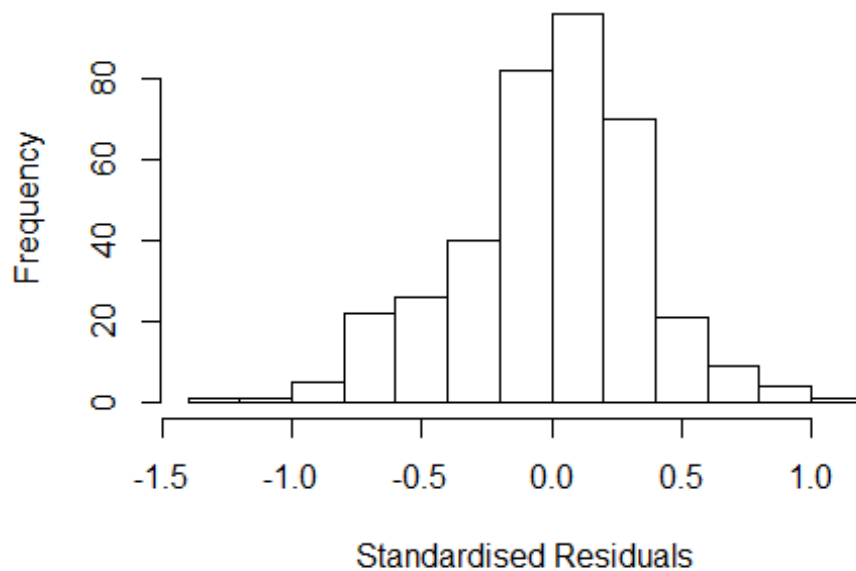
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2531 -0.1985  0.0291  0.2158  1.0858
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0582    0.1175   9.01  <2e-16 ***
## FirstAuthorFemale1 -0.0300    0.0385  -0.78   0.437
## Year1997          0.3880    0.1613   2.41   0.017 *
## Year1998          0.3366    0.2013   1.67   0.095 .
## Year1999          0.1005    0.1322   0.76   0.448
## Year2000          0.0610    0.2254   0.27   0.787
## Year2001          0.3169    0.1310   2.42   0.016 *
## Year2002          0.0967    0.1386   0.70   0.486
## Year2003          0.1228    0.1345   0.91   0.362
## Year2004          0.0102    0.1393   0.07   0.942
## Year2005          0.0828    0.1322   0.63   0.531
## Year2006          0.0885    0.1268   0.70   0.486
```

```

## Year2007          0.1528      0.1355      1.13      0.260
## Year2008          0.1209      0.1424      0.85      0.396
## Year2009          0.1866      0.1325      1.41      0.160
## Year2010          0.1244      0.1293      0.96      0.337
## Year2011          0.0716      0.1263      0.57      0.571
## Year2012          0.1949      0.1336      1.46      0.146
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.316
## Multiple R-squared:  0.0565, Adjusted R-squared:  0.012
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 350 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0815 0.8500 0.9520 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      2.65e-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 1.21 1      1.100
## Year      1.21 16      1.006

```


Residuals from last author



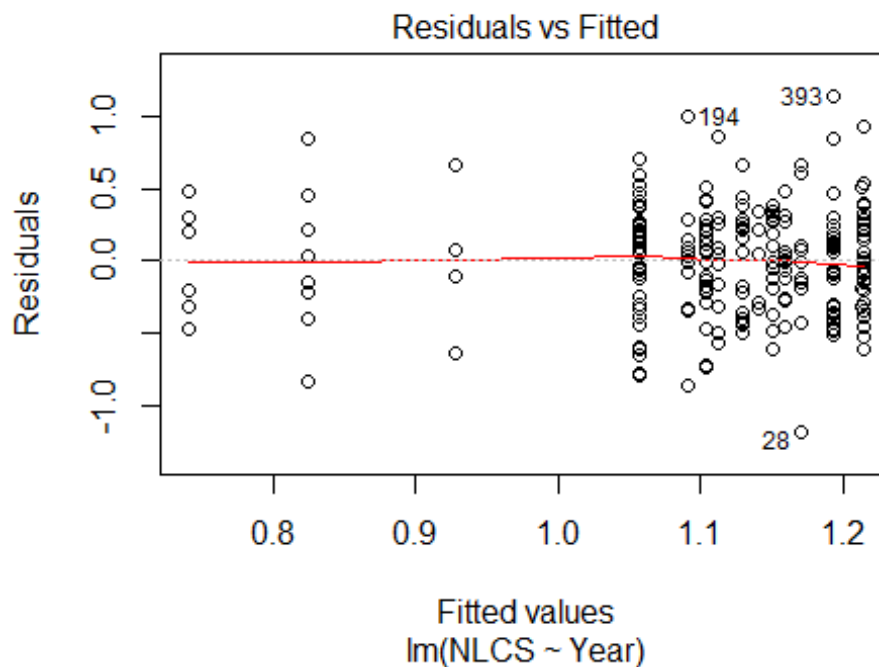
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2341 -0.2018 0.0362 0.2133 1.0533
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0471 0.1173 8.93 <2e-16 ***
## LastAuthorFemale1 0.0245 0.0569 0.43 0.667
## Year1997 0.3818 0.1607 2.38 0.018 *
## Year1998 0.3219 0.2053 1.57 0.118
## Year1999 0.1017 0.1338 0.76 0.447
## Year2000 0.0801 0.2340 0.34 0.732
## Year2001 0.3200 0.1296 2.47 0.014 *
## Year2002 0.0947 0.1385 0.68 0.495
## Year2003 0.1204 0.1356 0.89 0.375
## Year2004 0.0101 0.1398 0.07 0.942
## Year2005 0.0783 0.1337 0.59 0.558
## Year2006 0.0887 0.1285 0.69 0.490
```

```

## Year2007          0.1451      0.1360      1.07      0.287
## Year2008          0.1139      0.1443      0.79      0.431
## Year2009          0.1814      0.1336      1.36      0.176
## Year2010          0.1193      0.1307      0.91      0.362
## Year2011          0.0628      0.1283      0.49      0.625
## Year2012          0.1870      0.1345      1.39      0.165
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.317
## Multiple R-squared:  0.0554, Adjusted R-squared:  0.0108
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 354 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0962 0.8620 0.9530 0.8870 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      2.65e-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: 378"
## [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
##   11   9  10  12  11  15  14  22  12  16  18  34  34  32  42
## 2011 2012
##   31  29
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    5    4    6    3    5   10   16    9   12   11   15   28   23   28
## 2011 2012

```

```
## 25 18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 6 5 4 6 3 5 8 16 8 11 10 13 27 21 24
## 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 = 13, df = 16, p-value = 0.6
```



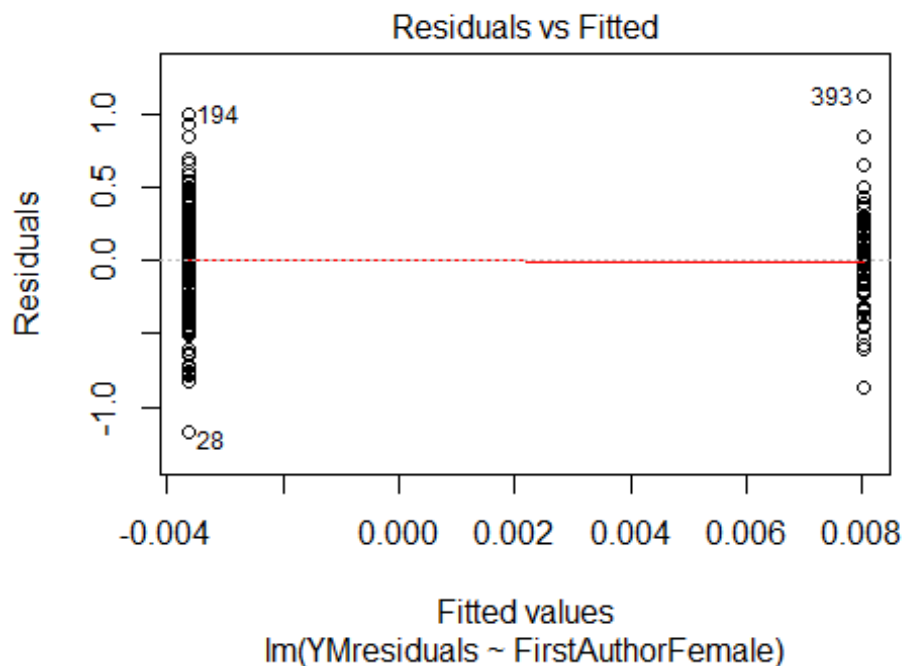
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.1, df = 1, p-value = 0.3

## [1] "Female first author team size 2018 geometric mean: 4.44062371460629"
## [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
```

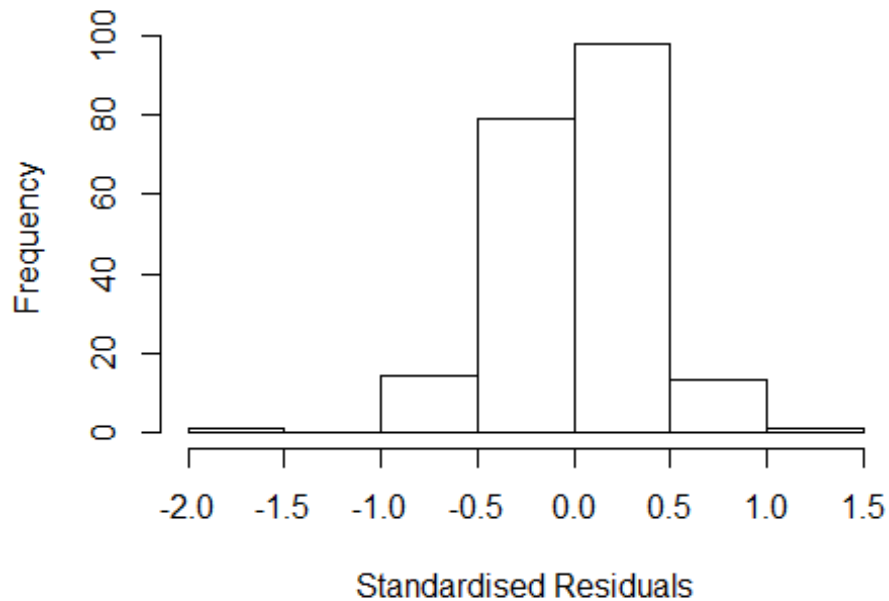
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 19, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.46410161513775"
## [1] "Male last author team size 2018 geometric mean: 3.98249845181447"

## Warning in wilcox.test.default(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 = 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 1.401  1      1.184
## LastAuthorFemale  1.863  1      1.365
## UniqueAuthors    4.416  4      1.204
## Year              7.167 16      1.063
```

Residuals from first and last author and team size



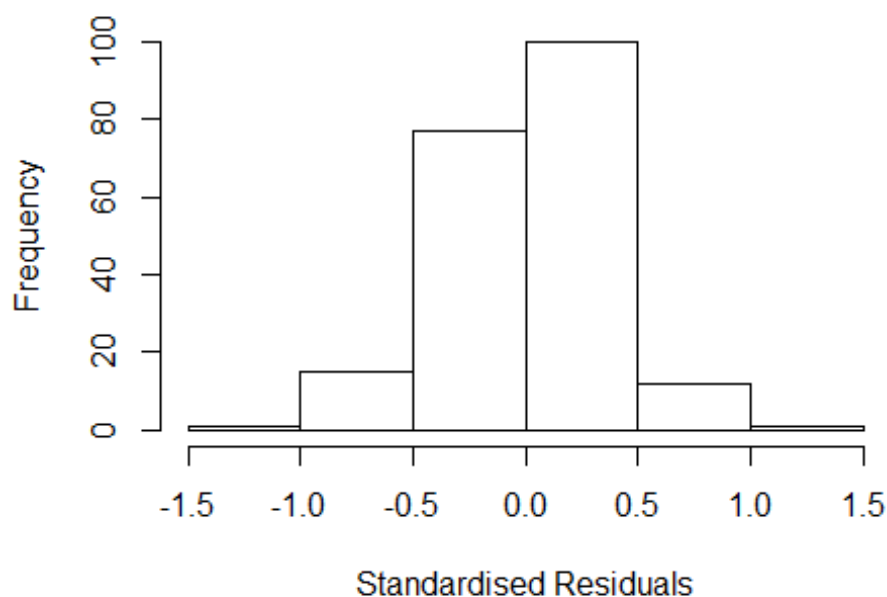
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5588 -0.2366 0.0279 0.2089 1.0706
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.4580 0.2241 2.04 0.04243 *
## FirstAuthorFemale1 0.0145 0.0540 0.27 0.78840
## LastAuthorFemale1 0.0114 0.0922 0.12 0.90178
## UniqueAuthors2 0.2932 0.1656 1.77 0.07824 .
## UniqueAuthors3 0.3405 0.1607 2.12 0.03548 *
## UniqueAuthors4 0.2443 0.1652 1.48 0.14082
## UniqueAuthors5 0.3554 0.1640 2.17 0.03156 *
## Year1997 0.7603 0.2158 3.52 0.00054 ***
## Year1998 0.1262 0.3303 0.38 0.70294
## Year1999 0.3577 0.2075 1.72 0.08647 .
```

```

## Year2000          0.4056      0.3152      1.29  0.19978
## Year2001          0.3792      0.2193      1.73  0.08541 .
## Year2002          0.3438      0.2439      1.41  0.16042
## Year2003          0.3354      0.2054      1.63  0.10424
## Year2004         -0.0315      0.2302     -0.14  0.89126
## Year2005          0.3781      0.1872      2.02  0.04481 *
## Year2006          0.3523      0.1880      1.87  0.06261 .
## Year2007          0.3905      0.1965      1.99  0.04842 *
## Year2008          0.4520      0.1750      2.58  0.01060 *
## Year2009          0.3880      0.1739      2.23  0.02688 *
## Year2010          0.4256      0.1771      2.40  0.01724 *
## Year2011          0.3049      0.1806      1.69  0.09300 .
## Year2012          0.4043      0.1866      2.17  0.03153 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.344
## Multiple R-squared:  0.143, Adjusted R-squared:  0.04
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 191 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0043 0.8790 0.9570 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      4.85e-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 1.343 1      1.159
## LastAuthorFemale  1.712 1      1.308
## Year              2.233 16      1.025

```

Residuals from first and last author



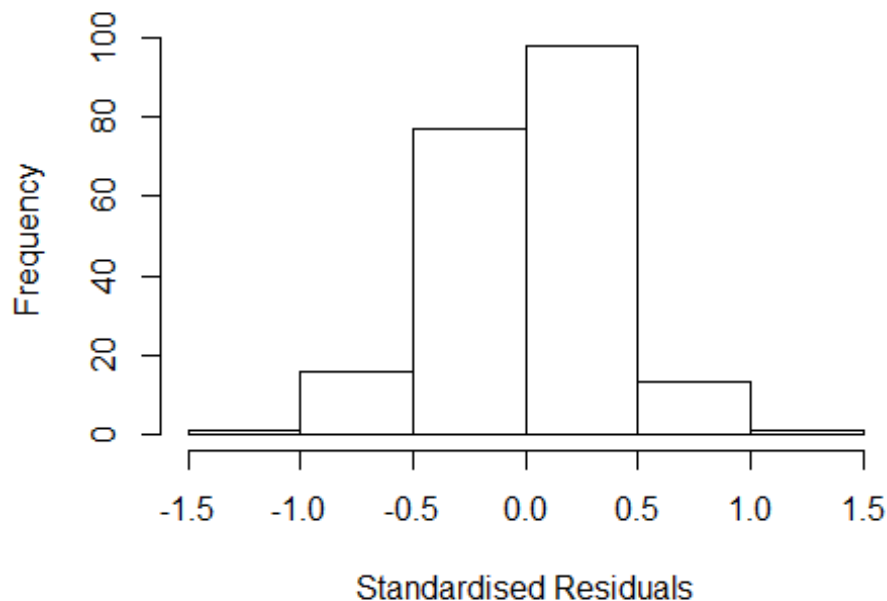
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4419 -0.2680 0.0361 0.2268 1.1273
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.73733 0.16359 4.51 1.2e-05 ***
## FirstAuthorFemale1 0.01531 0.05525 0.28 0.7819
## LastAuthorFemale1 -0.02646 0.09347 -0.28 0.7774
## Year1997 0.70459 0.29551 2.38 0.0181 *
## Year1998 0.17804 0.33998 0.52 0.6011
## Year1999 0.39498 0.20985 1.88 0.0614 .
## Year2000 0.45811 0.31108 1.47 0.1425
## Year2001 0.41148 0.21179 1.94 0.0535 .
## Year2002 0.36586 0.24259 1.51 0.1332
## Year2003 0.36006 0.20552 1.75 0.0814 .
## Year2004 0.00332 0.22561 0.01 0.9883
## Year2005 0.42114 0.18472 2.28 0.0237 *
```

```

## Year2006          0.38652      0.18466      2.09      0.0377 *
## Year2007          0.45119      0.19098      2.36      0.0192 *
## Year2008          0.48049      0.17555      2.74      0.0068 **
## Year2009          0.37621      0.17805      2.11      0.0359 *
## Year2010          0.44410      0.17264      2.57      0.0109 *
## Year2011          0.34786      0.18498      1.88      0.0616 .
## Year2012          0.45165      0.18916      2.39      0.0179 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.351
## Multiple R-squared:  0.111, Adjusted R-squared:  0.0251
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 193 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0534 0.8910 0.9500 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      4.85e-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 1.317 1      1.148
## Year      1.317 16      1.009

```


Residuals from first author



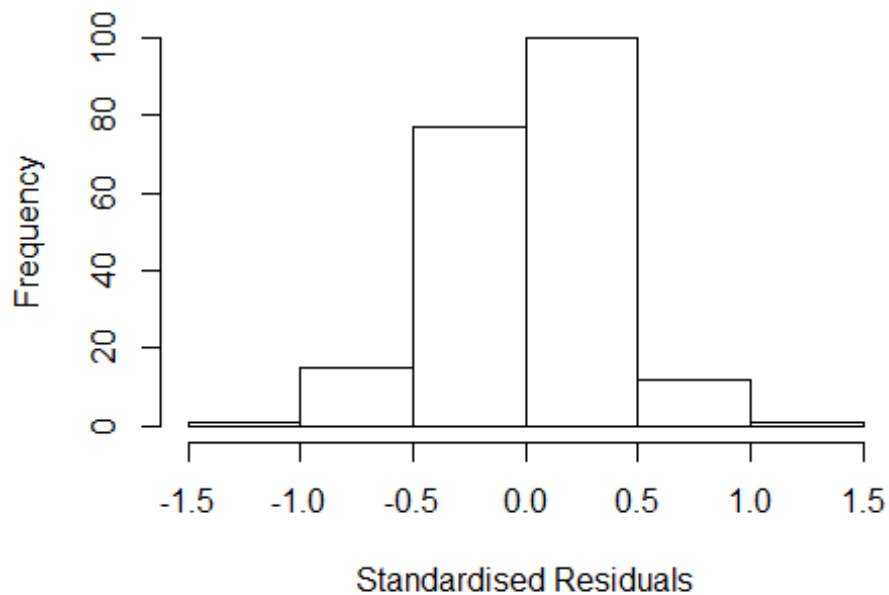
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.4335 -0.2628  0.0333  0.2336  1.1265
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.737257   0.162458   4.54    1e-05 ***
## FirstAuthorFemale1 0.015813   0.054979   0.29    0.7739
## Year1997        0.696287   0.310811   2.24    0.0262 *
## Year1998        0.167349   0.328318   0.51    0.6108
## Year1999        0.396006   0.209209   1.89    0.0599 .
## Year2000        0.432627   0.292492   1.48    0.1408
## Year2001        0.405321   0.207204   1.96    0.0519 .
## Year2002        0.368250   0.239289   1.54    0.1255
## Year2003        0.352563   0.201686   1.75    0.0821 .
## Year2004        0.000467   0.222479   0.00    0.9983
## Year2005        0.418243   0.182644   2.29    0.0231 *
## Year2006        0.388234   0.184296   2.11    0.0365 *
```

```

## Year2007          0.443978    0.187792    2.36    0.0191 *
## Year2008          0.480905    0.174495    2.76    0.0064 **
## Year2009          0.373750    0.176706    2.12    0.0357 *
## Year2010          0.444475    0.171859    2.59    0.0105 *
## Year2011          0.339295    0.179392    1.89    0.0601 .
## Year2012          0.448101    0.187497    2.39    0.0178 *
## ---
## 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:  0.0283
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 191 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0722 0.8890 0.9530 0.9050 0.9870 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      4.85e-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 1.734 1          1.317
## Year          1.734 16          1.017

```

Residuals from last author



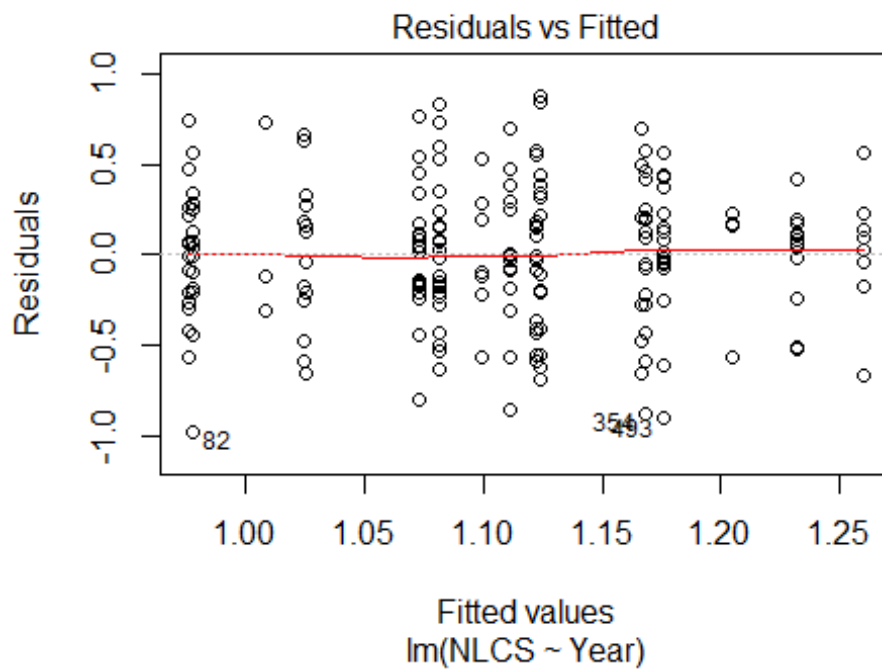
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4518 -0.2759 0.0389 0.2230 1.1406
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.74041 0.16558 4.47 1.3e-05 ***
## LastAuthorFemale1 -0.02853 0.09358 -0.30 0.761
## Year1997 0.71135 0.29148 2.44 0.016 *
## Year1998 0.18927 0.34268 0.55 0.581
## Year1999 0.39885 0.21566 1.85 0.066 .
## Year2000 0.47037 0.30916 1.52 0.130
## Year2001 0.40903 0.21342 1.92 0.057 .
## Year2002 0.36391 0.24759 1.47 0.143
## Year2003 0.36693 0.20302 1.81 0.072 .
## Year2004 0.00183 0.22945 0.01 0.994
## Year2005 0.42599 0.18597 2.29 0.023 *
## Year2006 0.38659 0.18995 2.04 0.043 *
```

```

## Year2007      0.45243      0.19179      2.36      0.019 *
## Year2008      0.48058      0.17923      2.68      0.008 **
## Year2009      0.37987      0.17829      2.13      0.034 *
## Year2010      0.44304      0.17784      2.49      0.014 *
## Year2011      0.35240      0.18451      1.91      0.058 .
## Year2012      0.45359      0.19104      2.37      0.019 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.342
## Multiple R-squared:  0.112, Adjusted R-squared:  0.0321
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 194 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0321 0.8840 0.9480 0.8990 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      4.85e-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: 206"
## [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
##   10   8   11   16   10   10   9   8   18   17   21   14   23   18   32
## 2011 2012
##   20   23
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7    4    9   14    7    6    4    7   14   15   16   14   22   17   26
## 2011 2012

```

```
## 13 19
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 6 3 9 11 6 4 2 6 12 14 15 14 20 14 24
## 2011 2012
## 12 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 = 11, df = 16, p-value = 0.8
```



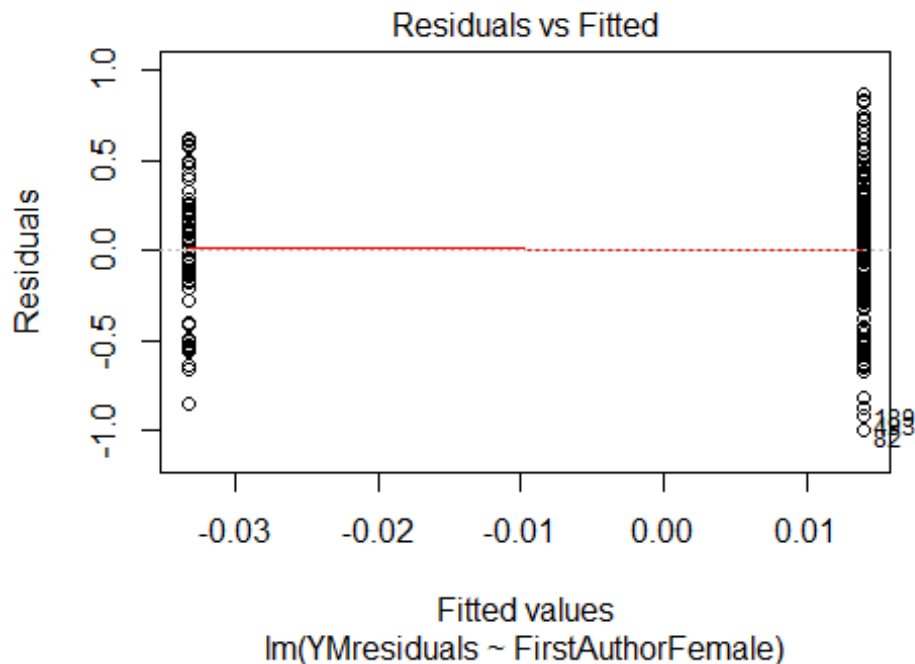
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.82, df = 1, p-value = 0.4

## [1] "Female first author team size 2018 geometric mean: 3.64112840605216"
## [1] "Male first author team size 2018 geometric mean: 6.04380027117127"

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

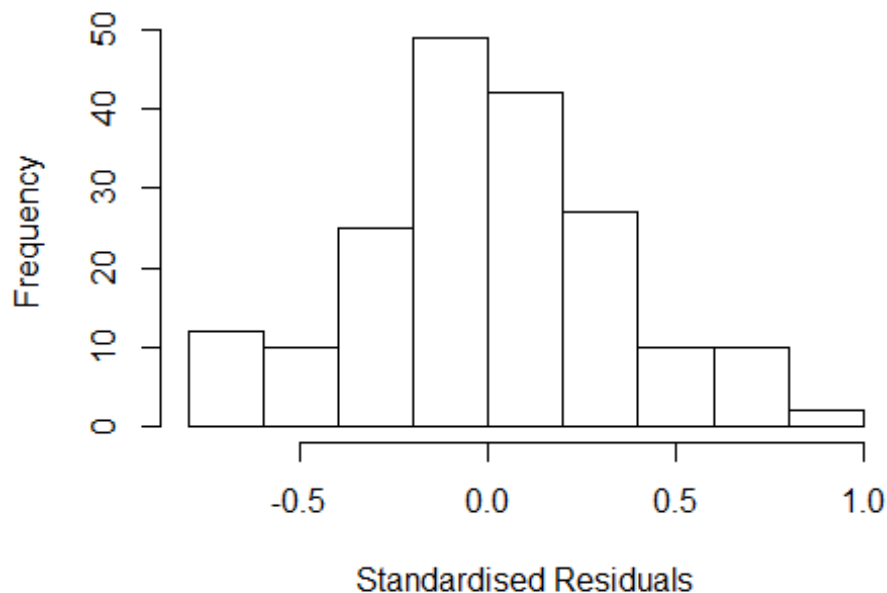
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7.5, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.82028452835046"
## [1] "Male last author team size 2018 geometric mean: 4.63677548024262"

## Warning in wilcox.test.default(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 = 0.7
## 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 1.591 1      1.261
## LastAuthorFemale  1.619 1      1.272
## UniqueAuthors    7.960 4      1.296
## Year             13.252 16     1.084
```

Residuals from first and last author and team size



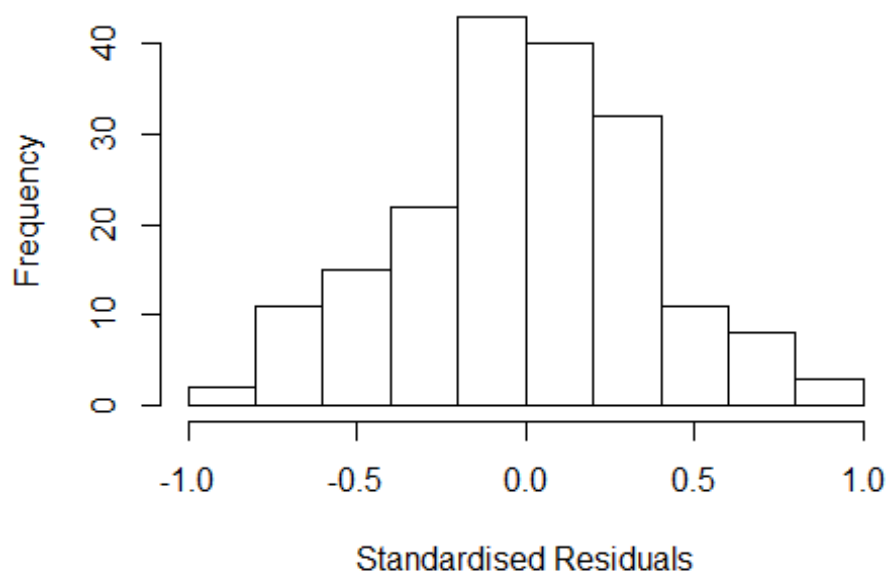
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.79393 -0.20469 -0.00736 0.22374 0.91632
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9248 0.2850 3.25 0.0014 **
## FirstAuthorFemale1 -0.0508 0.0595 -0.85 0.3947
## LastAuthorFemale1 -0.0299 0.0590 -0.51 0.6137
## UniqueAuthors2 -0.0945 0.2395 -0.39 0.6938
## UniqueAuthors3 -0.0220 0.2285 -0.10 0.9235
## UniqueAuthors4 -0.1546 0.2320 -0.67 0.5060
## UniqueAuthors5 0.1180 0.2196 0.54 0.5917
## Year1997 0.0991 0.3765 0.26 0.7927
## Year1998 0.4003 0.2055 1.95 0.0530 .
## Year1999 0.1084 0.1959 0.55 0.5810
```

```

## Year2000          0.2222      0.2056      1.08      0.2814
## Year2001          0.2333      0.5538      0.42      0.6741
## Year2002          0.0690      0.2754      0.25      0.8026
## Year2003          0.1708      0.2305      0.74      0.4597
## Year2004          0.2146      0.2065      1.04      0.3002
## Year2005          0.2254      0.2610      0.86      0.3891
## Year2006          0.1876      0.2020      0.93      0.3543
## Year2007          0.3662      0.1814      2.02      0.0452 *
## Year2008          0.1330      0.1794      0.74      0.4597
## Year2009          0.1382      0.1961      0.70      0.4819
## Year2010          0.1799      0.1852      0.97      0.3329
## Year2011          0.0318      0.2026      0.16      0.8753
## Year2012          0.1965      0.1986      0.99      0.3238
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.335
## Multiple R-squared:  0.146, Adjusted R-squared:  0.032
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 169 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.434  0.871  0.949  0.898  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 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.372 1      1.171
## LastAuthorFemale  1.516 1      1.231
## Year              2.018 16      1.022

```


Residuals from first and last author



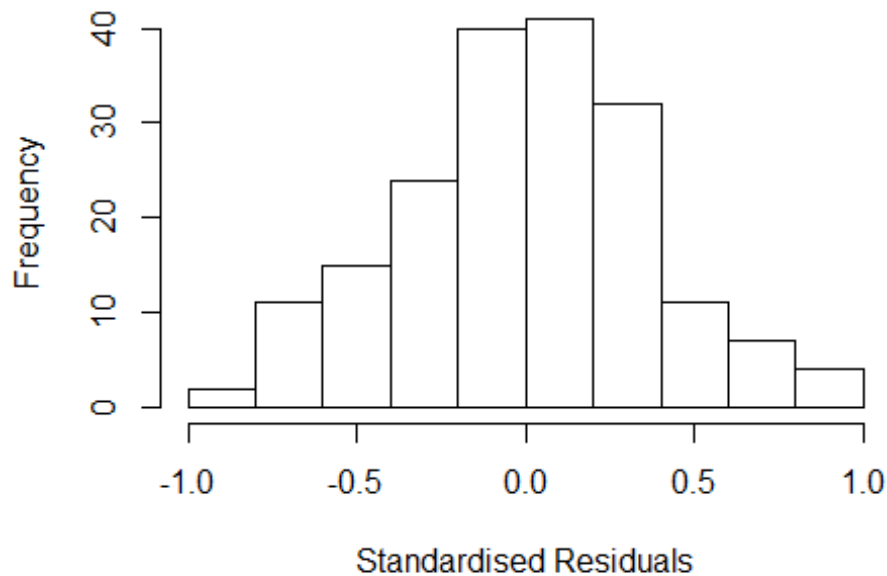
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.90888 -0.22930 0.00522 0.23050 0.86206
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8768 0.1799 4.87 2.5e-06 ***
## FirstAuthorFemale1 -0.0551 0.0601 -0.92 0.361
## LastAuthorFemale1 -0.0581 0.0595 -0.98 0.330
## Year1997 0.0682 0.3995 0.17 0.865
## Year1998 0.4259 0.2124 2.00 0.047 *
## Year1999 0.1720 0.1989 0.86 0.388
## Year2000 0.1654 0.2158 0.77 0.445
## Year2001 0.2651 0.4083 0.65 0.517
## Year2002 0.1287 0.3690 0.35 0.728
## Year2003 0.1895 0.2317 0.82 0.415
## Year2004 0.2558 0.2231 1.15 0.253
## Year2005 0.2632 0.2715 0.97 0.334
```

```

## Year2006          0.2520      0.2090      1.21      0.230
## Year2007          0.4042      0.1928      2.10      0.038 *
## Year2008          0.1919      0.1925      1.00      0.320
## Year2009          0.2443      0.1997      1.22      0.223
## Year2010          0.2248      0.2012      1.12      0.265
## Year2011          0.0897      0.2140      0.42      0.676
## Year2012          0.3021      0.2062      1.47      0.145
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.353
## Multiple R-squared:  0.0693, Adjusted R-squared:  -0.0304
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 168 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.488  0.853   0.955   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      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 1.388 1      1.178
## Year              1.388 16      1.010

```

Residuals from first author



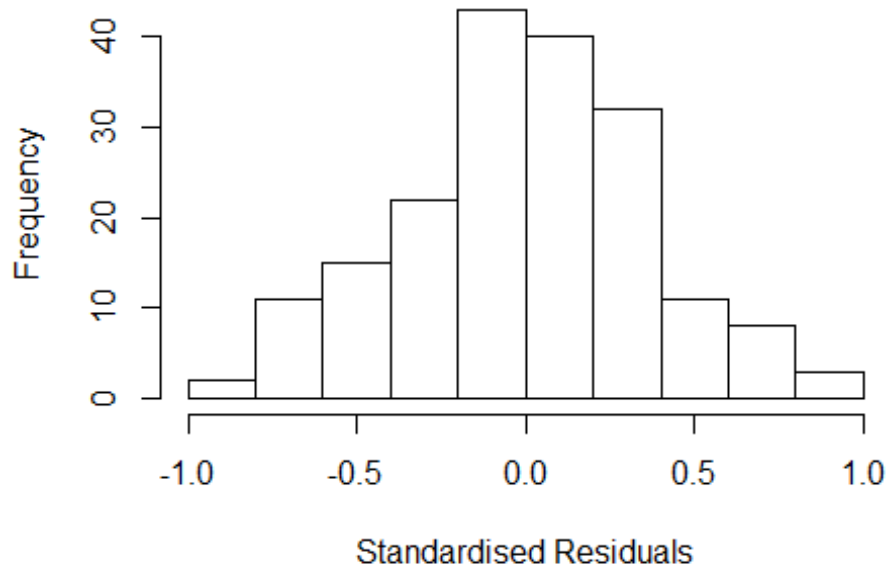
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.90024 -0.22104  0.00114  0.23942  0.86102
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.8759    0.1799   4.87 2.6e-06 ***
## FirstAuthorFemale1 -0.0589    0.0607  -0.97  0.333
## Year1997        0.0646    0.3999   0.16  0.872
## Year1998        0.4135    0.2087   1.98  0.049 *
## Year1999        0.1678    0.1993   0.84  0.401
## Year2000        0.1682    0.2158   0.78  0.437
## Year2001        0.2652    0.4116   0.64  0.520
## Year2002        0.1296    0.3739   0.35  0.729
## Year2003        0.1915    0.2319   0.83  0.410
## Year2004        0.2370    0.2202   1.08  0.283
## Year2005        0.2651    0.2734   0.97  0.334
## Year2006        0.2358    0.2089   1.13  0.261
```

```

## Year2007          0.3925      0.1927      2.04      0.043 *
## Year2008          0.1691      0.1896      0.89      0.374
## Year2009          0.2357      0.2008      1.17      0.242
## Year2010          0.2090      0.2000      1.05      0.297
## Year2011          0.0636      0.2068      0.31      0.759
## Year2012          0.2944      0.2052      1.43      0.153
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.349
## Multiple R-squared:  0.0669, Adjusted R-squared:  -0.027
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 168 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.485  0.837   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      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 1.536 1      1.239
## Year              1.536 16      1.013

```

Residuals from last author



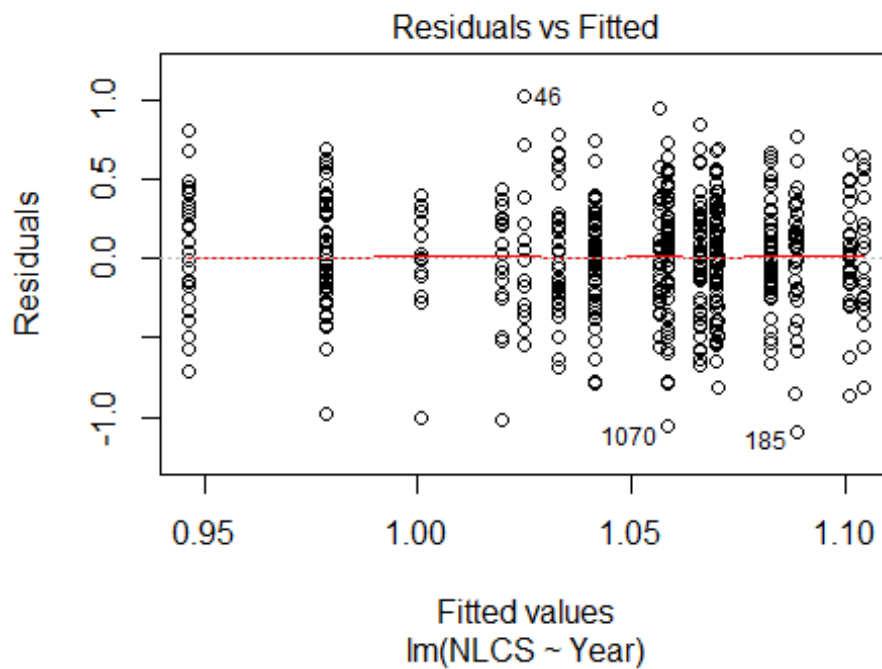
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9037 -0.2127 0.0085 0.2244 0.8852
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8661 0.1874 4.62 7.5e-06 ***
## LastAuthorFemale1 -0.0632 0.0601 -1.05 0.295
## Year1997 0.0738 0.4035 0.18 0.855
## Year1998 0.4204 0.2177 1.93 0.055 .
## Year1999 0.1723 0.2055 0.84 0.403
## Year2000 0.1504 0.2230 0.67 0.501
## Year2001 0.2748 0.4155 0.66 0.509
## Year2002 0.1394 0.3749 0.37 0.710
## Year2003 0.2014 0.2378 0.85 0.398
## Year2004 0.2582 0.2297 1.12 0.263
## Year2005 0.2507 0.2813 0.89 0.374
## Year2006 0.2374 0.2143 1.11 0.270
```

```

## Year2007          0.3931      0.2000      1.97      0.051 .
## Year2008          0.1784      0.1993      0.90      0.372
## Year2009          0.2369      0.2086      1.14      0.258
## Year2010          0.2204      0.2080      1.06      0.291
## Year2011          0.0800      0.2172      0.37      0.713
## Year2012          0.3076      0.2122      1.45      0.149
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.348
## Multiple R-squared:  0.0672, Adjusted R-squared:  -0.0267
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 163 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.480  0.841   0.947   0.893   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      5.35e-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 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
##   25   26   23   27   30   40   32   44   44   40   50   61   72   75   85
## 2011 2012
##   85   79
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   19   17   15   21   14   25   15   32   29   32   37   50   67   59   69
## 2011 2012

```

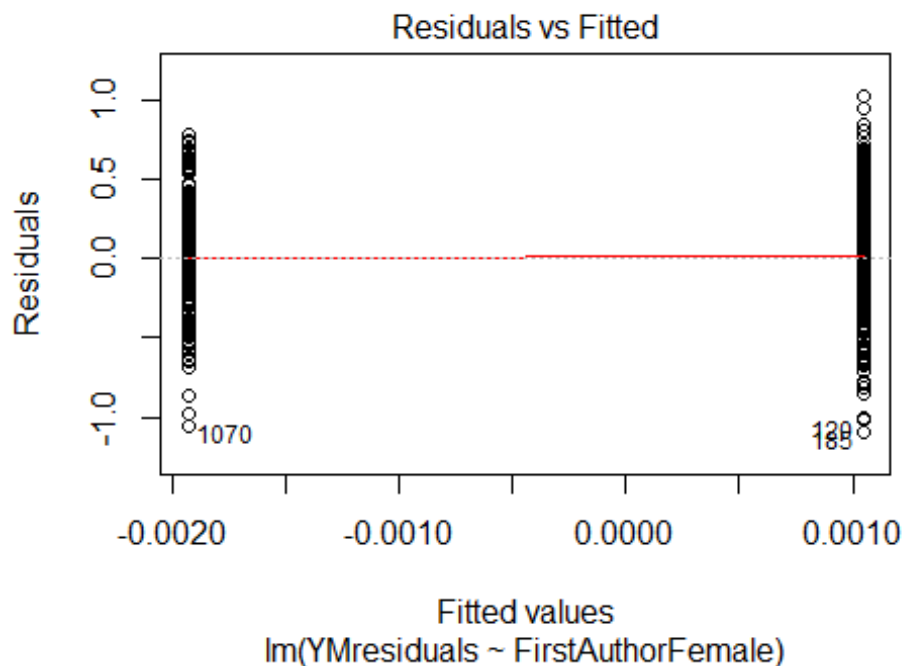
```
## 72 56
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 15 15 13 16 12 22 14 27 28 28 34 44 57 55 62
## 2011 2012
## 66 48
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 17, df = 16, p-value = 0.4
```



```
##
## 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: 4.67948669187545"
## [1] "Male first author team size 2018 geometric mean: 3.3147249799804"
## Warning in wilcox.test.default(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 = 0.08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.37951913988789"
## [1] "Male last author team size 2018 geometric mean: 3.58030867271064"

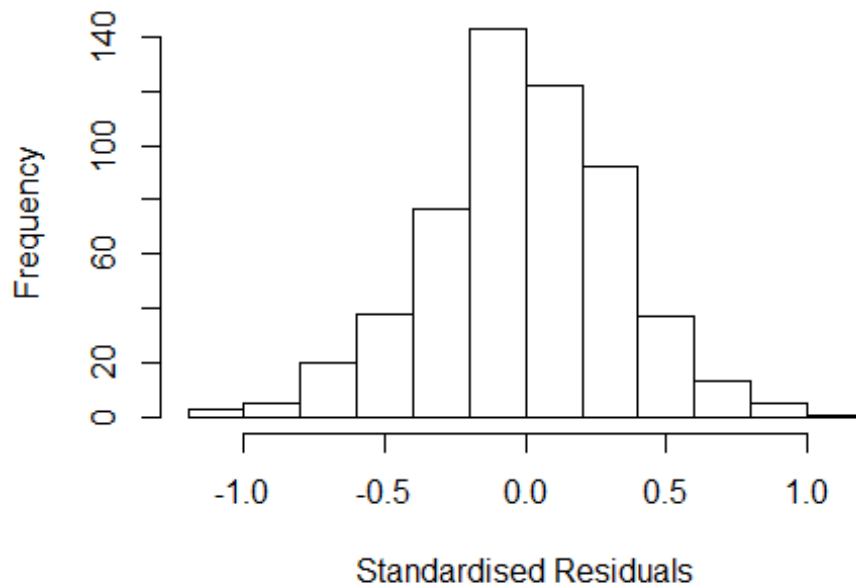
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 30, p-value = 0.5
## 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 | 1.239 | 1 | 1.113 |
| LastAuthorFemale | 1.176 | 1 | 1.085 |
| UniqueAuthors | 2.028 | 4 | 1.092 |
| Year | 2.629 | 16 | 1.031 |

Residuals from first and last author and team size



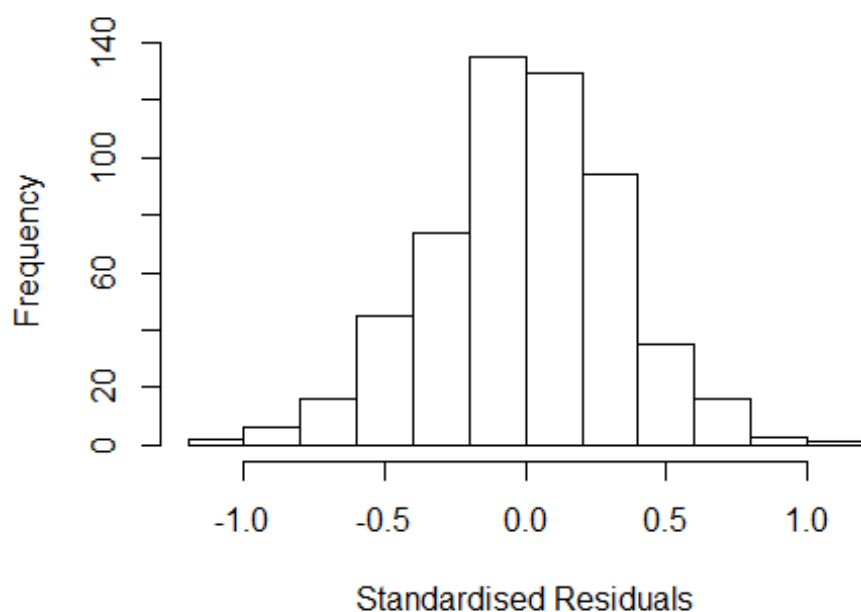
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0817 -0.2067 -0.0104 0.2099 1.1070
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.92054 0.20461 4.50 8.4e-06 ***
## FirstAuthorFemale1 -0.02794 0.02947 -0.95 0.34
## LastAuthorFemale1 -0.00984 0.03536 -0.28 0.78
## UniqueAuthors2 0.15060 0.15106 1.00 0.32
## UniqueAuthors3 0.14955 0.15315 0.98 0.33
## UniqueAuthors4 0.12051 0.15311 0.79 0.43
## UniqueAuthors5 0.23931 0.15316 1.56 0.12
## Year1997 -0.10208 0.16703 -0.61 0.54
## Year1998 0.02397 0.16161 0.15 0.88
## Year1999 0.03793 0.14250 0.27 0.79
```

```

## Year2000      0.00358    0.14947    0.02    0.98
## Year2001      0.04070    0.14522    0.28    0.78
## Year2002      0.02964    0.14538    0.20    0.84
## Year2003     -0.09911    0.15307   -0.65    0.52
## Year2004      0.03127    0.13701    0.23    0.82
## Year2005      0.02539    0.14655    0.17    0.86
## Year2006     -0.04776    0.13924   -0.34    0.73
## Year2007     -0.01682    0.13208   -0.13    0.90
## Year2008      0.00169    0.12894    0.01    0.99
## Year2009     -0.02913    0.12865   -0.23    0.82
## Year2010     -0.01865    0.12753   -0.15    0.88
## Year2011     -0.10514    0.12701   -0.83    0.41
## Year2012     -0.02190    0.14061   -0.16    0.88
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.31
## Multiple R-squared:  0.0432, Adjusted R-squared:  0.00373
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 48 weights are ~= 1. The remaining 508 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.174  0.865  0.949  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      1.80e-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 1.172 1      1.083
## LastAuthorFemale  1.142 1      1.069
## Year              1.328 16      1.009

```

Residuals from first and last author



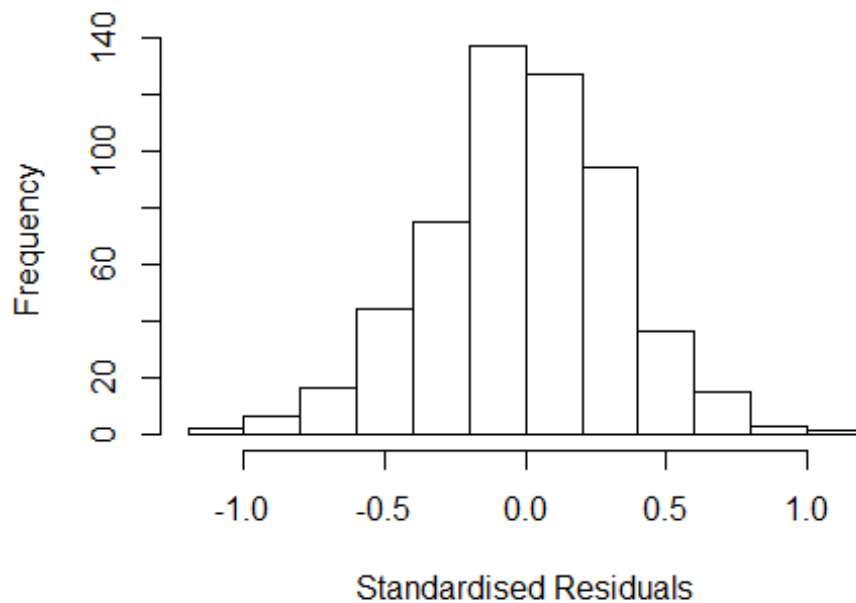
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.12896 -0.21225 -0.00106  0.21206  1.08318
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.082821   0.117588   9.21   <2e-16 ***
## FirstAuthorFemale1 -0.019740   0.029571  -0.67    0.50
## LastAuthorFemale1 -0.010669   0.036114  -0.30    0.77
## Year1997        -0.120000   0.160267  -0.75    0.45
## Year1998         0.024650   0.160570   0.15    0.88
## Year1999         0.016903   0.136175   0.12    0.90
## Year2000        -0.031963   0.142288  -0.22    0.82
## Year2001         0.046143   0.138252   0.33    0.74
## Year2002         0.022294   0.140996   0.16    0.87
## Year2003        -0.117494   0.148642  -0.79    0.43
## Year2004         0.040783   0.132166   0.31    0.76
## Year2005        -0.004287   0.136078  -0.03    0.97
```

```

## Year2006      -0.044013    0.137675   -0.32    0.75
## Year2007      -0.022608    0.127723   -0.18    0.86
## Year2008       0.000227    0.123731    0.00    1.00
## Year2009      -0.026202    0.124433   -0.21    0.83
## Year2010      -0.010550    0.123125   -0.09    0.93
## Year2011      -0.086345    0.123303   -0.70    0.48
## Year2012      -0.002530    0.133529   -0.02    0.98
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.315
## Multiple R-squared:  0.0179, Adjusted R-squared:  -0.015
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 45 weights are ~= 1. The remaining 511 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.171  0.865   0.947   0.896   0.987   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      1.80e-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 1.164 1      1.079
## Year              1.164 16      1.005

```

Residuals from first author

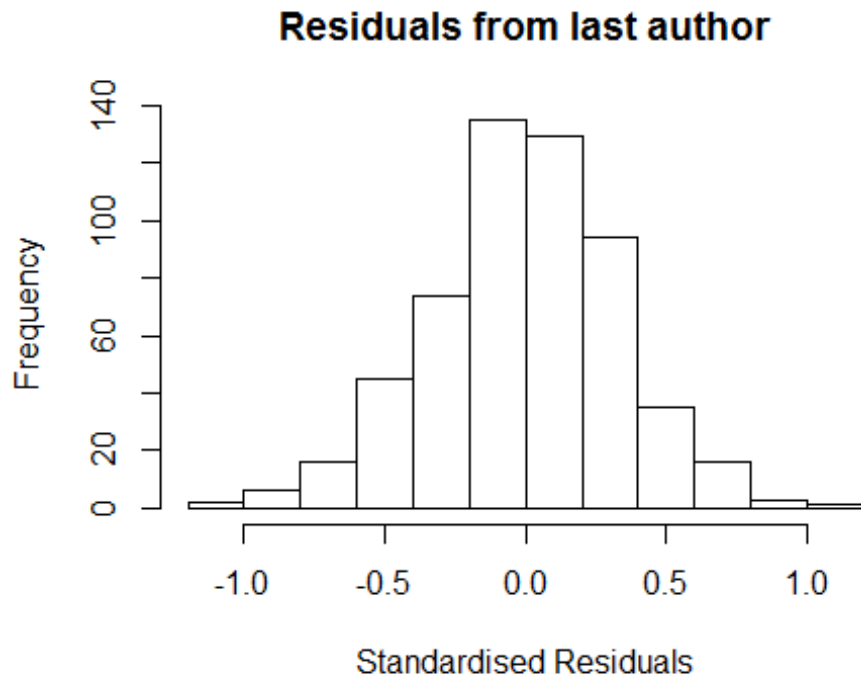


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.12800 -0.21255 -0.00357 0.21427 1.08542
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.08066 0.11673 9.26 <2e-16 ***
## FirstAuthorFemale1 -0.02052 0.02949 -0.70 0.49
## Year1997 -0.12007 0.16019 -0.75 0.45
## Year1998 0.02456 0.16031 0.15 0.88
## Year1999 0.01732 0.13644 0.13 0.90
## Year2000 -0.03014 0.14147 -0.21 0.83
## Year2001 0.04734 0.13805 0.34 0.73
## Year2002 0.02301 0.14085 0.16 0.87
## Year2003 -0.11625 0.14848 -0.78 0.43
## Year2004 0.04260 0.13136 0.32 0.75
## Year2005 -0.00342 0.13600 -0.03 0.98
## Year2006 -0.04435 0.13783 -0.32 0.75
```

```

## Year2007      -0.02163    0.12741   -0.17    0.87
## Year2008      0.00108    0.12351    0.01    0.99
## Year2009     -0.02482    0.12394   -0.20    0.84
## Year2010     -0.01165    0.12363   -0.09    0.92
## Year2011     -0.08773    0.12369   -0.71    0.48
## Year2012     -0.00279    0.13369   -0.02    0.98
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.314
## Multiple R-squared:  0.0178, Adjusted R-squared:  -0.0132
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 44 weights are ~= 1. The remaining 512 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.169  0.865  0.947  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      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 1.129 1          1.063
## 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
## -1.12425 -0.21217 -0.00187 0.21259 1.08653
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.08126 0.11739 9.21 <2e-16 ***
## LastAuthorFemale1 -0.01270 0.03599 -0.35 0.72
## Year1997 -0.12179 0.16069 -0.76 0.45
## Year1998 0.01851 0.16113 0.11 0.91
## Year1999 0.00942 0.13462 0.07 0.94
## Year2000 -0.03407 0.14169 -0.24 0.81
## Year2001 0.04299 0.13751 0.31 0.75
## Year2002 0.01865 0.14057 0.13 0.89
## Year2003 -0.12296 0.14789 -0.83 0.41
## Year2004 0.03332 0.13109 0.25 0.80
## Year2005 -0.00920 0.13473 -0.07 0.95
## Year2006 -0.04842 0.13743 -0.35 0.72
```

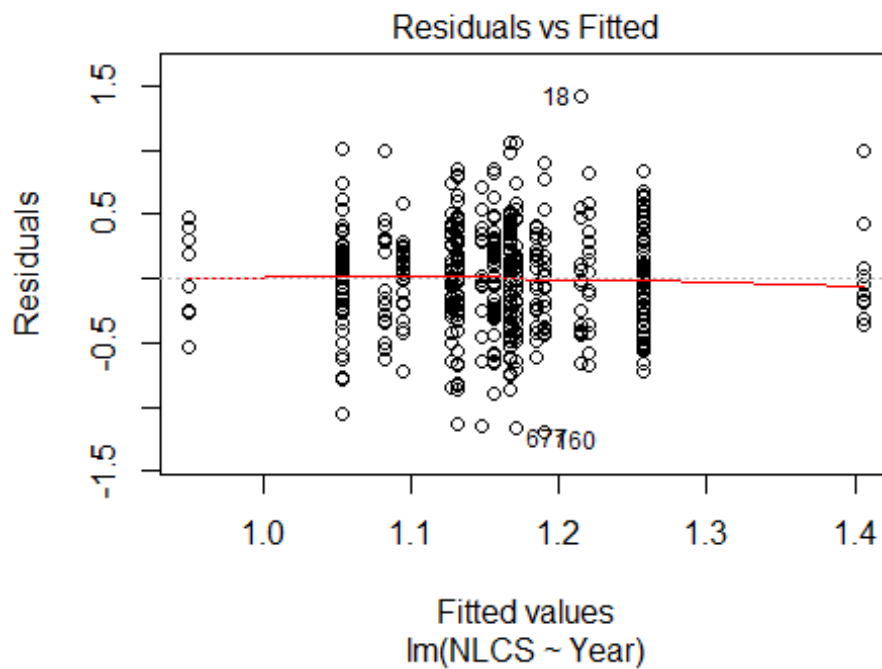
```

## Year2007          -0.02959      0.12656    -0.23      0.82
## Year2008          -0.00960      0.12161    -0.08      0.94
## Year2009          -0.03209      0.12337    -0.26      0.79
## Year2010          -0.01568      0.12255    -0.13      0.90
## Year2011          -0.09083      0.12274    -0.74      0.46
## Year2012          -0.00559      0.13335    -0.04      0.97
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.314
## Multiple R-squared:  0.0172, Adjusted R-squared:  -0.0139
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 43 weights are ~= 1. The remaining 513 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.173  0.865  0.948  0.896  0.987  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.80e-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 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
##   24   21   20   35   32   26   30   40   38   55   51   48   64   77   85
## 2011 2012
##   87   82
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   13    9   11   22   18   15   16   26   23   42   34   35   47   58   55
## 2011 2012

```



```
## 63 58
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 13 8 11 20 18 13 14 23 21 36 33 33 42 50 52
## 2011 2012
## 57 45
## [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 = 0.03
```



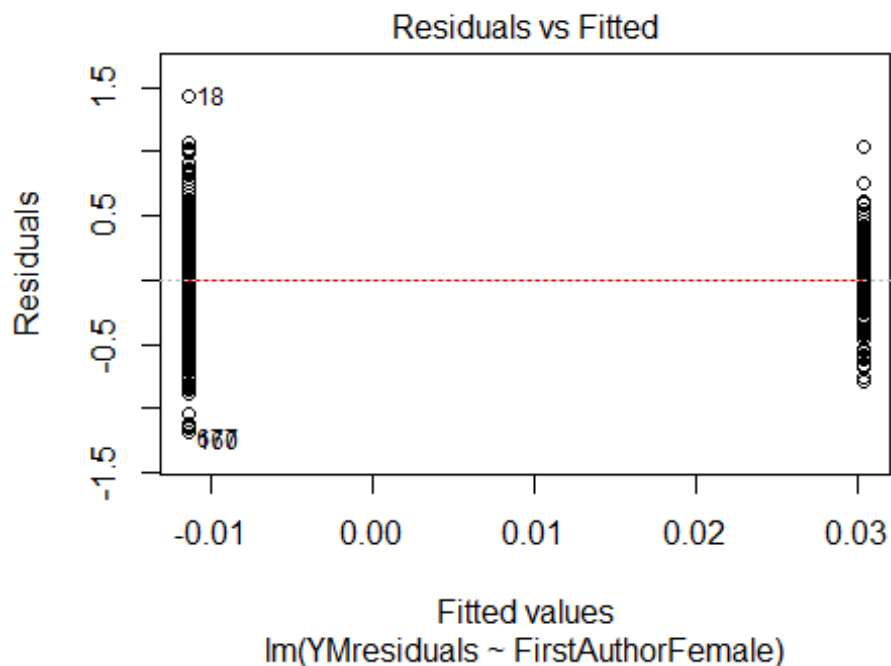
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 11, df = 1, p-value = 0.001

## [1] "Female first author team size 2018 geometric mean: 4.61789292741588"
## [1] "Male first author team size 2018 geometric mean: 3.72335777023232"

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

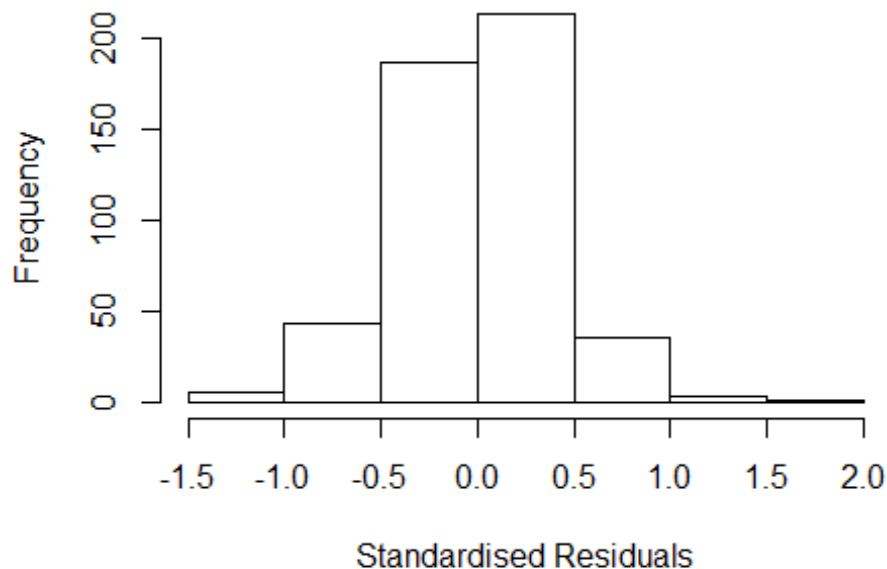
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 57, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.4920019719995"
## [1] "Male last author team size 2018 geometric mean: 3.44237271690962"

## Warning in wilcox.test.default(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 = 0.07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
## Year as factors"
##
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.144 1      1.070
## LastAuthorFemale  1.336 1      1.156
## UniqueAuthors    2.120 4      1.098
## Year              2.634 16     1.031
```

Residuals from first and last author and team size



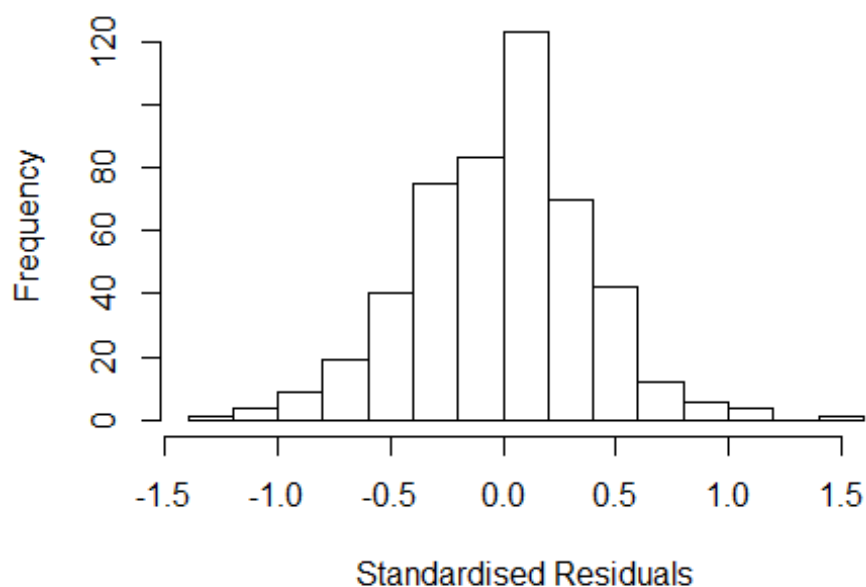
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3162 -0.2503 0.0216 0.2471 1.5418
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.19602 0.14924 8.01 9e-15 ***
## FirstAuthorFemale1 0.03066 0.03697 0.83 0.407
## LastAuthorFemale1 -0.06729 0.04200 -1.60 0.110
## UniqueAuthors2 -0.10485 0.09993 -1.05 0.295
## UniqueAuthors3 -0.13677 0.09664 -1.42 0.158
## UniqueAuthors4 -0.20377 0.09894 -2.06 0.040 *
## UniqueAuthors5 -0.07140 0.09565 -0.75 0.456
## Year1997 -0.11113 0.19387 -0.57 0.567
## Year1998 0.29065 0.15035 1.93 0.054 .
## Year1999 0.14570 0.15311 0.95 0.342
```

```

## Year2000      -0.00826    0.16251   -0.05    0.959
## Year2001      0.16542    0.16618    1.00    0.320
## Year2002      0.11887    0.16448    0.72    0.470
## Year2003      0.02844    0.13522    0.21    0.834
## Year2004      0.14598    0.13553    1.08    0.282
## Year2005      0.09033    0.13930    0.65    0.517
## Year2006      0.16975    0.13674    1.24    0.215
## Year2007      0.12023    0.14343    0.84    0.402
## Year2008      0.05612    0.12737    0.44    0.660
## Year2009      0.15117    0.12868    1.17    0.241
## Year2010      0.02343    0.13336    0.18    0.861
## Year2011      0.06114    0.13215    0.46    0.644
## Year2012      0.05032    0.14622    0.34    0.731
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.358
## Multiple R-squared:  0.0575, Adjusted R-squared:  0.013
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 32 weights are ~= 1. The remaining 457 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0236 0.8690 0.9490 0.8980 0.9850 0.9990
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          2.04e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.102 1 1.050
## LastAuthorFemale 1.266 1 1.125
## Year 1.385 16 1.010

```

Residuals from first and last author



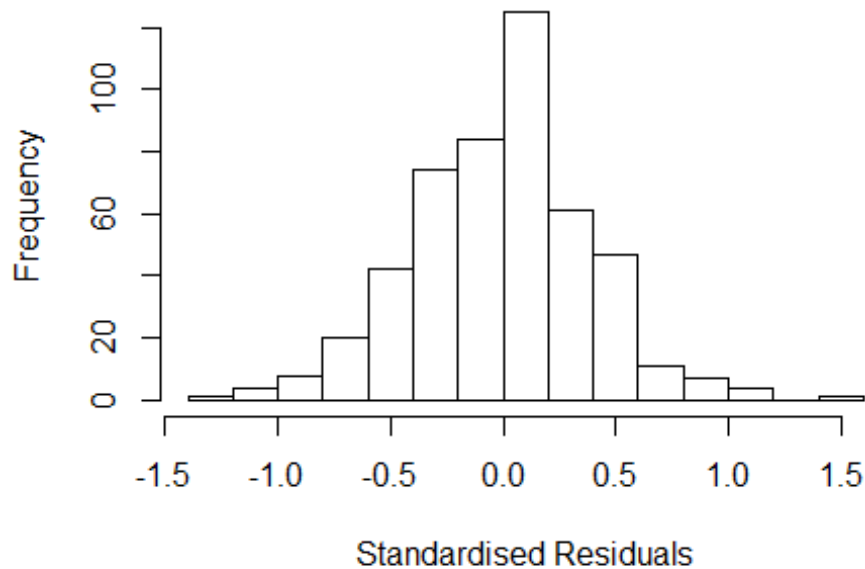
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2402 -0.2490  0.0323  0.2326  1.5563
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0767     0.1219   8.83  <2e-16 ***
## FirstAuthorFemale1  0.0253     0.0362   0.70   0.485
## LastAuthorFemale1 -0.0753     0.0418  -1.80   0.072 .
## Year1997          -0.1181     0.1889  -0.63   0.532
## Year1998           0.2907     0.1538   1.89   0.059 .
## Year1999           0.1502     0.1501   1.00   0.317
## Year2000          -0.0254     0.1658  -0.15   0.878
## Year2001           0.1636     0.1716   0.95   0.341
## Year2002           0.1327     0.1650   0.80   0.422
## Year2003           0.0307     0.1364   0.22   0.822
## Year2004           0.1450     0.1370   1.06   0.290
## Year2005           0.0968     0.1406   0.69   0.491
```

```

## Year2006          0.1697      0.1371      1.24      0.216
## Year2007          0.1193      0.1466      0.81      0.416
## Year2008          0.0582      0.1293      0.45      0.653
## Year2009          0.1627      0.1304      1.25      0.213
## Year2010          0.0197      0.1349      0.15      0.884
## Year2011          0.0743      0.1338      0.56      0.579
## Year2012          0.0496      0.1477      0.34      0.737
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.361
## Multiple R-squared:  0.0389, Adjusted R-squared:  0.00209
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 34 weights are ~= 1. The remaining 455 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0235 0.8660 0.9520 0.8980 0.9850 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      2.04e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.101 1      1.049
## Year              1.101 16      1.003

```

Residuals from first author



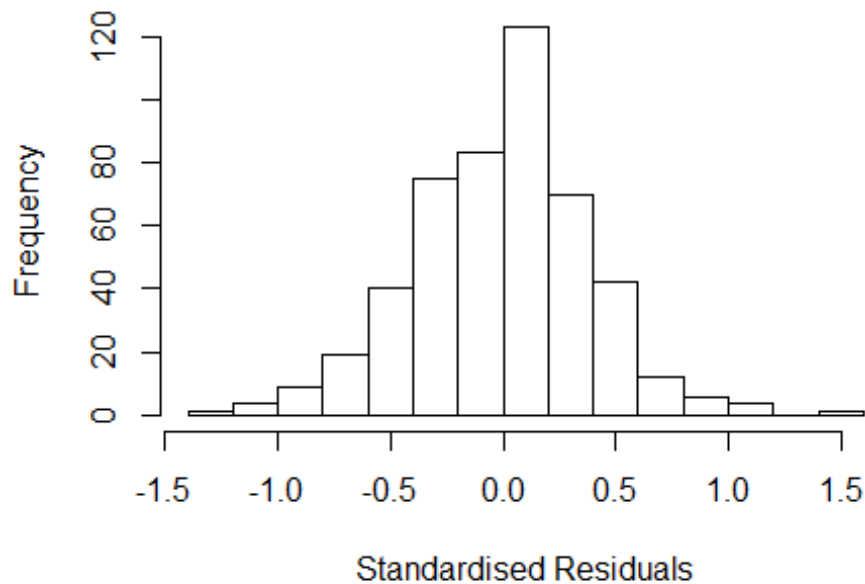
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.231 -0.250 0.027 0.228 1.555
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0782 0.1222 8.82 <2e-16 ***
## FirstAuthorFemale1 0.0218 0.0365 0.60 0.550
## Year1997 -0.1197 0.1888 -0.63 0.527
## Year1998 0.2572 0.1493 1.72 0.086 .
## Year1999 0.1494 0.1502 0.99 0.320
## Year2000 -0.0300 0.1643 -0.18 0.855
## Year2001 0.1532 0.1726 0.89 0.375
## Year2002 0.1317 0.1652 0.80 0.426
## Year2003 0.0273 0.1373 0.20 0.843
## Year2004 0.1344 0.1381 0.97 0.331
## Year2005 0.0904 0.1404 0.64 0.520
## Year2006 0.1556 0.1372 1.13 0.258
```

```

## Year2007          0.1036      0.1462      0.71      0.479
## Year2008          0.0428      0.1293      0.33      0.741
## Year2009          0.1542      0.1305      1.18      0.238
## Year2010         -0.0031      0.1338     -0.02      0.982
## Year2011          0.0523      0.1328      0.39      0.694
## Year2012          0.0292      0.1464      0.20      0.842
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.363
## Multiple R-squared:  0.0344, Adjusted R-squared:  -0.000416
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 38 weights are ~= 1. The remaining 451 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.027  0.864   0.952   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      2.04e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.257 1          1.121
## Year            1.257 16          1.007

```


Residuals from last author



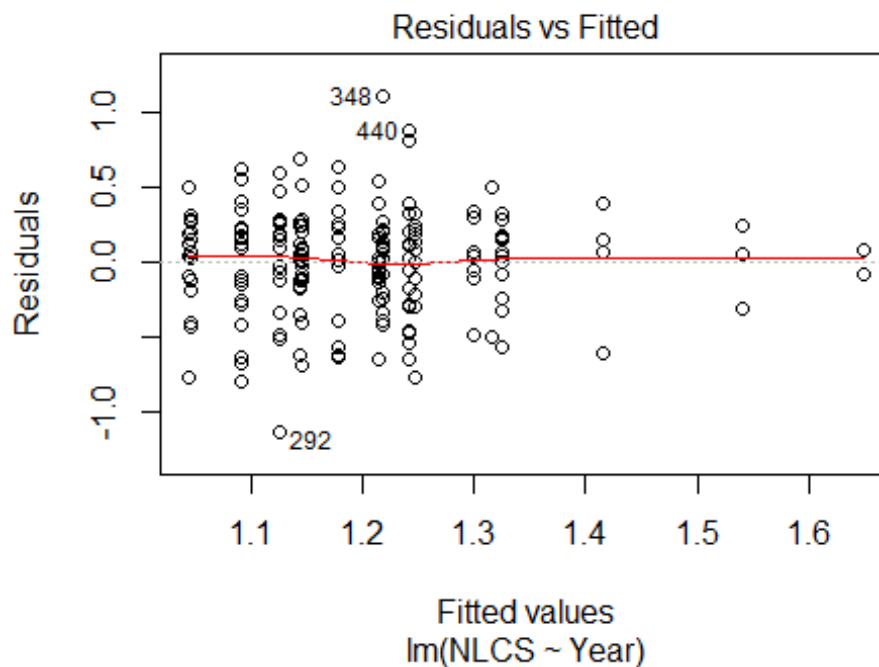
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2465 -0.2518 0.0285 0.2384 1.5514
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0816 0.1209 8.95 <2e-16 ***
## LastAuthorFemale1 -0.0737 0.0423 -1.74 0.082 .
## Year1997 -0.1230 0.1900 -0.65 0.518
## Year1998 0.2923 0.1554 1.88 0.061 .
## Year1999 0.1479 0.1493 0.99 0.322
## Year2000 -0.0236 0.1662 -0.14 0.887
## Year2001 0.1650 0.1705 0.97 0.334
## Year2002 0.1319 0.1649 0.80 0.424
## Year2003 0.0330 0.1360 0.24 0.808
## Year2004 0.1479 0.1372 1.08 0.281
## Year2005 0.0987 0.1403 0.70 0.482
## Year2006 0.1714 0.1367 1.25 0.210
```

```

## Year2007          0.1219      0.1465      0.83      0.406
## Year2008          0.0638      0.1287      0.50      0.621
## Year2009          0.1660      0.1299      1.28      0.202
## Year2010          0.0196      0.1343      0.15      0.884
## Year2011          0.0782      0.1334      0.59      0.558
## Year2012          0.0513      0.1478      0.35      0.729
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.36
## Multiple R-squared:  0.0382, Adjusted R-squared:  0.00352
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 32 weights are ~= 1. The remaining 457 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0233 0.8650 0.9530 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      2.04e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 489"
## [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
##   15   10    5   16    5   10   18   21   13   13   17   19   15   21   31
## 2011 2012
##   26   29
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    3    2   11    2    4   14   17   12    8   13   14   12   17   18
## 2011 2012

```

```
## 22 18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 6 3 2 11 2 4 13 17 12 8 12 13 12 16 15
## 2011 2012
## 19 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 = 14, df = 16, p-value = 0.6
```



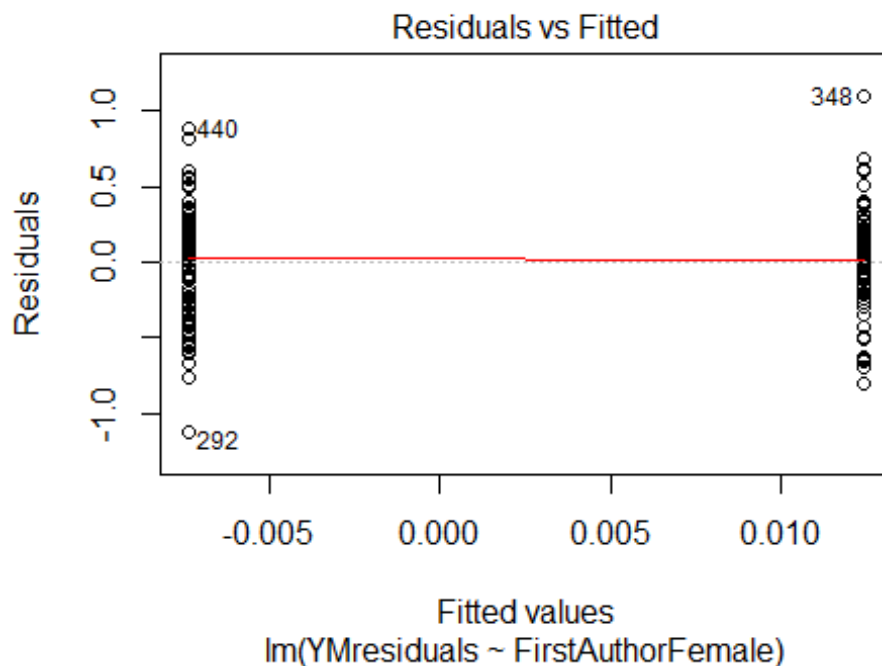
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.2, df = 1, p-value = 0.7

## [1] "Female first author team size 2018 geometric mean: 3.93597934253086"
## [1] "Male first author team size 2018 geometric mean: 2.56947031424688"

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

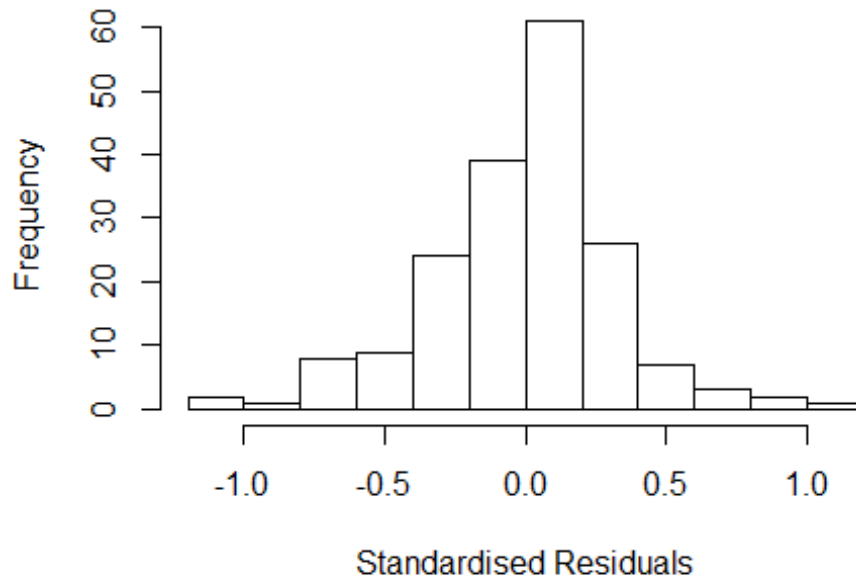
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 14, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.51784835276224"
## [1] "Male last author team size 2018 geometric mean: 2.32998610150143"

## Warning in wilcox.test.default(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 = 0.2
## 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 1.553 1      1.246
## LastAuthorFemale  1.817 1      1.348
## UniqueAuthors    6.647 4      1.267
## Year             10.174 16     1.075
```

Residuals from first and last author and team size



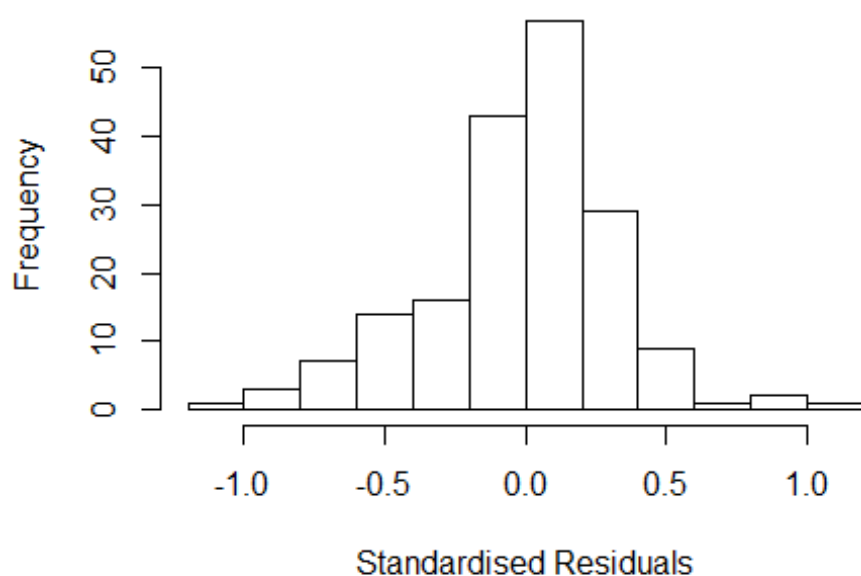
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1530 -0.1801 0.0287 0.1740 1.1832
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.31781 0.21000 6.28 3.1e-09 ***
## FirstAuthorFemale1 0.01982 0.05345 0.37 0.711
## LastAuthorFemale1 -0.08001 0.14215 -0.56 0.574
## UniqueAuthors2 -0.04222 0.16334 -0.26 0.796
## UniqueAuthors3 -0.18482 0.15100 -1.22 0.223
## UniqueAuthors4 -0.19527 0.15088 -1.29 0.197
## UniqueAuthors5 -0.19240 0.14617 -1.32 0.190
## Year1997 0.31562 0.26412 1.19 0.234
## Year1998 0.04990 0.64574 0.08 0.938
## Year1999 -0.06267 0.18652 -0.34 0.737
```

```

## Year2000          0.44471      0.20766      2.14      0.034 *
## Year2001          0.28375      0.27540      1.03      0.304
## Year2002          0.03453      0.20774      0.17      0.868
## Year2003         -0.03199      0.17358     -0.18      0.854
## Year2004         -0.11448      0.17331     -0.66      0.510
## Year2005          0.15873      0.17777      0.89      0.373
## Year2006          0.10603      0.16867      0.63      0.530
## Year2007          0.15156      0.15743      0.96      0.337
## Year2008          0.20781      0.16756      1.24      0.217
## Year2009          0.03049      0.16632      0.18      0.855
## Year2010         -0.00438      0.16274     -0.03      0.979
## Year2011         -0.01644      0.16970     -0.10      0.923
## Year2012          0.03291      0.19128      0.17      0.864
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.292
## Multiple R-squared:  0.147, Adjusted R-squared:  0.0299
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 167 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0626 0.8760 0.9580 0.8860 0.9860 0.9990
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          5.46e-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 1.488 1          1.220
## LastAuthorFemale  1.348 1          1.161
## Year              1.950 16          1.021

```

Residuals from first and last author



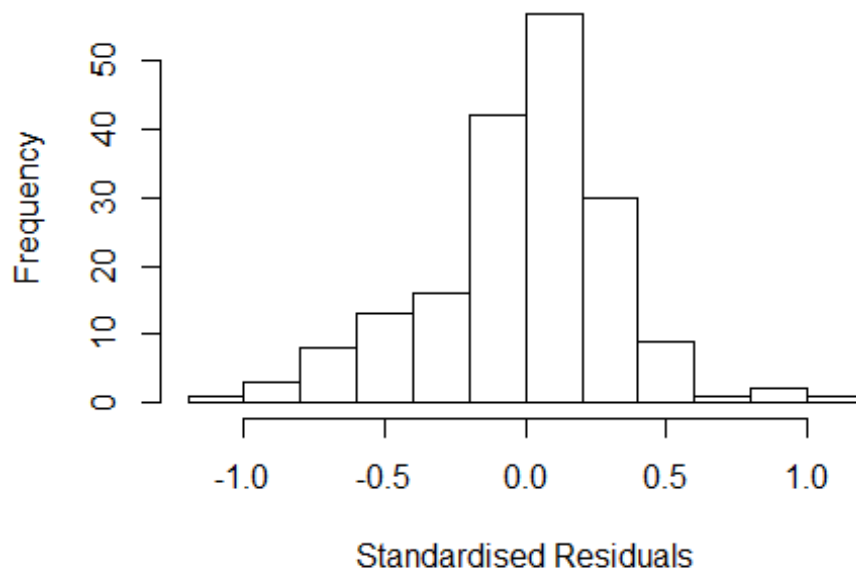
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1732 -0.1703 0.0269 0.1894 1.1752
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1058 0.1546 7.15 2.7e-11 ***
## FirstAuthorFemale1 0.0137 0.0527 0.26 0.7958
## LastAuthorFemale1 -0.0374 0.1347 -0.28 0.7814
## Year1997 0.4410 0.2078 2.12 0.0353 *
## Year1998 0.2226 0.8230 0.27 0.7871
## Year1999 0.0292 0.1793 0.16 0.8709
## Year2000 0.5432 0.1659 3.27 0.0013 **
## Year2001 0.3630 0.2435 1.49 0.1379
## Year2002 0.1060 0.2101 0.50 0.6146
## Year2003 0.0456 0.1555 0.29 0.7697
## Year2004 -0.0519 0.1676 -0.31 0.7573
## Year2005 0.1979 0.1755 1.13 0.2612
```

```

## Year2006          0.1482      0.1625      0.91      0.3633
## Year2007          0.1930      0.1525      1.27      0.2073
## Year2008          0.2297      0.1633      1.41      0.1614
## Year2009          0.0674      0.1684      0.40      0.6893
## Year2010          0.0294      0.1610      0.18      0.8555
## Year2011          0.0144      0.1732      0.08      0.9336
## Year2012          0.0849      0.1795      0.47      0.6368
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.294
## Multiple R-squared:  0.118, Adjusted R-squared:  0.0213
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 165 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.075  0.864   0.955   0.885   0.985   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      5.46e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.452 1      1.205
## Year              1.452 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
## -1.1725 -0.1703 0.0271 0.1915 1.1799
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.10940 0.15344 7.23 1.7e-11 ***
## FirstAuthorFemale1 0.01074 0.05220 0.21 0.8373
## Year1997 0.43745 0.20701 2.11 0.0361 *
## Year1998 0.20173 0.97624 0.21 0.8365
## Year1999 0.02192 0.17708 0.12 0.9017
## Year2000 0.53960 0.16487 3.27 0.0013 **
## Year2001 0.36151 0.24327 1.49 0.1392
## Year2002 0.09651 0.20236 0.48 0.6340
## Year2003 0.04296 0.15441 0.28 0.7812
## Year2004 -0.05950 0.16626 -0.36 0.7209
## Year2005 0.19498 0.17454 1.12 0.2656
## Year2006 0.14322 0.16108 0.89 0.3752
```

```

## Year2007          0.19154    0.15137    1.27    0.2075
## Year2008          0.22715    0.16230    1.40    0.1635
## Year2009          0.06312    0.16621    0.38    0.7046
## Year2010          0.02396    0.16024    0.15    0.8813
## Year2011          0.00462    0.16778    0.03    0.9780
## Year2012          0.08117    0.17998    0.45    0.6526
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.292
## Multiple R-squared:  0.118, Adjusted R-squared:  0.027
## Convergence in 30 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 166 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0653 0.8660 0.9560 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      5.46e-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 1.326 1      1.152
## Year              1.326 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
## -1.1796 -0.1702  0.0261  0.1829  1.1842

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.11764    0.13831    8.08 1.3e-13 ***
## LastAuthorFemale1 -0.03362    0.13321   -0.25 0.80108
## Year1997        0.42923    0.19623    2.19 0.03012 *
## Year1998        0.21567    0.80620    0.27 0.78941
## Year1999        0.02127    0.17106    0.12 0.90118
## Year2000        0.53136    0.15089    3.52 0.00055 ***
## Year2001        0.35624    0.23958    1.49 0.13894
## Year2002        0.10107    0.20514    0.49 0.62288
## Year2003        0.04052    0.14946    0.27 0.78665
## Year2004       -0.06087    0.15712   -0.39 0.69893
## Year2005        0.18884    0.16378    1.15 0.25059
## Year2006        0.14065    0.15449    0.91 0.36393
## Year2007        0.18915    0.14787    1.28 0.20263
## Year2008        0.22195    0.15524    1.43 0.15468
## Year2009        0.06196    0.16222    0.38 0.70300
## Year2010        0.02216    0.15426    0.14 0.88595
## Year2011        0.00782    0.16472    0.05 0.96221
## Year2012        0.07706    0.17339    0.44 0.65733
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.292
## Multiple R-squared:  0.119, Adjusted R-squared:  0.0279
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 165 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0621 0.8560 0.9560 0.8830 0.9840 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      5.46e-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: 183"
## [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]"
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    4    8    3    5    3    9    3    3   11   18   12   13   33   21   17
## 2012
##   13
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    3    6    2    4    2    4    0    1    9   14    9    9   26   16   14
## 2012
##   11
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    3    6    2    4    2    3    0    1    7   10    7    8   22   15   12
## 2012
##   10
## [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: 1.5157165665104"

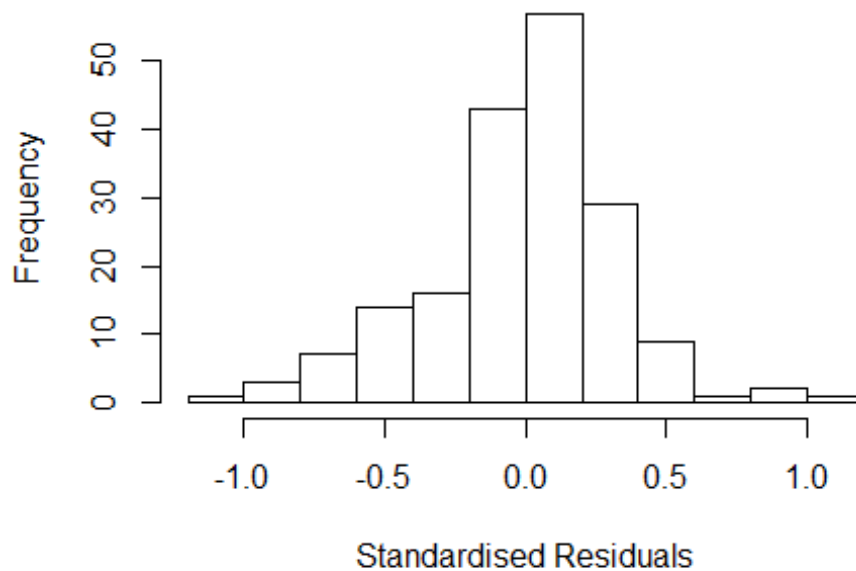
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 17, p-value = 0.08
## 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: 1.73851050644476"

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

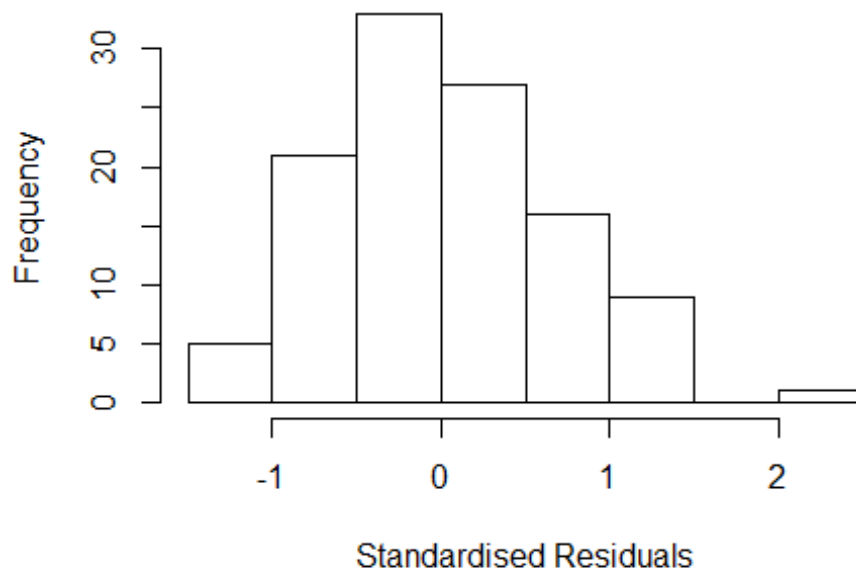
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 10, p-value = 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 2.665 1      1.632
## LastAuthorFemale 3.291 1      1.814
## UniqueAuthors    17.396 4      1.429
## Year              36.121 14     1.137
```

Residuals from first and last author and team size



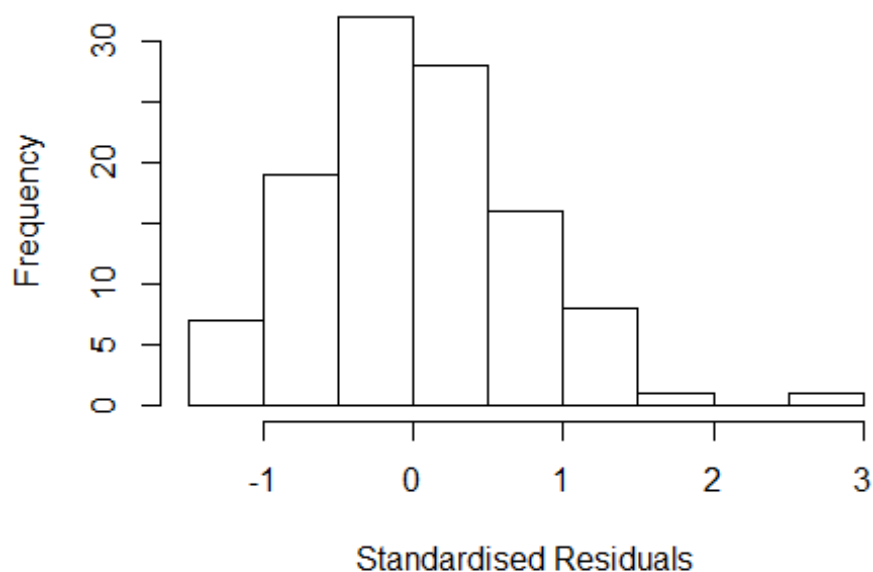
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3846 -0.4490 -0.0475 0.4663 2.4226
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.173402 0.197587 5.94 5.2e-08 ***
## FirstAuthorFemale1 0.591670 0.177303 3.34 0.00123 **
## LastAuthorFemale1 -0.092286 0.190301 -0.48 0.62888
## UniqueAuthors2 0.135951 0.192172 0.71 0.48110
## UniqueAuthors3 0.336607 0.201552 1.67 0.09834 .
## UniqueAuthors4 0.539915 0.243236 2.22 0.02892 *
## UniqueAuthors5 0.371114 0.270337 1.37 0.17319
## Year1998 0.075231 0.424296 0.18 0.85966
## Year1999 -0.039877 0.532143 -0.07 0.94043
## Year2000 -0.249401 0.527061 -0.47 0.63721
```

```

## Year2001      -0.989934    0.355667    -2.78    0.00654 **
## Year2002      -0.927684    0.814515    -1.14    0.25772
## Year2004      -0.841402    0.197587    -4.26    5.0e-05 ***
## Year2005      -0.415903    0.303174    -1.37    0.17349
## Year2006      -0.111341    0.300030    -0.37    0.71143
## Year2007      -0.855919    0.242091    -3.54    0.00064 ***
## Year2008      -0.255516    0.339712    -0.75    0.45390
## Year2009      -0.756669    0.238301    -3.18    0.00204 **
## Year2010      -0.033127    0.291681    -0.11    0.90983
## Year2011      -0.373169    0.302465    -1.23    0.22047
## Year2012       0.000383    0.370822     0.00    0.99918
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.706
## Multiple R-squared:  0.283, Adjusted R-squared:  0.126
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 98 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.215  0.893   0.947   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      8.93e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.489 1      1.578
## LastAuthorFemale  3.312 1      1.820
## Year              4.174 14      1.052

```

Residuals from first and last author



```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 225 84861860507 3.933 2012      1700      1      2.598
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2943 -0.4987 -0.0266  0.4452  2.5983
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.28867    0.14522   8.87 4.2e-14 ***
## FirstAuthorFemale1 0.63601    0.17564   3.62 0.00047 ***
## LastAuthorFemale1 -0.08039    0.18589  -0.43 0.66640
## Year1998         0.00560    0.40848   0.01 0.98908
## Year1999        -0.08717    0.57013  -0.15 0.87880
## Year2000        -0.18371    0.48277  -0.38 0.70440
## Year2001        -0.85167    0.35974  -2.37 0.01994 *
## Year2002        -0.78992    0.57943  -1.36 0.17602
## Year2004        -0.95667    0.14522  -6.59 2.5e-09 ***
## Year2005        -0.44244    0.29315  -1.51 0.13454
## Year2006        -0.02699    0.26832  -0.10 0.92009
## Year2007        -0.73479    0.22071  -3.33 0.00124 **
```

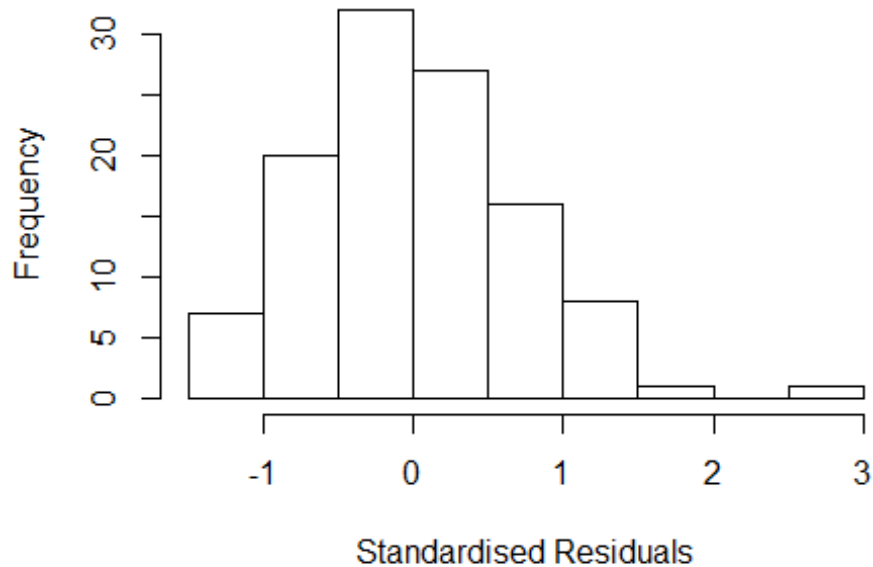


```

## Year2008          -0.20883    0.31325   -0.67  0.50659
## Year2009          -0.72939    0.20324   -3.59  0.00053 ***
## Year2010          -0.00504    0.27812   -0.02  0.98557
## Year2011          -0.26420    0.28313   -0.93  0.35310
## Year2012           0.04599    0.31471    0.15  0.88411
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.76
## Multiple R-squared:  0.242, Adjusted R-squared:  0.114
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 102 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.219  0.886  0.954  0.923  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      8.93e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##  nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##  trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.072 1          1.439
## Year              2.072 14          1.026

```

Residuals from first author



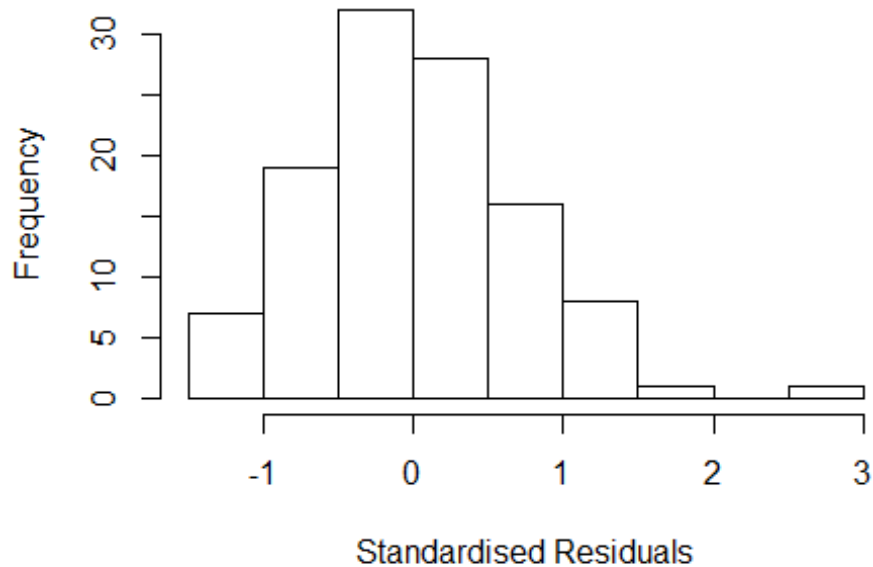
```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 225 84861860507 3.933 2012    1700      1    2.598
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2943 -0.4987 -0.0308  0.4623  2.5984
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.28867    0.14522   8.87 3.9e-14 ***
## FirstAuthorFemale1 0.59945    0.16258   3.69 0.00038 ***
## Year1998        0.00561    0.40840   0.01 0.98906
## Year1999       -0.08717    0.57014  -0.15 0.87880
## Year2000       -0.17117    0.47839  -0.36 0.72127
## Year2001       -0.85167    0.35974  -2.37 0.01992 *
## Year2002       -0.78994    0.57934  -1.36 0.17591
## Year2004       -0.95667    0.14522  -6.59 2.4e-09 ***
## Year2005       -0.45605    0.29170  -1.56 0.12125
## Year2006       -0.02256    0.26845  -0.08 0.93320
## Year2007       -0.74707    0.21869  -3.42 0.00093 ***
## Year2008       -0.20726    0.30943  -0.67 0.50460
```

```

## Year2009          -0.73507      0.20200      -3.64  0.00044 ***
## Year2010          -0.03209      0.25269      -0.13  0.89921
## Year2011          -0.26910      0.28048      -0.96  0.33975
## Year2012           0.04597      0.31456       0.15  0.88413
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.76
## Multiple R-squared:  0.241, Adjusted R-squared:  0.122
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 101 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.219  0.887   0.954   0.922   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      8.93e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 2.34 1          1.530
## Year             2.34 14          1.031

```

Residuals from last author



```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 225 84861860507 3.933 2012    1700      1    2.598
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3124 -0.5003 -0.0138  0.4706  2.5700
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.28876    0.14493   8.89 3.6e-14 ***
## LastAuthorFemale1 0.27366    0.19220   1.42 0.15774
## Year1998      -0.00299    0.40026  -0.01 0.99406
## Year1999      -0.08726    0.55812  -0.16 0.87609
## Year2000       0.02363    0.48911   0.05 0.96157
## Year2001      -0.85176    0.35706  -2.39 0.01902 *
## Year2002      -0.77650    0.56818  -1.37 0.17493
## Year2004      -0.95676    0.14493  -6.60 2.2e-09 ***
## Year2005      -0.35502    0.28150  -1.26 0.21030
## Year2006       0.04806    0.28824   0.17 0.86793
## Year2007      -0.78752    0.22994  -3.42 0.00091 ***
## Year2008      -0.08532    0.32634  -0.26 0.79431
```

```

## Year2009          -0.58560      0.20799      -2.82   0.00591 **
## Year2010           0.01099      0.30473       0.04   0.97130
## Year2011          -0.02403      0.26977      -0.09   0.92919
## Year2012           0.07422      0.32116       0.23   0.81772
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.815
## Multiple R-squared:  0.172, Adjusted R-squared:  0.0424
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 102 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.299  0.888  0.962  0.927  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      8.93e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 112"
## [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]"
##
## 2005 2006 2007 2008 2009 2010 2011 2012
##    1    5    1    3    1    4    2    4
##
## 2005 2006 2007 2008 2009 2010 2011 2012
##    1    5    1    2    0    4    2    3
##
## 2005 2006 2007 2008 2009 2010 2011 2012
##    1    5    1    1    0    4    1    2
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"

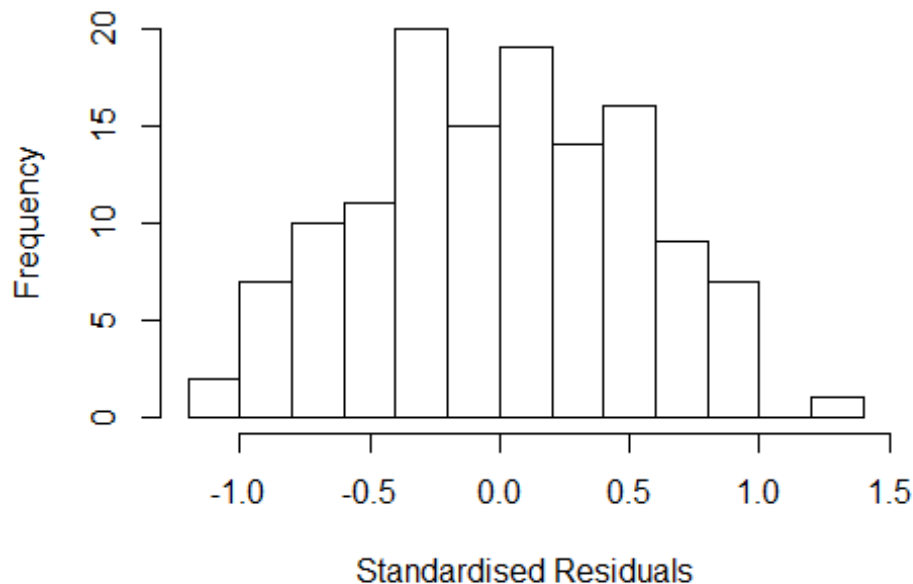
```

```

## [1] "Female first author team size 2018 geometric mean: NaN"
## [1] "Male first author team size 2018 geometric mean: 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: 15"
## [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
##    3    1    7    6    7    9   11    3   10   12   14   24   20   24   20
## 2011 2012
##   16    8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    1    7    5    5    1    6    3    9   10    9   19   16   20   13
## 2011 2012
##   14    7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    1    6    4    5    1    5    3    9    8    9   17   13   18   11
## 2011 2012
##   12    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: 1.81712059283214"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -5.823e+14  1          NaN
## LastAuthorFemale  4.399e+00  1          2.097
## UniqueAuthors    -2.303e+15  4          NaN
## Year              -1.559e+16 16          NaN

```

Residuals from first and last author and team size



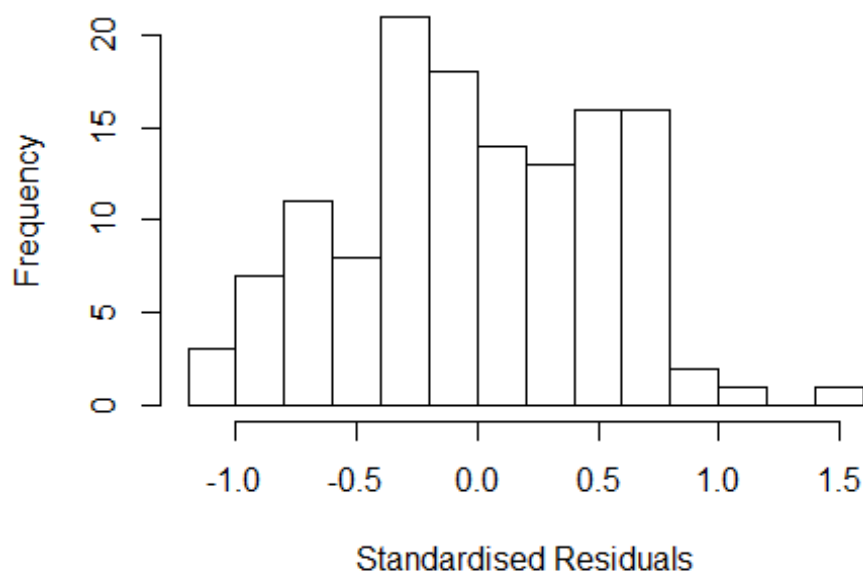
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.0693 -0.3760  0.0132  0.3980  1.3226
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.69235    0.37876    1.83   0.070 .
## FirstAuthorFemale1 -0.24739    0.15373   -1.61   0.110
## LastAuthorFemale1  0.35890    0.18582    1.93   0.056 .
## UniqueAuthors2     0.14765    0.13625    1.08   0.281
## UniqueAuthors3    -0.00219    0.15906   -0.01   0.989
## UniqueAuthors4     0.30546    0.19895    1.54   0.128
## UniqueAuthors5     0.21258    0.32928    0.65   0.520
## Year1997          -0.18535    0.37876   -0.49   0.626
## Year1998          -0.02459    0.49418   -0.05   0.960
## Year1999           0.15100    0.45922    0.33   0.743
```

```

## Year2000          1.06479      0.43630      2.44      0.016 *
## Year2001          0.36538      0.42752      0.85      0.395
## Year2002          0.39625      0.43428      0.91      0.364
## Year2003         -0.06120      0.50415     -0.12      0.904
## Year2004          0.51811      0.40991      1.26      0.209
## Year2005          0.33162      0.47466      0.70      0.486
## Year2006          0.49288      0.43871      1.12      0.264
## Year2007          0.23288      0.41334      0.56      0.574
## Year2008         -0.09516      0.41629     -0.23      0.820
## Year2009          0.22932      0.42170      0.54      0.588
## Year2010          0.22961      0.42791      0.54      0.593
## Year2011          0.15303      0.43706      0.35      0.727
## Year2012         -0.02302      0.45706     -0.05      0.960
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.592
## Multiple R-squared:  0.219, Adjusted R-squared:  0.0601
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 119 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.597  0.890  0.951  0.927  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      7.63e-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 1.298e+13  1      3.602e+06
## LastAuthorFemale  2.248e+00  1      1.499e+00
## Year              2.569e+13 16      2.625e+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
## -1.059 -0.361 -0.010 0.418 1.420
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7321 0.4312 1.70 0.092 .
## FirstAuthorFemale1 -0.2677 0.1557 -1.72 0.088 .
## LastAuthorFemale1 0.3312 0.1836 1.80 0.074 .
## Year1997 -0.2251 0.4312 -0.52 0.603
## Year1998 -0.0139 0.5313 -0.03 0.979
## Year1999 0.1470 0.5056 0.29 0.772
## Year2000 1.0890 0.4650 2.34 0.021 *
## Year2001 0.4936 0.4587 1.08 0.284
## Year2002 0.3600 0.4779 0.75 0.453
## Year2003 0.1256 0.5746 0.22 0.827
## Year2004 0.6126 0.4540 1.35 0.180
## Year2005 0.4038 0.5056 0.80 0.426
```

```

## Year2006          0.6434      0.4647      1.38      0.169
## Year2007          0.3098      0.4486      0.69      0.491
## Year2008         -0.0671      0.4575     -0.15      0.884
## Year2009          0.2859      0.4563      0.63      0.532
## Year2010          0.3024      0.4639      0.65      0.516
## Year2011          0.1964      0.4763      0.41      0.681
## Year2012          0.0387      0.4876      0.08      0.937
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.615
## Multiple R-squared:  0.197, Adjusted R-squared:  0.0677
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 119 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.573  0.890  0.956  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      7.63e-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 -1.159e+15  1      NaN
## Year              -1.159e+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

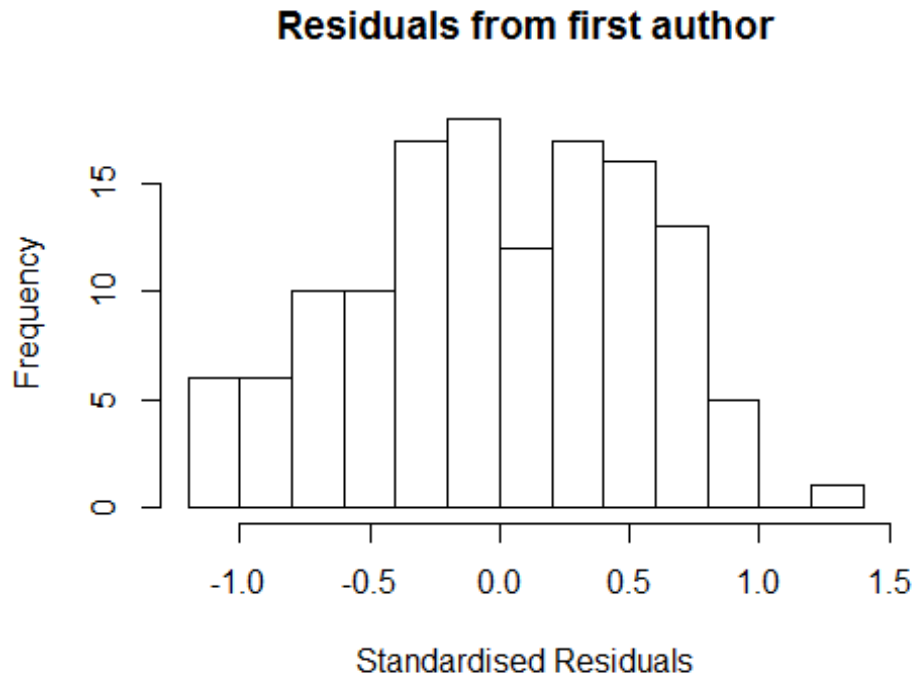
```

```

## -1.06e+00 -3.68e-01 -1.28e-15 4.22e-01 1.31e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.7320    0.4329   1.69   0.094 .
## FirstAuthorFemale1 -0.1702    0.1664  -1.02   0.308
## Year1997        -0.2250    0.4329  -0.52   0.604
## Year1998         0.0970    0.5478   0.18   0.860
## Year1999         0.1474    0.5091   0.29   0.773
## Year2000         1.0890    0.4655   2.34   0.021 *
## Year2001         0.3963    0.4638   0.85   0.395
## Year2002         0.3402    0.4841   0.70   0.484
## Year2003         0.1261    0.5784   0.22   0.828
## Year2004         0.6414    0.4577   1.40   0.164
## Year2005         0.4704    0.5129   0.92   0.361
## Year2006         0.6309    0.4680   1.35   0.180
## Year2007         0.2884    0.4504   0.64   0.523
## Year2008        -0.0283    0.4557  -0.06   0.951
## Year2009         0.3297    0.4557   0.72   0.471
## Year2010         0.3027    0.4669   0.65   0.518
## Year2011         0.2703    0.4820   0.56   0.576
## Year2012         0.0845    0.4811   0.18   0.861
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.612
## Multiple R-squared:  0.175, Adjusted R-squared:  0.0508
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 123 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.626 0.895 0.957 0.930 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      7.63e-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"

```

```
## Warning in lmrob.S(x, y, control = control, mf = mf): S refinements did
not
## converge (to refine.tol=1e-07) in 1000 (= k.max) steps
```



```
## [1] "Sample size for the above analysis: 131"
## [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
##    2    8    7    3    2    4    5    7    9   13   17   13   18   17   15
## 2011 2012
##   11    6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    1    1    2    1    2    2    6    6    9   12   11   14   16    9
## 2011 2012
##    9    5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    1    1    2    1    2    2    5    4    6   12   10   11   16    8
## 2011 2012
##    9    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: 1.25992104989487"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.725e+02 1      13.132
## LastAuthorFemale  6.216e+00 1       2.493
## UniqueAuthors    6.169e+28 4     3969.853
## Year              5.605e+30 15     10.591

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##      Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##      k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.07877 -0.31749 -0.00561  0.33673  1.19256
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.65e+00   9.08e-08  1.82e+07 < 2e-16 ***
## FirstAuthorFemale1 -3.94e-01   1.84e-01 -2.15e+00  0.03523 *
## LastAuthorFemale1  1.06e-01   2.87e-01  3.70e-01  0.71301
## UniqueAuthors2    2.29e-01   2.31e-01  9.90e-01  0.32307
## UniqueAuthors3    4.56e-01   1.95e-01  2.34e+00  0.02203 *
## UniqueAuthors4    6.48e-01   2.37e-01  2.73e+00  0.00791 **
## UniqueAuthors5    5.38e-01   2.23e-01  2.41e+00  0.01849 *
## Year1998          -1.09e+00   2.31e-01 -4.73e+00  1.1e-05 ***
## Year1999          -1.32e+00   2.73e-01 -4.84e+00  7.3e-06 ***
## Year2000          -1.44e+00   1.95e-01 -7.40e+00  2.0e-10 ***
## Year2001          -7.53e-01   1.57e-01 -4.80e+00  8.4e-06 ***
## Year2002          -1.33e+00   2.19e-01 -6.05e+00  5.8e-08 ***
## Year2003          -9.62e-01   2.57e-01 -3.74e+00  0.00037 ***
## Year2004          -3.05e-01   1.14e-01 -2.68e+00  0.00907 **
## Year2005          -5.73e-01   2.85e-01 -2.01e+00  0.04804 *
## Year2006          -8.04e-01   1.63e-01 -4.94e+00  4.9e-06 ***
## Year2007          -9.43e-01   1.71e-01 -5.53e+00  4.8e-07 ***
## Year2008          -8.81e-01   2.68e-01 -3.29e+00  0.00157 **
## Year2009          -5.62e-01   1.95e-01 -2.88e+00  0.00522 **
## Year2010          -8.99e-01   2.38e-01 -3.78e+00  0.00032 ***
## Year2011          -9.59e-01   1.94e-01 -4.94e+00  4.9e-06 ***
```

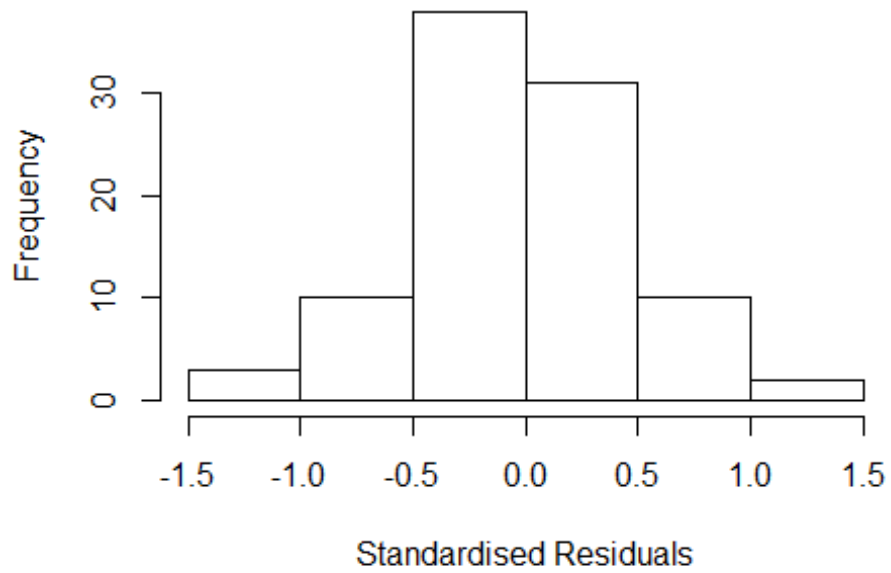
```

## Year2012          -1.03e+00  2.64e-01 -3.92e+00  0.00020 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.527
## Multiple R-squared:  0.266, Adjusted R-squared:  0.0523
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 82 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.588  0.887  0.955  0.913  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      1.06e-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"

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

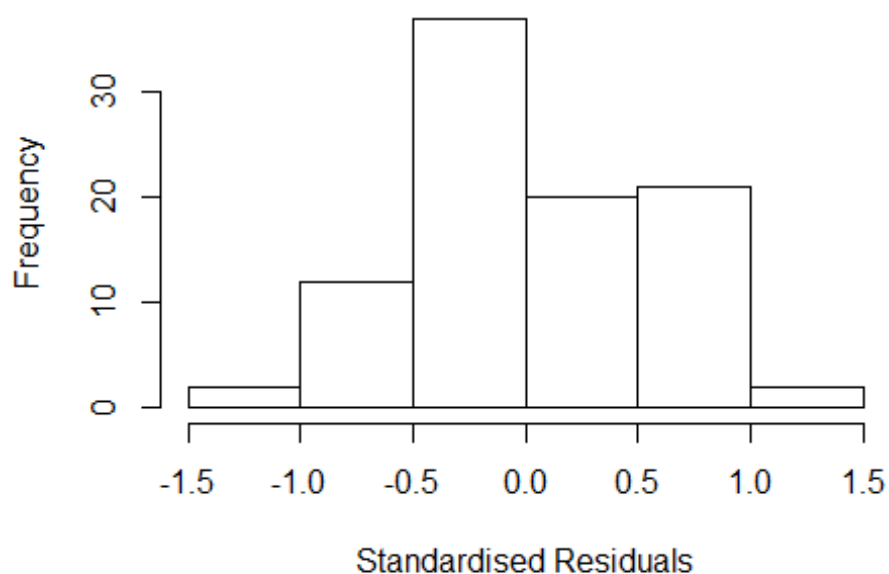
```

Residuals from first and last author and team size



| | | | | |
|----|-------------------|---------|----|---------------------|
| ## | | GVIF | Df | $GVIF^{(1/(2*Df))}$ |
| ## | FirstAuthorFemale | 7.971 | 1 | 2.823 |
| ## | LastAuthorFemale | 43.522 | 1 | 6.597 |
| ## | Year | 344.471 | 15 | 1.215 |

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
## -1.1508 -0.3826 -0.0235 0.4643 1.2133
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.65e+00 0.00e+00 Inf < 2e-16 ***
## FirstAuthorFemale1 -3.20e-01 2.04e-01 -1.57e+00 0.12113
## LastAuthorFemale1 1.48e-01 2.82e-01 5.20e-01 0.60215
## Year1998 -8.61e-01 5.74e-08 -1.50e+07 < 2e-16 ***
## Year1999 -8.08e-01 2.02e-01 -4.00e+00 0.00015 ***
## Year2000 -9.85e-01 2.24e-08 -4.39e+07 < 2e-16 ***
## Year2001 -4.48e-01 1.10e-01 -4.07e+00 0.00012 ***
## Year2002 -1.01e+00 1.43e-01 -7.05e+00 7.2e-10 ***
## Year2003 -5.37e-01 1.24e-01 -4.34e+00 4.4e-05 ***
## Year2004 -2.49e-01 6.29e-02 -3.96e+00 0.00017 ***
## Year2005 -1.76e-01 1.63e-01 -1.08e+00 0.28384
## Year2006 -6.02e-01 2.02e-01 -2.99e+00 0.00381 **
```



```

## Year2007          -7.29e-01    2.24e-01 -3.26e+00    0.00169 **
## Year2008          -6.86e-01    2.91e-01 -2.36e+00    0.02091 *
## Year2009          -3.30e-01    1.50e-01 -2.20e+00    0.03061 *
## Year2010          -6.54e-01    1.96e-01 -3.34e+00    0.00131 **
## Year2011          -5.95e-01    2.20e-01 -2.71e+00    0.00832 **
## Year2012          -8.71e-01    3.53e-01 -2.47e+00    0.01592 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.615
## Multiple R-squared:  0.149, Adjusted R-squared:  -0.0409
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 80 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.677  0.900   0.950   0.925   0.973   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          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 11.82  1      3.439
## Year              11.82 15      1.086
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.0603 -0.3968 -0.0235  0.4544  1.1915

```

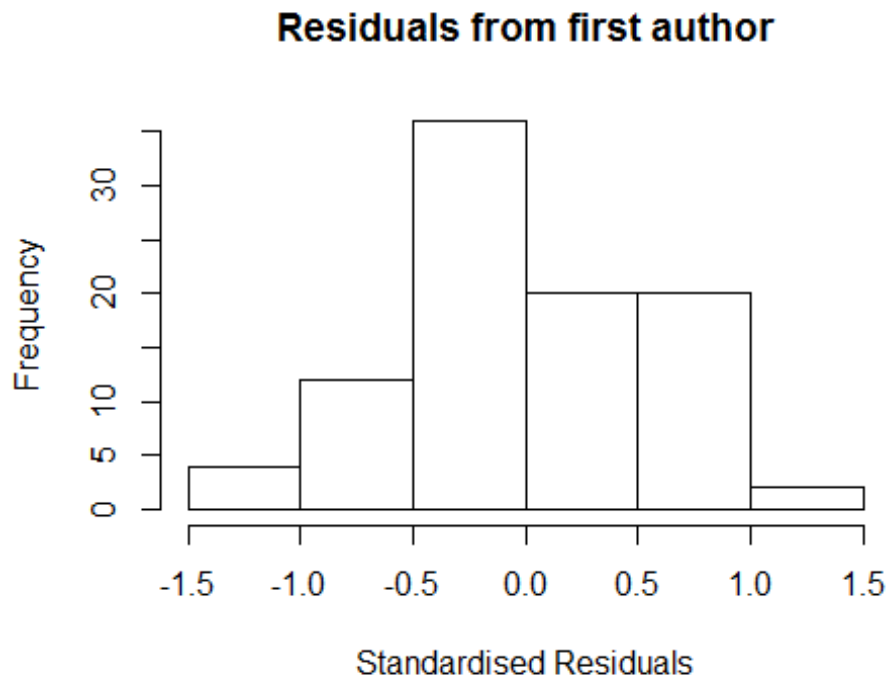
```

##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.65e+00   1.69e-08  9.81e+07 < 2e-16 ***
## FirstAuthorFemale1 -2.95e-01  2.07e-01 -1.43e+00  0.15810
## Year1998        -8.61e-01  2.77e-08 -3.11e+07 < 2e-16 ***
## Year1999        -8.20e-01  2.11e-01 -3.89e+00  0.00021 ***
## Year2000        -9.85e-01  2.51e-08 -3.93e+07 < 2e-16 ***
## Year2001        -4.60e-01  1.08e-01 -4.24e+00  6.1e-05 ***
## Year2002        -9.32e-01  7.43e-02 -1.26e+01 < 2e-16 ***
## Year2003        -5.37e-01  1.23e-01 -4.35e+00  4.1e-05 ***
## Year2004        -2.49e-01  6.29e-02 -3.96e+00  0.00017 ***
## Year2005        -1.58e-01  1.84e-01 -8.60e-01  0.39318
## Year2006        -5.93e-01  2.00e-01 -2.96e+00  0.00403 **
## Year2007        -7.13e-01  2.14e-01 -3.34e+00  0.00131 **
## Year2008        -6.42e-01  2.70e-01 -2.37e+00  0.02006 *
## Year2009        -3.14e-01  1.45e-01 -2.17e+00  0.03323 *
## Year2010        -6.44e-01  1.86e-01 -3.45e+00  0.00091 ***
## Year2011        -5.97e-01  2.17e-01 -2.75e+00  0.00745 **
## Year2012        -8.37e-01  3.73e-01 -2.24e+00  0.02799 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.643
## Multiple R-squared:  0.142, Adjusted R-squared:  -0.0359
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 83 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.712  0.905   0.957   0.933   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      1.06e-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"

```

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

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



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

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3595 -0.3783 -0.0235  0.4291  1.2326
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.6530      0.0000      Inf < 2e-16 ***
## LastAuthorFemale1  0.0960      0.3001   0.32  0.74992
```

```

## Year1998      -0.8610      0.0000      -Inf < 2e-16 ***
## Year1999      -0.9675      0.3066      -3.16 0.00228 **
## Year2000      -0.9850      0.0000      -Inf < 2e-16 ***
## Year2001      -0.6075      0.0721      -8.42 1.6e-12 ***
## Year2002      -0.9805      0.1553      -6.31 1.6e-08 ***
## Year2003      -0.5372      0.1238      -4.34 4.3e-05 ***
## Year2004      -0.2490      0.0629      -3.96 0.00017 ***
## Year2005      -0.1701      0.1688      -1.01 0.31688
## Year2006      -0.6905      0.1768      -3.91 0.00020 ***
## Year2007      -0.7250      0.2214      -3.27 0.00159 **
## Year2008      -0.7056      0.2869      -2.46 0.01615 *
## Year2009      -0.3895      0.1431      -2.72 0.00803 **
## Year2010      -0.6704      0.2188      -3.06 0.00300 **
## Year2011      -0.6319      0.2341      -2.70 0.00854 **
## Year2012      -0.8583      0.3641      -2.36 0.02093 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.615
## Multiple R-squared:  0.127, Adjusted R-squared:  -0.0543
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 85 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.604  0.900  0.956  0.927  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      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 1704"
## [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    4    1    6    5    1    5    8    4    3    4    4    7    5    7
##
## 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    4    1    1    3    0    4    7    2    2    3    2    4    5    2
##
## 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    4    1    1    3    0    4    7    2    2    3    2    3    5    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: 4.16016764610381"

## Warning in wilcox.test.default(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: 3"
## [1] "Male last author team size 2018 geometric mean: 4.16016764610381"

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

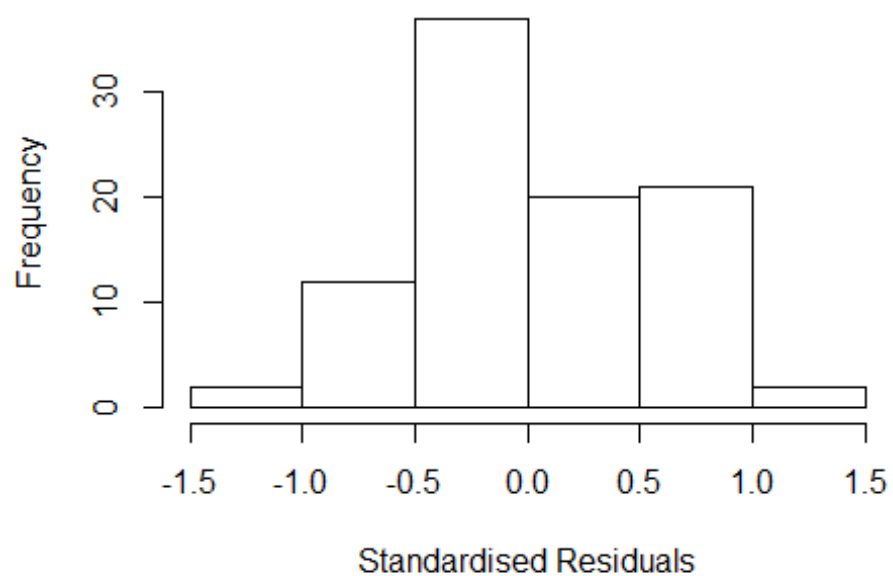
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0.5, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"

## 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

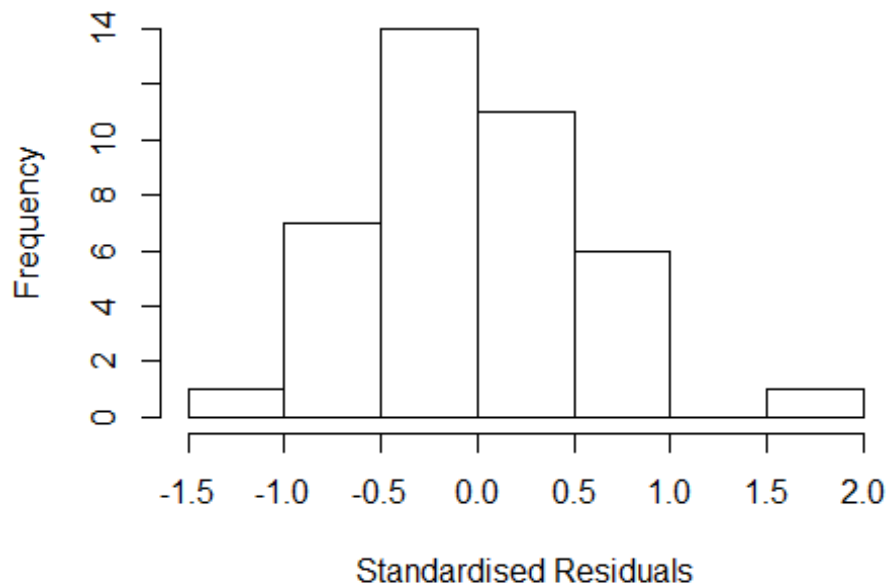
```

Residuals from last author



| | | | | |
|----|-------------------|------|----|-----------------|
| ## | | GVIF | Df | GVIF^(1/(2*Df)) |
| ## | FirstAuthorFemale | NaN | 1 | NaN |
| ## | LastAuthorFemale | NaN | 1 | NaN |
| ## | UniqueAuthors | NaN | 4 | NaN |
| ## | Year | NaN | 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
## -1.25e+00 -3.74e-01 -8.08e-16 4.08e-01 1.72e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.28e-15 3.74e-08 0.00e+00 1.000
## FirstAuthorFemale1 -5.18e-01 3.92e-01 -1.32e+00 0.201
## LastAuthorFemale1 2.92e-01 2.93e-01 9.90e-01 0.332
## UniqueAuthors2 2.13e-01 3.63e-01 5.90e-01 0.564
## UniqueAuthors3 7.33e-01 4.10e-01 1.79e+00 0.089 .
## UniqueAuthors4 6.89e-01 4.80e-01 1.44e+00 0.167
## UniqueAuthors5 1.73e+00 2.69e-01 6.42e+00 2.9e-06 ***
## Year1999 6.35e-01 5.21e-01 1.22e+00 0.237
## Year2000 3.05e-01 0.00e+00 Inf < 2e-16 ***
## Year2001 1.06e+00 3.63e-08 2.93e+07 < 2e-16 ***
```

```

## Year2002          1.20e+00  4.25e-01  2.83e+00  0.010 *
## Year2004          1.60e+00  5.43e-01  2.95e+00  0.008 **
## Year2005          8.48e-01  5.30e-01  1.60e+00  0.125
## Year2006          1.66e-01  4.37e-01  3.80e-01  0.708
## Year2007          6.20e-01  5.44e-01  1.14e+00  0.267
## Year2008          5.18e-01  3.62e-01  1.43e+00  0.168
## Year2009          4.00e-01  8.51e-01  4.70e-01  0.643
## Year2010          7.88e-01  3.42e-01  2.30e+00  0.032 *
## Year2011          2.96e-01  2.35e-01  1.26e+00  0.224
## Year2012         -2.68e-01  8.99e-01 -3.00e-01  0.769
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.833
## Multiple R-squared:  0.435, Adjusted R-squared:  -0.102
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 35 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.650  0.931  0.973   0.951   0.985   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          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 4.451 1      2.110
## LastAuthorFemale  6.720 1      2.592
## Year              10.675 13      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 + LastAuthorFemale + Year, data =

```



```

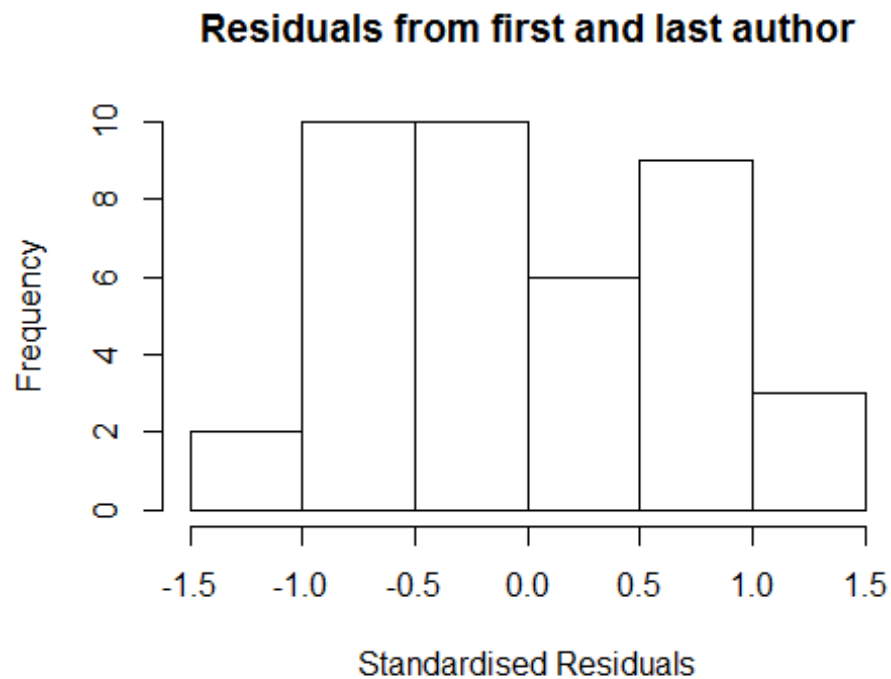
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q   Median      3Q      Max
## -1.36e+00 -6.33e-01 -1.09e-15  5.42e-01  1.28e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.29e-15   1.54e-08  0.00e+00  1.00000
## FirstAuthorFemale1 -2.23e-01   4.88e-01 -4.60e-01  0.65122
## LastAuthorFemale1  2.55e-01   3.15e-01  8.10e-01  0.42546
## Year1999         7.97e-01   4.54e-01  1.75e+00  0.09229 .
## Year2000         3.05e-01   2.31e-08  1.32e+07 < 2e-16 ***
## Year2001         1.06e+00   2.44e-08  4.37e+07 < 2e-16 ***
## Year2002         1.40e+00   5.05e-01  2.78e+00  0.01035 *
## Year2004         1.71e+00   4.36e-01  3.91e+00  0.00065 ***
## Year2005         1.29e+00   2.47e-01  5.24e+00  2.3e-05 ***
## Year2006         6.39e-01   4.84e-01  1.32e+00  0.19895
## Year2007         8.33e-01   3.92e-01  2.13e+00  0.04399 *
## Year2008         1.13e+00   3.97e-01  2.85e+00  0.00879 **
## Year2009         1.15e+00   6.12e-01  1.88e+00  0.07223 .
## Year2010         1.18e+00   2.48e-01  4.74e+00  8.0e-05 ***
## Year2011         7.12e-01   4.48e-01  1.59e+00  0.12512
## Year2012         5.01e-01   6.70e-01  7.50e-01  0.46173
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 1.07
## Multiple R-squared:  0.232, Adjusted R-squared:  -0.248
## Convergence in 9 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 34 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.859  0.950  0.968  0.958  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      2.50e-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"

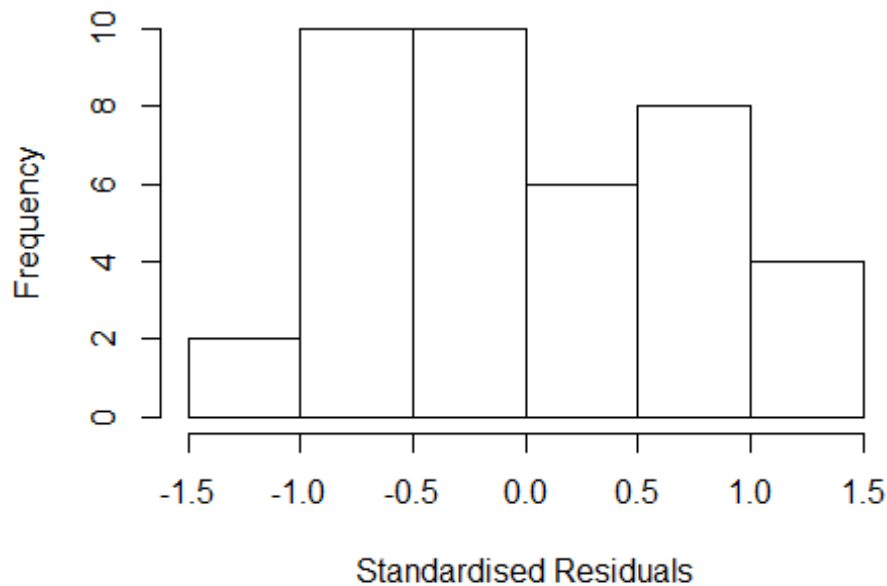
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

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



```
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1          NaN
## Year              NaN 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
## -1.43e+00 -6.33e-01 -5.55e-17 5.71e-01 1.24e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.06e-15 0.00e+00 -Inf < 2e-16 ***
## FirstAuthorFemale1 -1.00e-01 5.22e-01 -0.19 0.84936
## Year1999 7.96e-01 4.65e-01 1.71 0.09942 .
## Year2000 3.05e-01 0.00e+00 Inf < 2e-16 ***
## Year2001 1.07e+00 0.00e+00 Inf < 2e-16 ***
## Year2002 1.38e+00 5.62e-01 2.45 0.02146 *
## Year2004 1.78e+00 4.59e-01 3.88 0.00068 ***
## Year2005 1.33e+00 2.28e-01 5.82 4.5e-06 ***
## Year2006 6.39e-01 4.89e-01 1.31 0.20297
## Year2007 8.34e-01 3.95e-01 2.11 0.04491 *
## Year2008 1.17e+00 4.43e-01 2.64 0.01422 *
## Year2009 1.28e+00 4.82e-01 2.66 0.01358 *
```

```

## Year2010          1.22e+00  2.73e-01  4.47  0.00015 ***
## Year2011          7.67e-01  4.36e-01  1.76  0.09107 .
## Year2012          7.57e-01  5.99e-01  1.26  0.21810
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 1
## Multiple R-squared:  0.225, Adjusted R-squared:  -0.208
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 36 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.822  0.943  0.964  0.954  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      2.50e-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 2.088 1          1.445
## Year            2.088 13          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
## -1.38e+00 -5.98e-01  6.67e-16  5.42e-01  1.27e+00
##
## Coefficients:
##      Estimate Std. Error  t value Pr(>|t|)

```

```

## (Intercept)      -1.06e-15   3.19e-08 0.00e+00 1.00000
## LastAuthorFemale1 1.87e-01   3.08e-01 6.10e-01 0.54940
## Year1999          7.96e-01   4.60e-01 1.73e+00 0.09542 .
## Year2000          3.05e-01   4.00e-08 7.63e+06 < 2e-16 ***
## Year2001          1.07e+00   3.10e-08 3.44e+07 < 2e-16 ***
## Year2002          1.35e+00   6.31e-01 2.14e+00 0.04267 *
## Year2004          1.73e+00   4.41e-01 3.91e+00 0.00062 ***
## Year2005          1.30e+00   2.45e-01 5.32e+00 1.6e-05 ***
## Year2006          6.39e-01   4.85e-01 1.32e+00 0.19985
## Year2007          8.34e-01   3.93e-01 2.12e+00 0.04398 *
## Year2008          1.09e+00   3.59e-01 3.03e+00 0.00568 **
## Year2009          1.19e+00   5.82e-01 2.04e+00 0.05252 .
## Year2010          1.13e+00   1.85e-01 6.09e+00 2.3e-06 ***
## Year2011          6.47e-01   3.97e-01 1.63e+00 0.11636
## Year2012          5.70e-01   6.67e-01 8.50e-01 0.40098
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 1.05
## Multiple R-squared:  0.229, Adjusted R-squared:  -0.202
## Convergence in 9 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 33 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.848  0.943  0.966  0.955  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] "Sample size for the above analysis: 40"
## [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]"

```

```

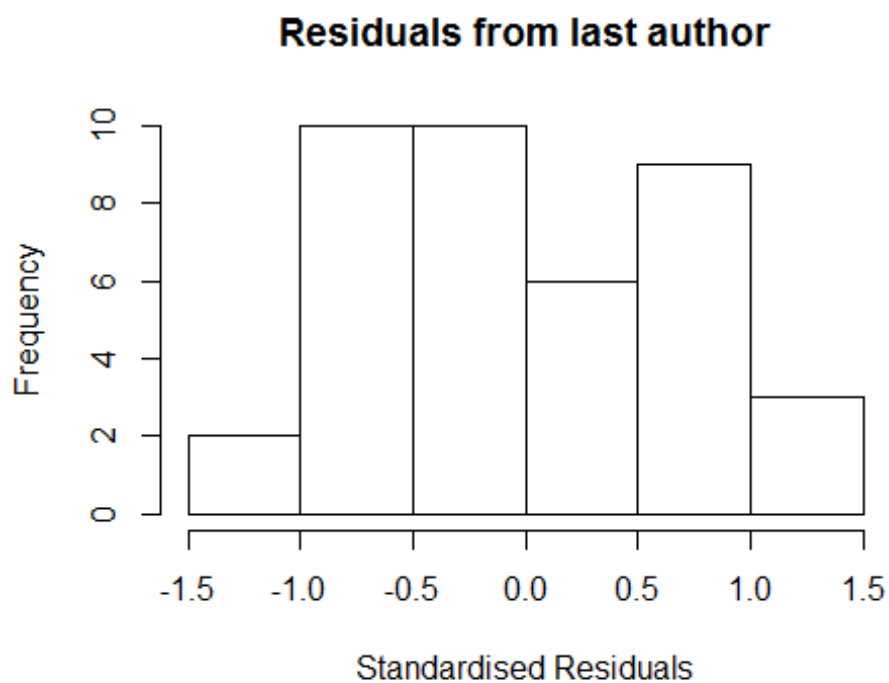
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    5    2    4    9    9   10   10   12   11   24   37   28   36   28
## 2011 2012
##   34   30
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    4    1    1    5    7    5    8   10    9   18   27   20   26   21
## 2011 2012
##   23   16
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    4    1    1    3    7    5    8    9    9   18   22   17   23   17
## 2011 2012
##   19   14
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.05976714390712"
## [1] "Male first author team size 2018 geometric mean: 3.34174480013438"

## Warning in wilcox.test.default(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 = 0.1
## 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: 3.06439349308142"

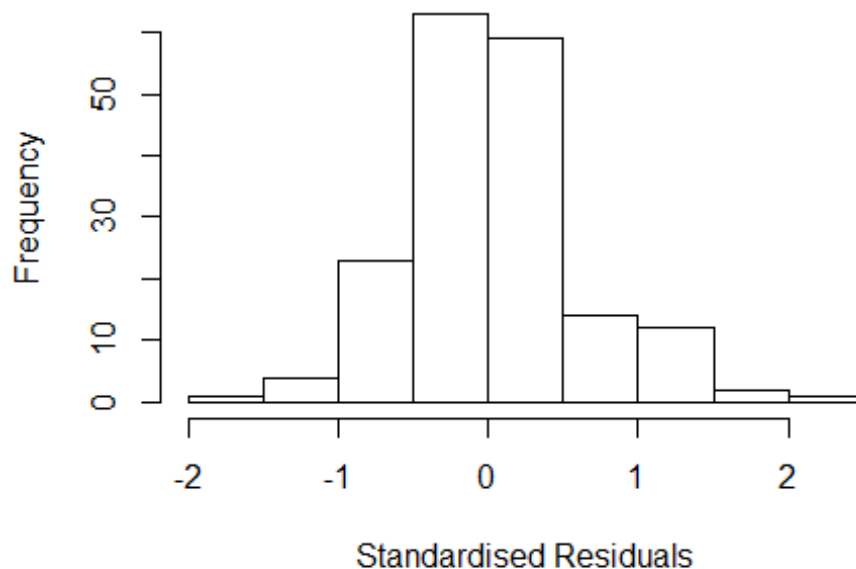
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 13, p-value = 0.7
## 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 9.314e+13  1      9.651e+06
## LastAuthorFemale  2.955e+00  1      1.719e+00
## UniqueAuthors    3.549e+14  4      6.588e+01
## Year              2.681e+15 16      3.035e+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
## -1.62e+00 -3.37e-01 -1.33e-15  3.36e-01  2.02e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.5452    0.2321    2.35 0.02006 *
## FirstAuthorFemale1 -0.0542    0.1667   -0.33 0.74548
## LastAuthorFemale1 -0.2056    0.1731   -1.19 0.23677
## UniqueAuthors2     0.1324    0.1296    1.02 0.30854
## UniqueAuthors3    -0.0891    0.1434   -0.62 0.53503
## UniqueAuthors4     0.3541    0.1851    1.91 0.05756 .
## UniqueAuthors5     0.0417    0.3605    0.12 0.90808
## Year1997          -0.0071    0.2604   -0.03 0.97828
## Year1998          -0.6234    0.2758   -2.26 0.02521 *
## Year1999          -0.5452    0.2321   -2.35 0.02006 *
```

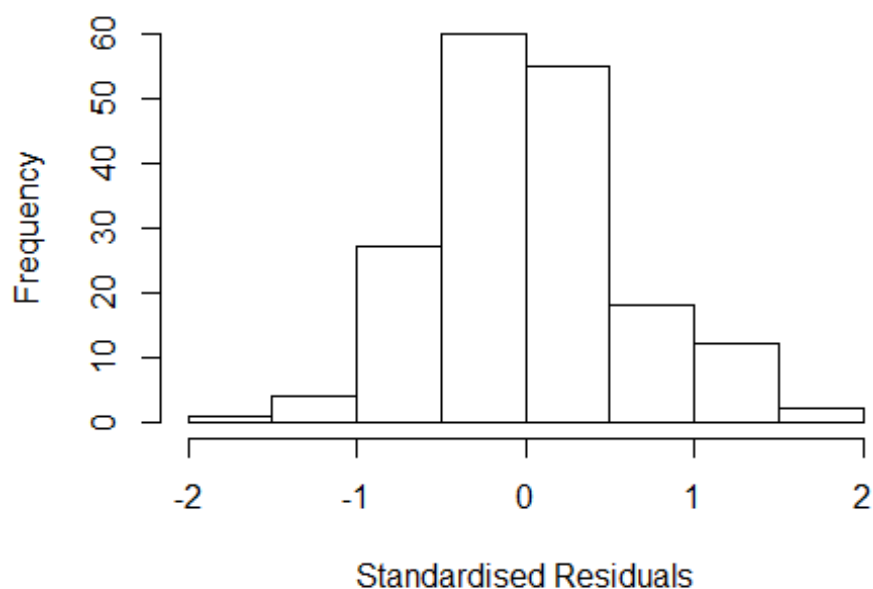


```

## Year2000          0.0982      0.3796      0.26  0.79614
## Year2001          0.6690      0.4319      1.55  0.12343
## Year2002          1.0654      0.2712      3.93  0.00013 ***
## Year2003          0.9029      0.2990      3.02  0.00296 **
## Year2004          0.6713      0.2770      2.42  0.01652 *
## Year2005          1.1675      0.2976      3.92  0.00013 ***
## Year2006          0.2964      0.3086      0.96  0.33825
## Year2007          0.8060      0.2715      2.97  0.00346 **
## Year2008          0.2744      0.2653      1.03  0.30272
## Year2009          0.4721      0.2913      1.62  0.10709
## Year2010          0.2307      0.2515      0.92  0.36058
## Year2011          0.2969      0.2531      1.17  0.24259
## Year2012          1.0764      0.3117      3.45  0.00071 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.544
## Multiple R-squared:  0.307, Adjusted R-squared:  0.21
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 165 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.137  0.860  0.959  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          5.59e-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 3.657e+14 1      1.912e+07
## LastAuthorFemale  2.807e+00 1      1.675e+00
## Year              6.048e+14 16      2.897e+00

```

Residuals from first and last author



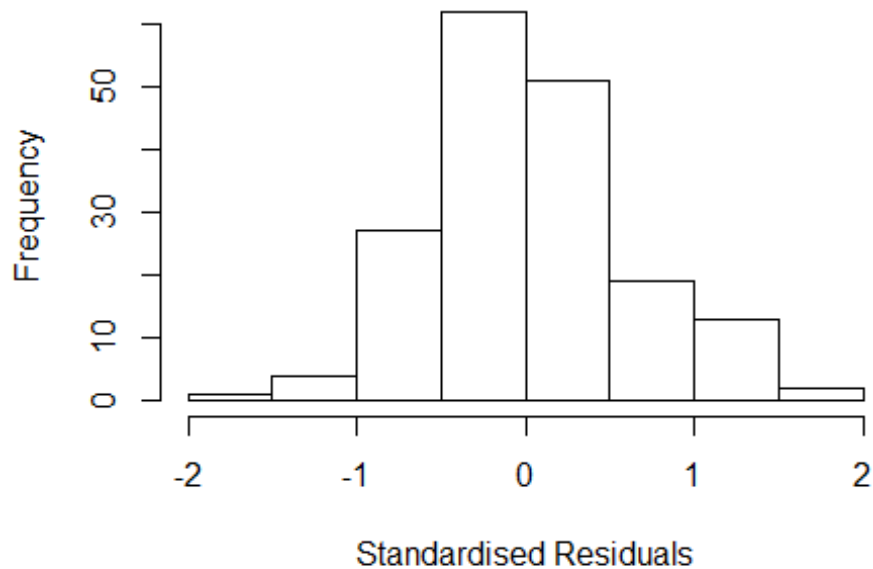
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.70685 -0.37536 -0.00313 0.37631 1.77446
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.63391 0.25753 2.46 0.0149 *
## FirstAuthorFemale1 -0.08731 0.16069 -0.54 0.5876
## LastAuthorFemale1 -0.21751 0.18370 -1.18 0.2382
## Year1997 0.00383 0.27878 0.01 0.9891
## Year1998 -0.54660 0.29056 -1.88 0.0618 .
## Year1999 -0.63391 0.25753 -2.46 0.0149 *
## Year2000 0.00584 0.38993 0.01 0.9881
## Year2001 0.63606 0.43883 1.45 0.1492
## Year2002 0.99886 0.30223 3.30 0.0012 **
## Year2003 0.85163 0.32833 2.59 0.0104 *
## Year2004 0.73512 0.30274 2.43 0.0163 *
## Year2005 1.07294 0.32562 3.30 0.0012 **
```

```

## Year2006          0.36763      0.31622      1.16      0.2467
## Year2007          0.82536      0.29731      2.78      0.0062 **
## Year2008          0.25575      0.29226      0.88      0.3828
## Year2009          0.43521      0.29116      1.49      0.1369
## Year2010          0.16522      0.27703      0.60      0.5518
## Year2011          0.28120      0.28175      1.00      0.3198
## Year2012          1.04168      0.31483      3.31      0.0012 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.61
## Multiple R-squared:  0.262, Adjusted R-squared:  0.179
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 164 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.377  0.865   0.960   0.910   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      5.59e-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 2.383e+13 1      4.882e+06
## Year              2.383e+13 16      2.618e+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
## -1.65596 -0.37479 -0.00707 0.36604 1.75097
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5705 0.3077 1.85 0.0656 .
## FirstAuthorFemale1 -0.1780 0.1285 -1.39 0.1679
## Year1997 0.0672 0.3257 0.21 0.8367
## Year1998 -0.3925 0.3077 -1.28 0.2040
## Year1999 -0.5705 0.3077 -1.85 0.0656 .
## Year2000 0.0694 0.4248 0.16 0.8704
## Year2001 0.6992 0.4856 1.44 0.1519
## Year2002 1.0568 0.3378 3.13 0.0021 **
## Year2003 0.9286 0.3635 2.55 0.0116 *
## Year2004 0.8019 0.3431 2.34 0.0206 *
## Year2005 1.0855 0.3804 2.85 0.0049 **
## Year2006 0.4545 0.3492 1.30 0.1949
```

```

## Year2007          0.9044      0.3335      2.71      0.0074 **
## Year2008          0.2825      0.3395      0.83      0.4066
## Year2009          0.4856      0.3402      1.43      0.1555
## Year2010          0.2043      0.3275      0.62      0.5336
## Year2011          0.3295      0.3295      1.00      0.3189
## Year2012          1.1197      0.3564      3.14      0.0020 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.607
## Multiple R-squared:  0.259, Adjusted R-squared:  0.181
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 165 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.385  0.864  0.959  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      5.59e-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 1.611 1      1.269
## Year      1.611 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
## -1.7121 -0.3724 -0.0192  0.3848  1.7980

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.6175    0.2695    2.29  0.0232 *
## LastAuthorFemale1 -0.2720    0.1426   -1.91  0.0581 .
## Year1997        0.0202    0.2899    0.07  0.9445
## Year1998       -0.6175    0.2695   -2.29  0.0232 *
## Year1999       -0.6175    0.2695   -2.29  0.0232 *
## Year2000        0.0224    0.3982    0.06  0.9552
## Year2001        0.6488    0.4338    1.50  0.1367
## Year2002        0.9898    0.3107    3.19  0.0017 **
## Year2003        0.8550    0.3413    2.50  0.0132 *
## Year2004        0.7283    0.3171    2.30  0.0229 *
## Year2005        1.0945    0.3340    3.28  0.0013 **
## Year2006        0.3605    0.3291    1.10  0.2750
## Year2007        0.8283    0.3083    2.69  0.0080 **
## Year2008        0.2707    0.3013    0.90  0.3702
## Year2009        0.4539    0.3002    1.51  0.1324
## Year2010        0.1830    0.2881    0.64  0.5262
## Year2011        0.2936    0.2933    1.00  0.3183
## Year2012        1.0433    0.3238    3.22  0.0015 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.606
## Multiple R-squared:  0.261, Adjusted R-squared:  0.183
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 165 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.359  0.864  0.959  0.909  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      5.59e-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: 179"
## [1] ""

```

```

## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1706"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    20    1   33   19   19   38   48   33   57   44   48   52   60   74   62
## 2011 2012
##    59    53
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     5     1   15     5   14   18   17   23   33   28   30   30   34   53   45
## 2011 2012
##    44    37
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     5     1   11     4   13   16   16   21   26   21   28   27   28   48   37
## 2011 2012
##    40    33
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.05976714390712"
## [1] "Male first author team size 2018 geometric mean: 2.38983586834283"

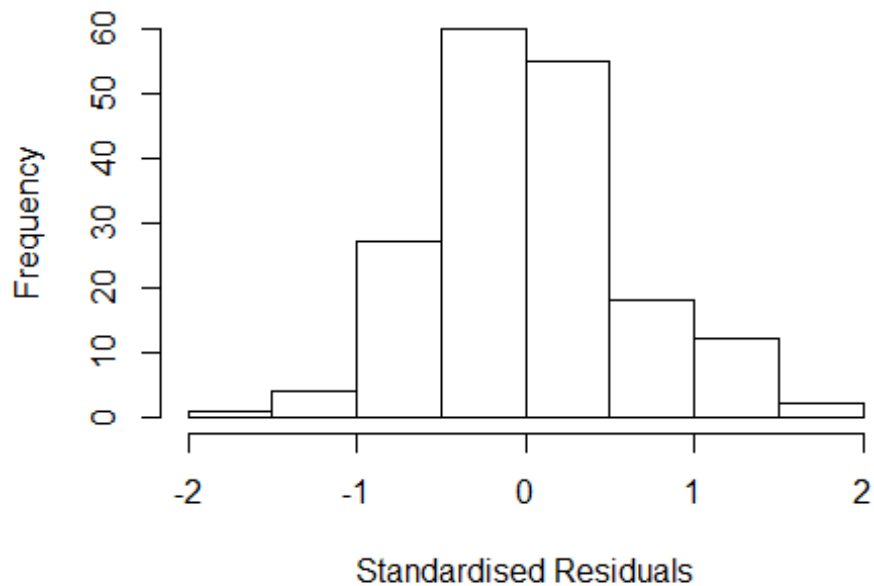
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## 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: 2.27950705695478"
## [1] "Male last author team size 2018 geometric mean: 2.31644820288969"

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

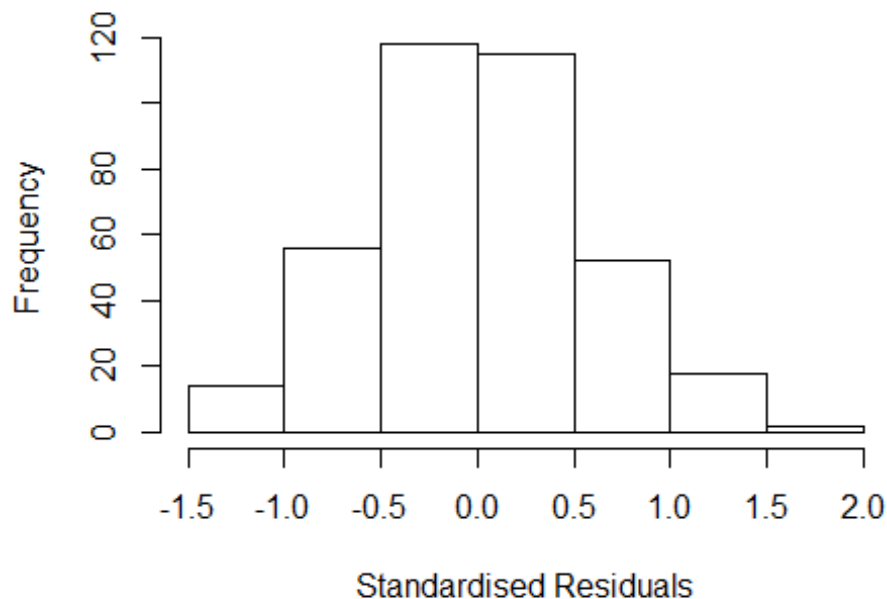
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 26, 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 1.428 1          1.195
## LastAuthorFemale  1.345 1          1.160
## UniqueAuthors    4.832 4          1.218
## Year              6.435 16         1.060
```


Residuals from first and last author and team size



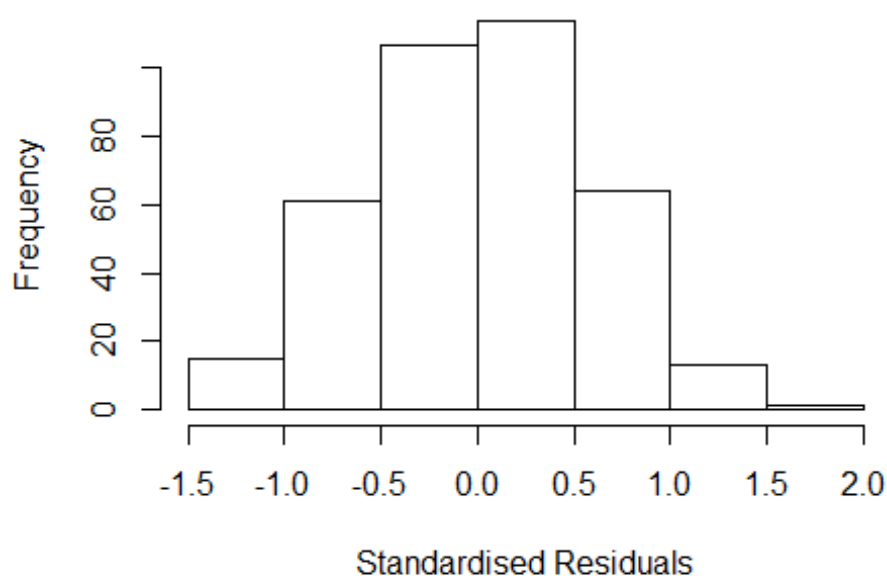
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.40e+00 -3.79e-01 -4.44e-16 3.54e-01 1.52e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.4609 0.1893 2.43 0.0154 *
## FirstAuthorFemale1 -0.1296 0.0986 -1.31 0.1896
## LastAuthorFemale1 0.0935 0.1151 0.81 0.4169
## UniqueAuthors2 0.2265 0.0901 2.51 0.0124 *
## UniqueAuthors3 0.3253 0.1025 3.17 0.0016 **
## UniqueAuthors4 0.3373 0.1208 2.79 0.0055 **
## UniqueAuthors5 0.2988 0.1153 2.59 0.0100 **
## Year1997 1.3806 0.2013 6.86 3.1e-11 ***
## Year1998 0.2419 0.2618 0.92 0.3562
## Year1999 0.5066 0.5008 1.01 0.3124
```

```

## Year2000          0.6830      0.2426      2.82      0.0051 **
## Year2001          0.4351      0.2514      1.73      0.0844 .
## Year2002          0.4990      0.2596      1.92      0.0554 .
## Year2003          0.5348      0.2190      2.44      0.0151 *
## Year2004          0.6395      0.2110      3.03      0.0026 **
## Year2005          0.7092      0.2676      2.65      0.0084 **
## Year2006          0.5664      0.2204      2.57      0.0106 *
## Year2007          0.5342      0.2404      2.22      0.0269 *
## Year2008          0.4091      0.2380      1.72      0.0865 .
## Year2009          0.4781      0.2172      2.20      0.0284 *
## Year2010          0.3211      0.2108      1.52      0.1285
## Year2011          0.2034      0.2096      0.97      0.3326
## Year2012          0.2669      0.2217      1.20      0.2294
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.575
## Multiple R-squared:  0.111, Adjusted R-squared:  0.0554
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 31 weights are ~= 1. The remaining 344 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.464  0.862  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      2.67e-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 1.381 1      1.175
## LastAuthorFemale  1.302 1      1.141
## Year              1.495 16      1.013

```

Residuals from first and last author



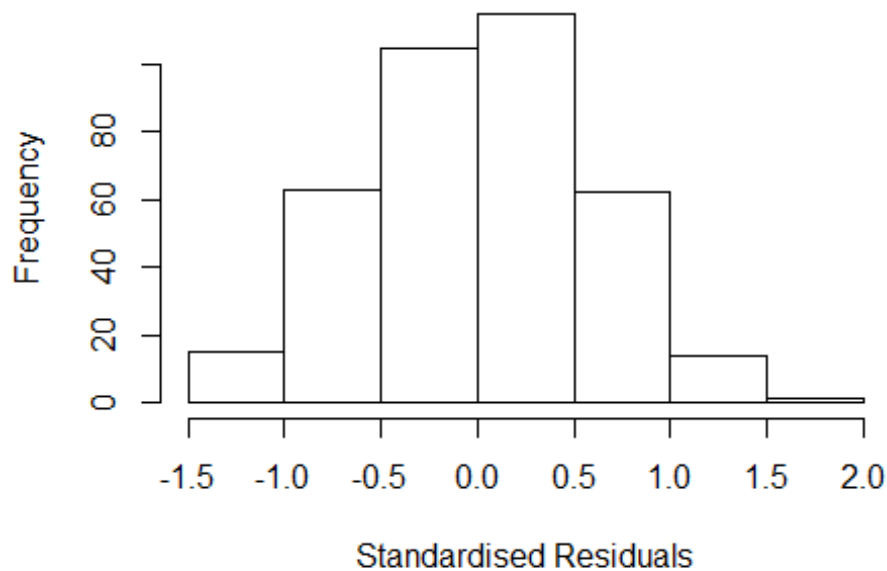
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4169 -0.4322 0.0193 0.3829 1.6376
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5869 0.1865 3.15 0.00179 **
## FirstAuthorFemale1 -0.1304 0.0982 -1.33 0.18528
## LastAuthorFemale1 0.0497 0.1129 0.44 0.66036
## Year1997 1.4811 0.1865 7.94 2.6e-14 ***
## Year1998 0.2858 0.2677 1.07 0.28646
## Year1999 0.5801 0.5381 1.08 0.28175
## Year2000 0.7278 0.2424 3.00 0.00287 **
## Year2001 0.5081 0.2506 2.03 0.04329 *
## Year2002 0.5381 0.2477 2.17 0.03047 *
## Year2003 0.6289 0.2122 2.96 0.00325 **
## Year2004 0.6854 0.2080 3.29 0.00108 **
## Year2005 0.8300 0.2635 3.15 0.00177 **
```

```

## Year2006          0.7138      0.2146      3.33  0.00097 ***
## Year2007          0.6175      0.2357      2.62  0.00917 **
## Year2008          0.4905      0.2349      2.09  0.03754 *
## Year2009          0.5863      0.2127      2.76  0.00613 **
## Year2010          0.4093      0.2102      1.95  0.05224 .
## Year2011          0.3164      0.2055      1.54  0.12451
## Year2012          0.3875      0.2166      1.79  0.07438 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.611
## Multiple R-squared:  0.0729, Adjusted R-squared:  0.026
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 41 weights are ~= 1. The remaining 334 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.453  0.867  0.949  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      2.67e-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 1.297 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
## -1.4187 -0.4271 0.0289 0.3789 1.6390
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5870 0.1866 3.15 0.0018 **
## FirstAuthorFemale1 -0.1203 0.0951 -1.27 0.2067
## Year1997 1.4810 0.1866 7.94 2.7e-14 ***
## Year1998 0.2941 0.2669 1.10 0.2713
## Year1999 0.5774 0.5426 1.06 0.2880
## Year2000 0.7299 0.2427 3.01 0.0028 **
## Year2001 0.5101 0.2504 2.04 0.0423 *
## Year2002 0.5450 0.2462 2.21 0.0275 *
## Year2003 0.6317 0.2127 2.97 0.0032 **
## Year2004 0.6896 0.2080 3.31 0.0010 **
## Year2005 0.8318 0.2655 3.13 0.0019 **
## Year2006 0.7137 0.2151 3.32 0.0010 **
```

```

## Year2007          0.6160      0.2351      2.62      0.0092 **
## Year2008          0.4957      0.2332      2.13      0.0342 *
## Year2009          0.5915      0.2118      2.79      0.0055 **
## Year2010          0.4126      0.2103      1.96      0.0506 .
## Year2011          0.3173      0.2055      1.54      0.1236
## Year2012          0.3964      0.2166      1.83      0.0681 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.609
## Multiple R-squared:  0.0725, Adjusted R-squared:  0.0284
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 43 weights are ~= 1. The remaining 332 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.449  0.866  0.947  0.908  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      2.67e-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 1.229 1      1.108
## Year      1.229 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
## -1.3913 -0.4207  0.0167  0.3888  1.6687

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.5870    0.1866   3.15  0.0018 **
## LastAuthorFemale1 0.0127    0.1114   0.11  0.9092
## Year1997       1.4810    0.1866   7.94 2.7e-14 ***
## Year1998       0.2773    0.2685   1.03  0.3025
## Year1999       0.5499    0.5802   0.95  0.3439
## Year2000       0.6944    0.2389   2.91  0.0039 **
## Year2001       0.4830    0.2547   1.90  0.0587 .
## Year2002       0.5331    0.2471   2.16  0.0316 *
## Year2003       0.6073    0.2106   2.88  0.0042 **
## Year2004       0.6702    0.2084   3.22  0.0014 **
## Year2005       0.8044    0.2668   3.02  0.0028 **
## Year2006       0.6987    0.2161   3.23  0.0013 **
## Year2007       0.5863    0.2324   2.52  0.0121 *
## Year2008       0.4885    0.2341   2.09  0.0376 *
## Year2009       0.5620    0.2121   2.65  0.0084 **
## Year2010       0.3967    0.2102   1.89  0.0599 .
## Year2011       0.3075    0.2050   1.50  0.1345
## Year2012       0.3854    0.2165   1.78  0.0759 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.609
## Multiple R-squared:  0.0682, Adjusted R-squared:  0.0239
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 37 weights are ~= 1. The remaining 338 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.433  0.866  0.950  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      2.67e-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: 375"
## [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]"
##
## 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    2    1    3    2    3    8    6    9    6   11   12    7    7
##
## 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    0    1    1    0    3    1    3    6    4    8    5   11    8    5    5
##
## 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    0    1    1    0    1    0    3    6    4    8    4   10    6    4    5
## [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: 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 = 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: 3.77976314968462"
## [1] "Male last author team size 2018 geometric mean: 2"

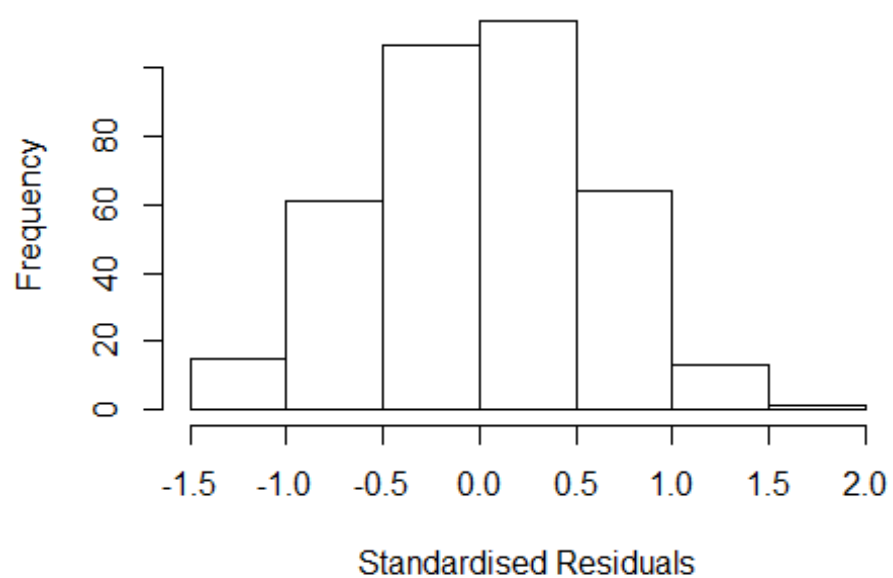
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 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"

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

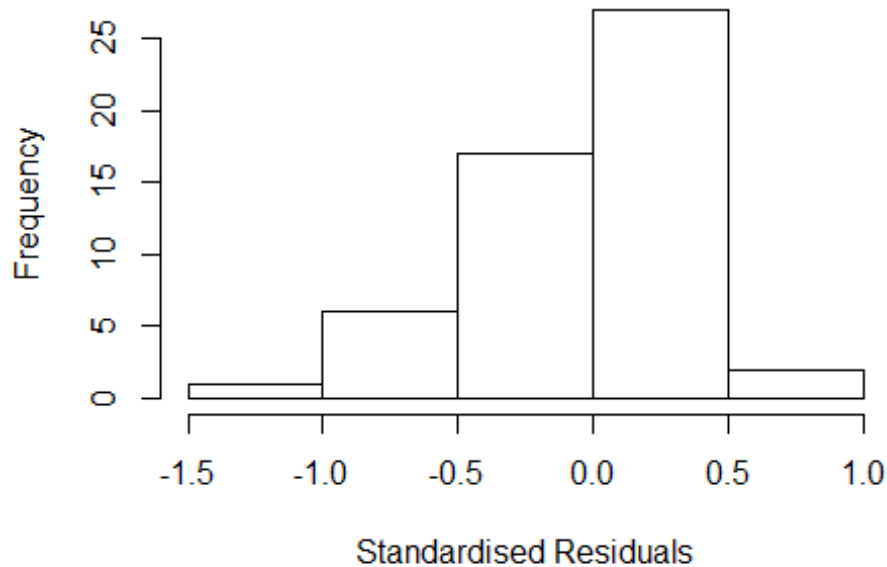
```


Residuals from last author



| ## | | GVIF | Df | $GVIF^{1/(2 \cdot Df)}$ |
|----|-------------------|-----------|----|-------------------------|
| ## | FirstAuthorFemale | 5.566e+00 | 1 | 2.359 |
| ## | LastAuthorFemale | 3.556e+00 | 1 | 1.886 |
| ## | UniqueAuthors | 1.483e+30 | 4 | 5907.356 |
| ## | Year | 8.250e+30 | 11 | 25.427 |

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
## -1.3796 -0.2997 0.0177 0.2375 0.8249
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8000 0.0000 Inf < 2e-16 ***
## FirstAuthorFemale1 -0.4210 0.2165 -1.95 0.060 .
## LastAuthorFemale1 0.6347 0.2375 2.67 0.011 *
## UniqueAuthors2 0.2878 0.2525 1.14 0.262
## UniqueAuthors3 0.1433 0.1912 0.75 0.459
## UniqueAuthors4 -0.1906 0.6336 -0.30 0.765
## UniqueAuthors5 0.3319 0.2603 1.27 0.211
## Year2000 1.3522 0.2525 5.36 5.5e-06 ***
## Year2002 1.6536 0.6336 2.61 0.013 *
## Year2004 0.3185 0.2616 1.22 0.232
```

```

## Year2005          0.6302      0.3237      1.95      0.060 .
## Year2006          0.2927      0.1694      1.73      0.093 .
## Year2007          0.5958      0.3230      1.84      0.074 .
## Year2008          0.6409      0.2399      2.67      0.011 *
## Year2009          0.5967      0.2724      2.19      0.035 *
## Year2010          0.3194      0.3410      0.94      0.355
## Year2011          0.3434      0.2684      1.28      0.209
## Year2012         -0.0231      0.3596     -0.06      0.949
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.433
## Multiple R-squared:  0.417, Adjusted R-squared:  0.134
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~ = 1. The remaining 44 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.288  0.908  0.954  0.912  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      1.89e-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 4.768 1      2.184
## LastAuthorFemale  1.968 1      1.403
## Year              6.839 11      1.091

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))

```

```

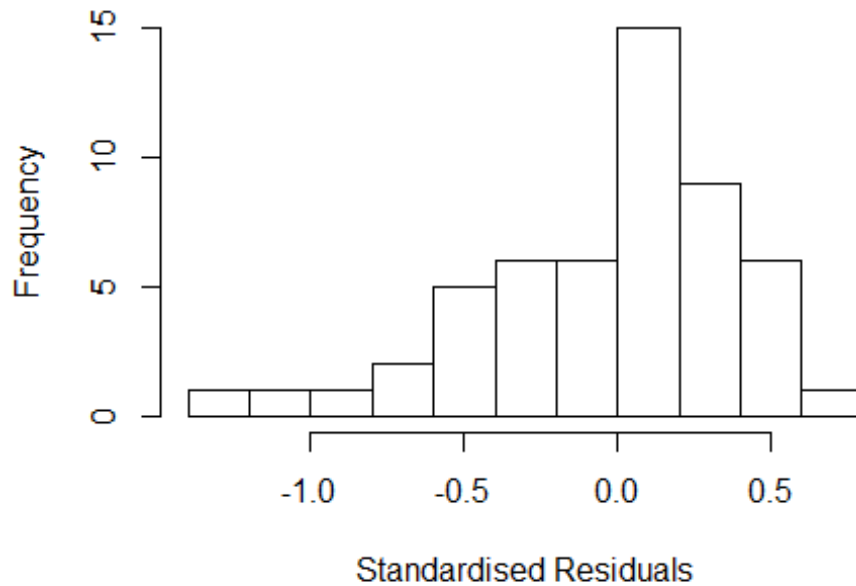
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3106 -0.3311  0.0649  0.2279  0.6184
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    8.00e-01   1.56e-08  5.13e+07 < 2e-16 ***
## FirstAuthorFemale1 -4.30e-01   1.89e-01 -2.27e+00  0.02862 *
## LastAuthorFemale1  5.83e-01   2.16e-01  2.71e+00  0.01007 *
## Year2000          1.64e+00   2.24e-08  7.34e+07 < 2e-16 ***
## Year2002          1.46e+00   7.80e-09  1.87e+08 < 2e-16 ***
## Year2004          5.15e-01   1.63e-01  3.16e+00  0.00303 **
## Year2005          8.02e-01   1.86e-01  4.30e+00  0.00011 ***
## Year2006          4.51e-01   1.04e-01  4.33e+00   1e-04 ***
## Year2007          8.15e-01   2.69e-01  3.03e+00  0.00433 **
## Year2008          8.57e-01   8.93e-02  9.60e+00   8e-12 ***
## Year2009          7.46e-01   1.91e-01  3.91e+00  0.00036 ***
## Year2010          5.97e-01   2.38e-01  2.50e+00  0.01662 *
## Year2011          5.67e-01   2.51e-01  2.26e+00  0.02949 *
## Year2012          2.26e-01   2.33e-01  9.70e-01  0.33811
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.439
## Multiple R-squared:  0.367, Adjusted R-squared:  0.156
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 47 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.352  0.907  0.957  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      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

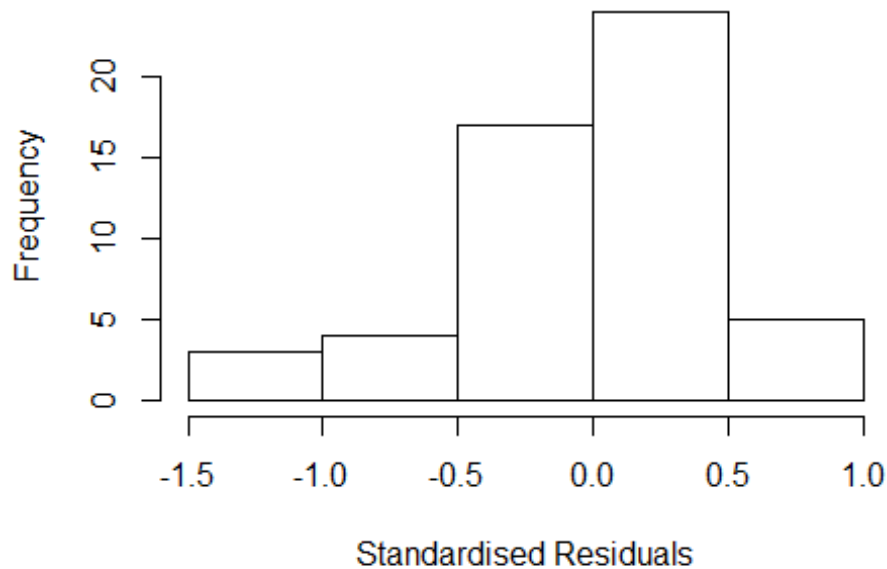
## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

Residuals from first and last author



```
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale  NaN  1          NaN
## Year              NaN 11          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
## -1.3926 -0.2719 0.0313 0.2284 0.9198
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8000 0.0000 Inf < 2e-16 ***
## FirstAuthorFemale1 -0.3272 0.2135 -1.53 0.13321
## Year2000 1.6400 0.0000 Inf < 2e-16 ***
## Year2002 1.4630 0.0000 Inf < 2e-16 ***
## Year2004 0.5136 0.1610 3.19 0.00277 **
## Year2005 0.8005 0.1818 4.40 7.7e-05 ***
## Year2006 0.4506 0.1037 4.35 9.3e-05 ***
## Year2007 0.8966 0.2588 3.46 0.00128 **
## Year2008 0.8309 0.0912 9.11 2.7e-11 ***
## Year2009 0.7609 0.1885 4.04 0.00024 ***
## Year2010 0.7003 0.1664 4.21 0.00014 ***
## Year2011 0.5681 0.2432 2.34 0.02458 *
```

```

## Year2012          0.3012      0.3116      0.97  0.33955
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.476
## Multiple R-squared:  0.289, Adjusted R-squared:  0.0758
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 42 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.371  0.888  0.951  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      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 1.709 1      1.307
## Year          1.709 11      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
## -1.3199 -0.3309  0.0351  0.2477  0.7671
##
## Coefficients:
##      Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    8.00e-01   2.82e-08  2.84e+07 < 2e-16 ***
## LastAuthorFemale1 5.05e-01   2.09e-01  2.42e+00  0.0201 *

```

```

## Year2000      1.64e+00  2.64e-08 6.22e+07 < 2e-16 ***
## Year2002      1.46e+00  3.15e-08 4.64e+07 < 2e-16 ***
## Year2004      5.14e-01  1.62e-01 3.17e+00  0.0029 **
## Year2005      8.02e-01  1.85e-01 4.33e+00  9.8e-05 ***
## Year2006      4.51e-01  1.04e-01 4.34e+00  9.5e-05 ***
## Year2007      8.24e-01  2.65e-01 3.11e+00  0.0035 **
## Year2008      7.49e-01  9.91e-02 7.55e+00  3.3e-09 ***
## Year2009      5.59e-01  1.62e-01 3.45e+00  0.0013 **
## Year2010      6.07e-01  2.26e-01 2.68e+00  0.0105 *
## Year2011      5.67e-01  2.49e-01 2.28e+00  0.0280 *
## Year2012      2.40e-01  2.38e-01 1.01e+00  0.3182
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.448
## Multiple R-squared:  0.324, Adjusted R-squared:  0.122
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 46 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.365  0.902  0.950  0.911  0.986  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.89e-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 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
##    5    1    3    7    7    6    8    6    7    7    10   19   17   10   22
## 2011 2012

```



```

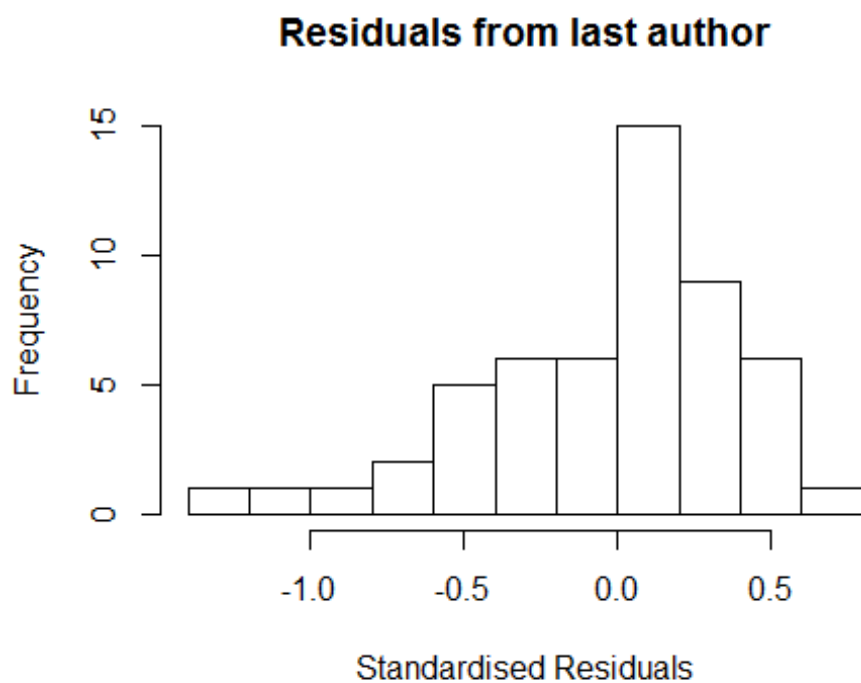
## 20 15
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4 1 2 4 4 5 6 4 4 5 9 13 11 8 14
## 2011 2012
## 13 11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3 1 2 4 4 4 6 4 4 5 9 10 9 6 11
## 2011 2012
## 10 7
## [1] "Heteroscedasticity 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: 2.35215804504935"

## Warning in wilcox.test.default(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: 3"
## [1] "Male last author team size 2018 geometric mean: 2.35215804504935"

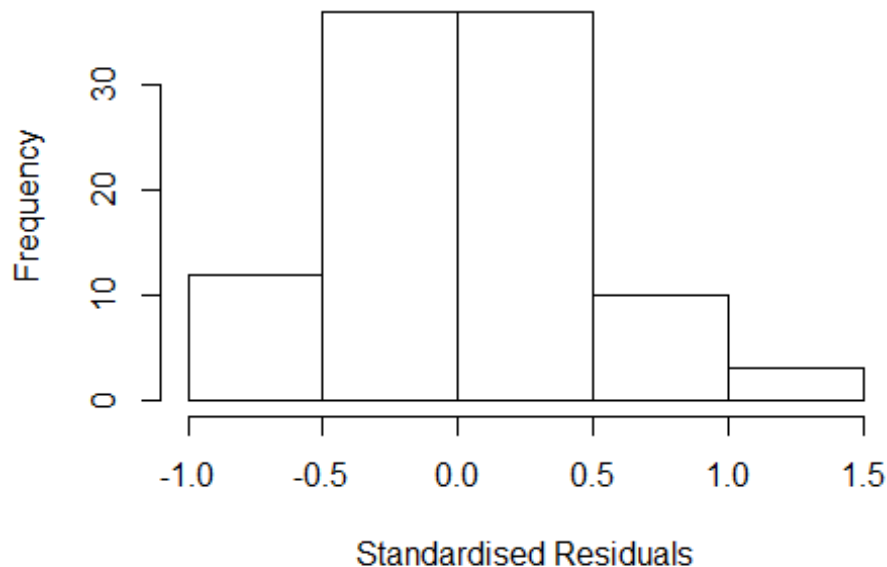
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## 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    4.637  1      2.153
## LastAuthorFemale     6.861  1      2.619
## UniqueAuthors      208.680  4      1.950
## Year                2049.135 16      1.269
```

Residuals from first and last author and team size



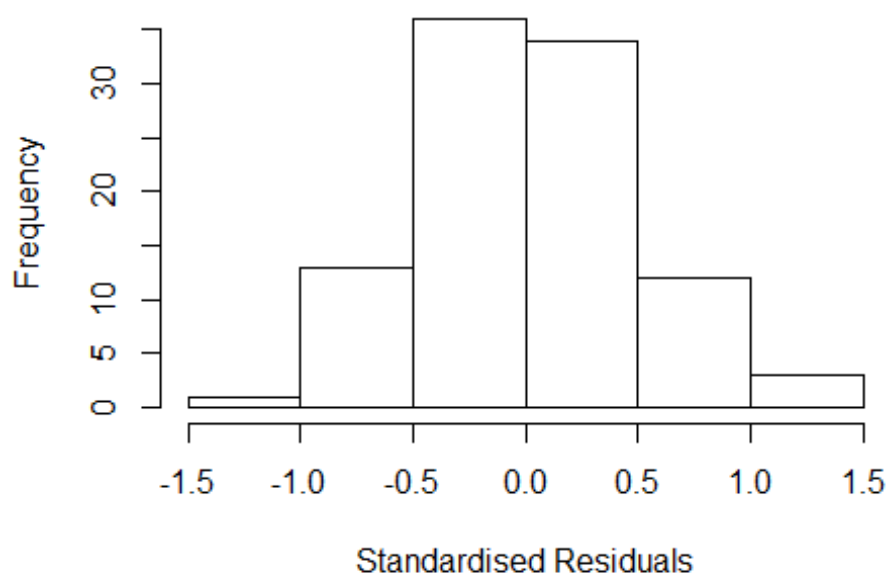
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8916 -0.2729 0.0184 0.2394 1.4783
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5000 0.4884 1.02 0.3092
## FirstAuthorFemale1 0.5748 0.3043 1.89 0.0627 .
## LastAuthorFemale1 -0.5389 0.3153 -1.71 0.0915 .
## UniqueAuthors2 -0.0921 0.2026 -0.45 0.6506
## UniqueAuthors3 -0.1040 0.1975 -0.53 0.6000
## UniqueAuthors4 0.0550 0.2288 0.24 0.8107
## UniqueAuthors5 -0.6754 0.2971 -2.27 0.0258 *
## Year1997 -0.4079 0.4550 -0.90 0.3729
## Year1998 -0.7414 0.5319 -1.39 0.1674
## Year1999 0.2744 0.5375 0.51 0.6111
```

```

## Year2000          0.7688      0.4607      1.67      0.0993 .
## Year2001          0.5596      0.5760      0.97      0.3343
## Year2002          0.1380      0.4876      0.28      0.7780
## Year2003          0.5484      0.4765      1.15      0.2534
## Year2004          1.2342      0.4388      2.81      0.0063 **
## Year2005          1.2969      0.5558      2.33      0.0223 *
## Year2006          0.3504      0.5458      0.64      0.5229
## Year2007          0.9632      0.4585      2.10      0.0390 *
## Year2008          0.6984      0.4770      1.46      0.1473
## Year2009          0.4407      0.4526      0.97      0.3332
## Year2010          0.4956      0.4466      1.11      0.2706
## Year2011          0.7355      0.4499      1.64      0.1062
## Year2012          0.2538      0.6259      0.41      0.6863
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.459
## Multiple R-squared:  0.447, Adjusted R-squared:  0.287
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 88 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.277  0.875   0.950   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          1.01e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.785 1          1.945
## LastAuthorFemale  6.162 1          2.482
## Year             11.144 16          1.078

```

Residuals from first and last author



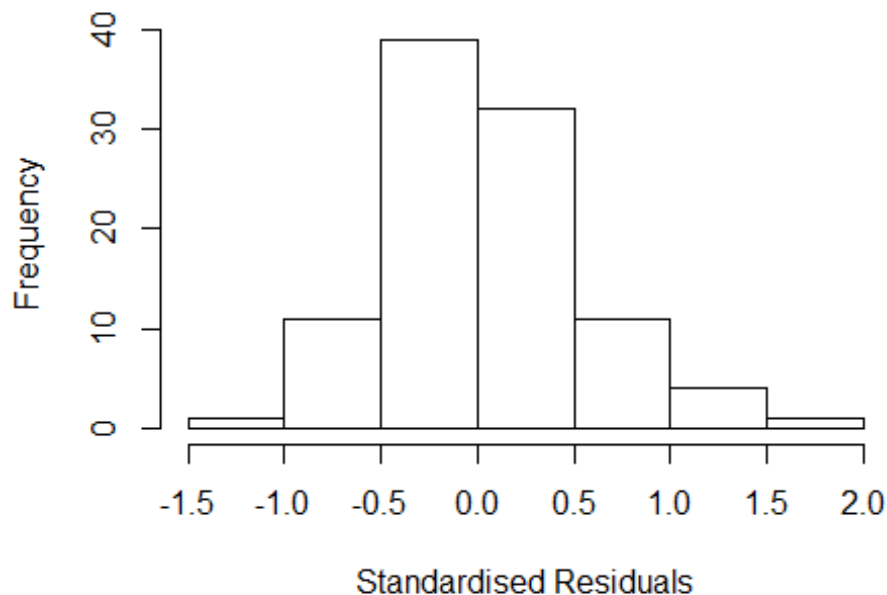
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.14e+00 -3.24e-01 5.55e-16 2.96e-01 1.45e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5250 0.4203 1.25 0.2153
## FirstAuthorFemale1 0.6188 0.2905 2.13 0.0362 *
## LastAuthorFemale1 -0.5194 0.3098 -1.68 0.0976 .
## Year1997 -0.5250 0.4203 -1.25 0.2153
## Year1998 -0.8344 0.5087 -1.64 0.1049
## Year1999 0.1635 0.4857 0.34 0.7373
## Year2000 0.6947 0.4236 1.64 0.1050
## Year2001 0.3074 0.7076 0.43 0.6652
## Year2002 0.0502 0.4576 0.11 0.9129
## Year2003 0.4280 0.4494 0.95 0.3437
## Year2004 1.1760 0.4265 2.76 0.0072 **
## Year2005 1.1644 0.5437 2.14 0.0353 *
```

```

## Year2006          0.1458      0.5274      0.28      0.7828
## Year2007          0.7878      0.4294      1.83      0.0703 .
## Year2008          0.5846      0.4442      1.32      0.1920
## Year2009          0.3346      0.4402      0.76      0.4494
## Year2010          0.3422      0.4425      0.77      0.4416
## Year2011          0.6144      0.4717      1.30      0.1964
## Year2012          0.1620      0.6139      0.26      0.7926
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.448
## Multiple R-squared:  0.398, Adjusted R-squared:  0.262
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 87 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.270  0.864  0.947  0.896  0.987  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      1.01e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
##   nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##   trace.lev      mts      compute.rd
##      0          1000          0
##           psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 5.3 1          2.302
## Year              5.3 16          1.054

```

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
## -1.0974 -0.2950 -0.0045 0.2950 1.6444
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5235 0.4224 1.24 0.2188
## FirstAuthorFemale1 0.4202 0.4198 1.00 0.3198
## Year1997 -0.5235 0.4224 -1.24 0.2188
## Year1998 -0.7336 0.5063 -1.45 0.1512
## Year1999 0.1655 0.4879 0.34 0.7354
## Year2000 0.6961 0.4257 1.64 0.1059
## Year2001 0.3093 0.7035 0.44 0.6614
## Year2002 0.0503 0.4591 0.11 0.9131
## Year2003 0.4295 0.4515 0.95 0.3442
## Year2004 1.1775 0.4285 2.75 0.0074 **
## Year2005 0.8750 0.4568 1.92 0.0589 .
## Year2006 0.2362 0.5748 0.41 0.6822
```

```

## Year2007          0.7892      0.4314      1.83      0.0710 .
## Year2008          0.5859      0.4463      1.31      0.1929
## Year2009          0.3720      0.4487      0.83      0.4095
## Year2010          0.2726      0.4454      0.61      0.5422
## Year2011          0.5739      0.4842      1.19      0.2394
## Year2012         -0.0279      0.7180     -0.04      0.9691
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.444
## Multiple R-squared:  0.376, Adjusted R-squared:  0.245
## Convergence in 31 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 86 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.140  0.860  0.937  0.886  0.986  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      1.01e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale -5.02e-14  1      NaN
## Year             -5.02e-14 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
## -1.19e+00 -2.36e-01 -1.11e-16  2.89e-01  1.74e+00

```



```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.5231    0.4227   1.24  0.2195
## LastAuthorFemale1 -0.2116    0.1845  -1.15  0.2548
## Year1997          -0.5231    0.4227  -1.24  0.2195
## Year1998          -0.5231    0.4227  -1.24  0.2195
## Year1999           0.1661    0.4879   0.34  0.7345
## Year2000           0.6966    0.4260   1.64  0.1059
## Year2001           0.3098    0.7171   0.43  0.6669
## Year2002           0.0503    0.4596   0.11  0.9131
## Year2003           0.4300    0.4517   0.95  0.3440
## Year2004           1.1779    0.4288   2.75  0.0074 **
## Year2005           0.9932    0.4936   2.01  0.0475 *
## Year2006           0.4232    0.5665   0.75  0.4572
## Year2007           0.7896    0.4317   1.83  0.0711 .
## Year2008           0.5863    0.4467   1.31  0.1930
## Year2009           0.4380    0.4476   0.98  0.3307
## Year2010           0.3517    0.4403   0.80  0.4267
## Year2011           0.6644    0.4685   1.42  0.1600
## Year2012          -0.0906    0.6146  -0.15  0.8832
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.443
## Multiple R-squared:  0.388, Adjusted R-squared:  0.26
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 87 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0892 0.8590 0.9490 0.8830 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      1.01e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 99"
## [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 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    1    1    3    2    5    1    4    2    8    4    8   10   11   13
## 2012
##    9
##
## 1996 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    1    1    3    1    2    1    4    1    6    2    6    9    9   11
## 2012
##    7
##
## 1996 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    1    1    3    0    2    1    4    1    6    2    6    7    7    9
## 2012
##    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: 1.90636858599387"

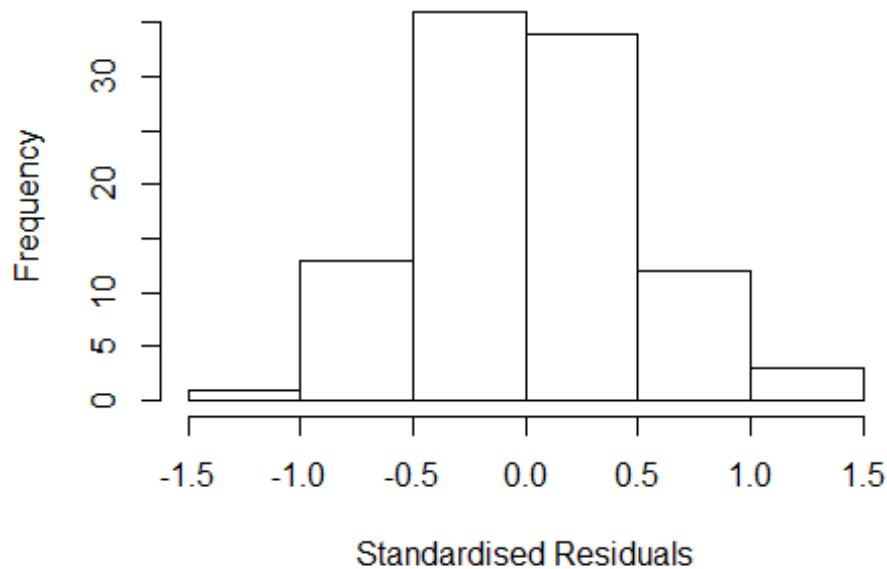
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 6.5, p-value = 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: 2.1192680453341"

## 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 = 0.2
## 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 1.946e+16  1      139500421
## LastAuthorFemale  8.473e+17  1      920472950
## UniqueAuthors    4.179e+31  4         8967
## Year              2.834e+45 14          42

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##      Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##      k.max = 1000))
## \--> method = "MM"
## Residuals:
```

```

##      Min      1Q  Median      3Q      Max
## -0.8901 -0.2984  0.0065  0.2451  0.8345
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   -4.90e-01   3.57e-01 -1.37e+00  0.17806
## FirstAuthorFemale1 -1.47e-03   3.38e-01  0.00e+00  0.99655
## LastAuthorFemale1  4.92e-01   5.36e-01  9.20e-01  0.36460
## UniqueAuthors2    1.01e-01   2.69e-01  3.80e-01  0.70885
## UniqueAuthors3    2.34e-01   2.48e-01  9.40e-01  0.35200
## UniqueAuthors4    4.07e-01   2.36e-01  1.73e+00  0.09323 .
## UniqueAuthors5    8.20e-01   2.39e-01  3.43e+00  0.00156 **
## Year1997          9.97e-01   3.57e-01  2.80e+00  0.00836 **
## Year1998          9.63e-01   3.43e-08  2.81e+07 < 2e-16 ***
## Year2000          1.59e+00   3.65e-01  4.37e+00  0.00011 ***
## Year2002          1.58e+00   2.08e-01  7.63e+00  6.0e-09 ***
## Year2003          2.42e+00   4.64e-01  5.21e+00  8.4e-06 ***
## Year2004          1.61e+00   3.97e-01  4.06e+00  0.00027 ***
## Year2005          2.14e+00   3.93e-01  5.44e+00  4.2e-06 ***
## Year2006          1.60e+00   4.90e-01  3.26e+00  0.00250 **
## Year2007          1.68e+00   4.39e-01  3.82e+00  0.00053 ***
## Year2008          1.04e+00   2.97e-01  3.52e+00  0.00123 **
## Year2009          1.46e+00   3.90e-01  3.73e+00  0.00068 ***
## Year2010          1.53e+00   4.02e-01  3.81e+00  0.00054 ***
## Year2011          1.01e+00   4.06e-01  2.49e+00  0.01759 *
## Year2012          1.30e+00   5.02e-01  2.58e+00  0.01428 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.607
## Multiple R-squared:  0.399, Adjusted R-squared:  0.0555
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~ = 1. The remaining 47 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.813  0.912  0.953  0.945  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          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"

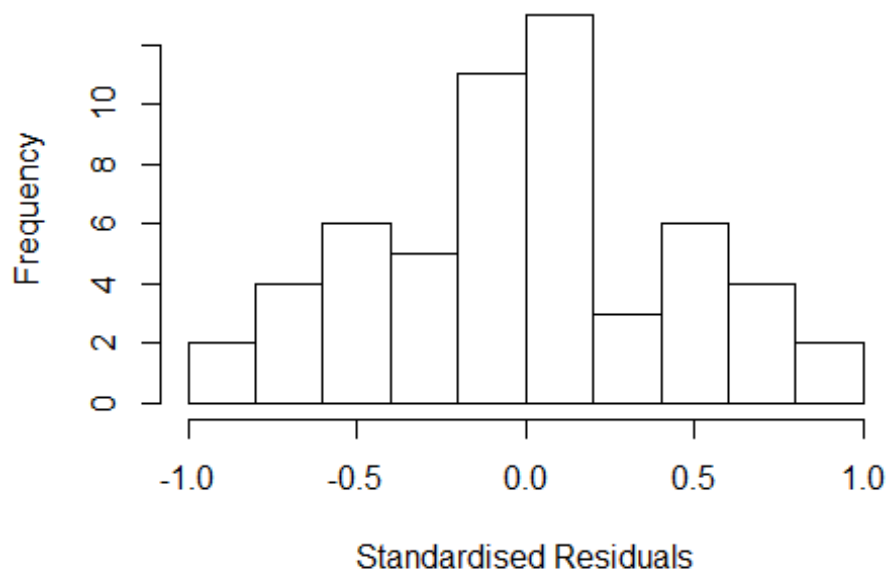
```

```
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Regression 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:
```

```

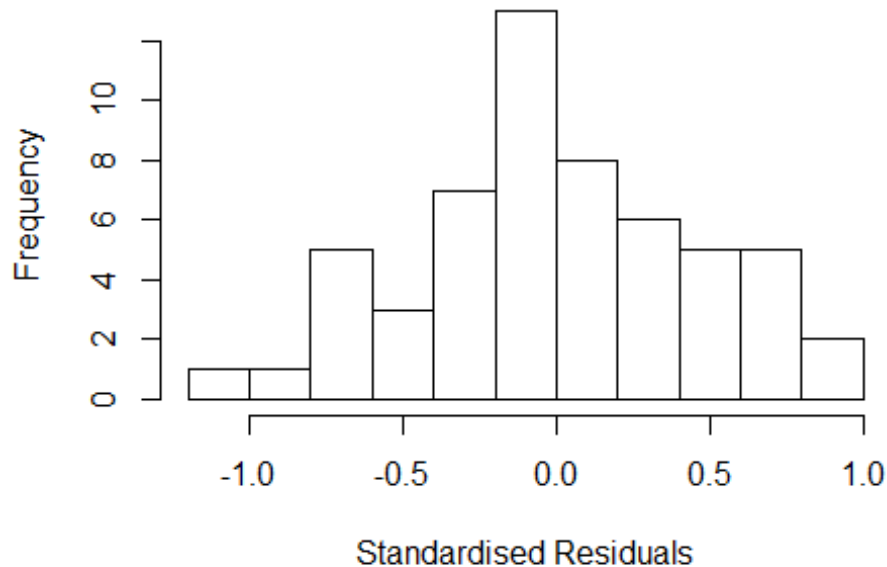
##           Min           1Q           Median           3Q           Max
## -1.09e+00 -3.12e-01  5.55e-17  3.62e-01  8.10e-01
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -0.422      0.354   -1.19  0.24002
## FirstAuthorFemale1 -0.139      0.318   -0.44  0.66433
## LastAuthorFemale1  0.561      0.481    1.17  0.25087
## Year1997         0.929      0.354    2.63  0.01225 *
## Year1998         0.963      0.000    Inf < 2e-16 ***
## Year2000         1.601      0.355    4.51  5.7e-05 ***
## Year2002         1.550      0.221    7.03  2.0e-08 ***
## Year2003         2.588      0.481    5.38  3.8e-06 ***
## Year2004         2.007      0.393    5.11  8.9e-06 ***
## Year2005         2.480      0.354    7.01  2.1e-08 ***
## Year2006         1.714      0.482    3.56  0.00101 **
## Year2007         1.929      0.446    4.33  0.00010 ***
## Year2008         1.055      0.296    3.56  0.00100 ***
## Year2009         1.571      0.397    3.95  0.00032 ***
## Year2010         1.678      0.376    4.46  6.8e-05 ***
## Year2011         1.203      0.436    2.76  0.00876 **
## Year2012         1.510      0.531    2.84  0.00705 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.58
## Multiple R-squared:  0.328, Adjusted R-squared:  0.0524
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 47 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.705  0.885  0.960  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           1.79e-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
```

Residuals from first and last author



```
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1             NaN
## Year              NaN 14            NaN

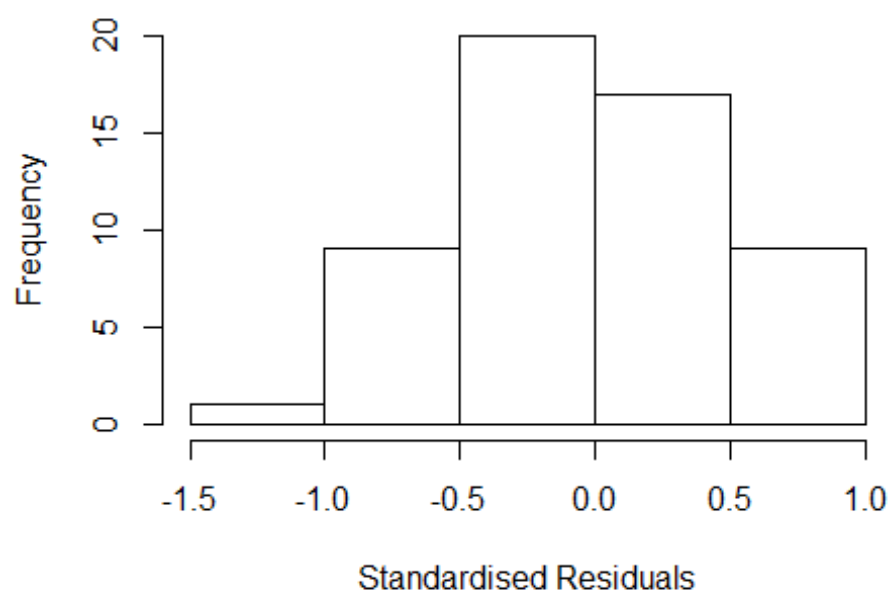
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q   Median      3Q      Max
## -1.05e+00 -3.61e-01  2.22e-16  3.61e-01  9.32e-01
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   -4.00e-02   2.47e-01 -1.60e-01  0.87234
## FirstAuthorFemale1  4.00e-02   2.47e-01  1.60e-01  0.87234
```

```

## Year1997          5.47e-01    2.47e-01    2.21e+00    0.03267 *
## Year1998          9.63e-01    1.16e-08    8.32e+07    < 2e-16 ***
## Year2000          1.22e+00    2.49e-01    4.90e+00    1.6e-05 ***
## Year2002          1.36e+00    3.12e-01    4.36e+00    8.9e-05 ***
## Year2003          2.03e+00    0.00e+00      Inf    < 2e-16 ***
## Year2004          1.58e+00    2.79e-01    5.67e+00    1.4e-06 ***
## Year2005          2.10e+00    2.47e-01    8.49e+00    1.7e-10 ***
## Year2006          1.26e+00    2.82e-01    4.46e+00    6.4e-05 ***
## Year2007          1.55e+00    3.66e-01    4.23e+00    0.00013 ***
## Year2008          9.20e-01    2.68e-01    3.44e+00    0.00139 **
## Year2009          1.14e+00    2.79e-01    4.10e+00    0.00020 ***
## Year2010          1.30e+00    2.82e-01    4.60e+00    4.2e-05 ***
## Year2011          8.20e-01    3.44e-01    2.38e+00    0.02198 *
## Year2012          1.09e+00    4.03e-01    2.71e+00    0.00987 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.604
## Multiple R-squared:  0.305, Adjusted R-squared:  0.045
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 48 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.742  0.912  0.954  0.936  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          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

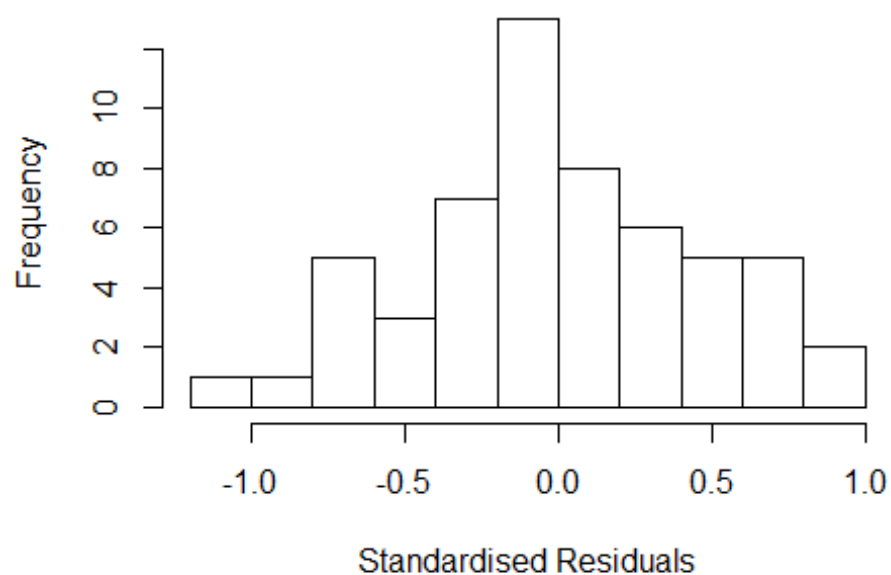
```


Residuals from first author



```
##          GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale  NaN  1          NaN
## Year              NaN 14          NaN
```

Residuals from last author

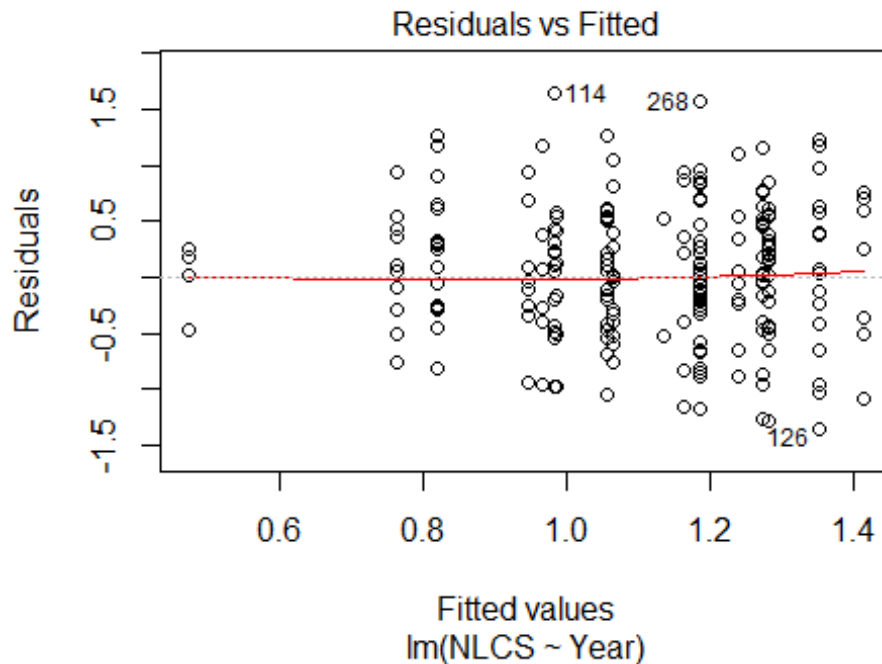


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.05e+00 -3.18e-01 2.22e-16 3.64e-01 8.12e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.420 0.348 -1.21 0.23527
## LastAuthorFemale1 0.420 0.348 1.21 0.23527
## Year1997 0.927 0.348 2.66 0.01115 *
## Year1998 0.963 0.000 Inf < 2e-16 ***
## Year2000 1.599 0.349 4.58 4.5e-05 ***
## Year2002 1.549 0.219 7.08 1.4e-08 ***
## Year2003 2.447 0.348 7.03 1.7e-08 ***
## Year2004 1.972 0.389 5.06 9.7e-06 ***
## Year2005 2.478 0.348 7.12 1.3e-08 ***
## Year2006 1.653 0.440 3.76 0.00055 ***
## Year2007 1.927 0.439 4.39 8.1e-05 ***
## Year2008 1.056 0.290 3.65 0.00076 ***
## Year2009 1.532 0.391 3.92 0.00033 ***
## Year2010 1.676 0.371 4.52 5.5e-05 ***
## Year2011 1.198 0.424 2.82 0.00737 **
## Year2012 1.473 0.497 2.96 0.00512 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.643
## Multiple R-squared: 0.318, Adjusted R-squared: 0.0626
## Convergence in 9 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~ 1. The remaining 47 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.770 0.921 0.958 0.944 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 1.79e-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 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
##    6    5   13    6   10   11    9   13    9   17   25   27   20   39   32
## 2011 2012
##   27   24
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    2    8    4    7    9    7   11    8   15   19   24   14   34   23
## 2011 2012
##   23   20
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    2    6    3    7    9    6   10    8   15   18   20   10   29   21
## 2011 2012
##   15   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 = 7.5, df = 16, p-value = 1

```



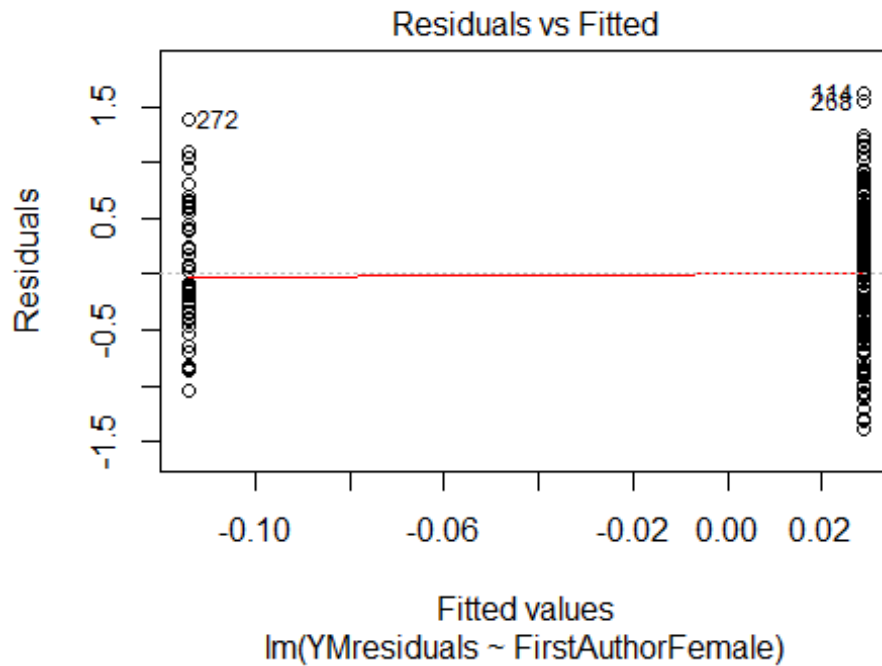
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.065, df = 1, p-value = 0.8

## [1] "Female first author team size 2018 geometric mean: 2.04767251107922"
## [1] "Male first author team size 2018 geometric mean: 3.07097855013232"

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

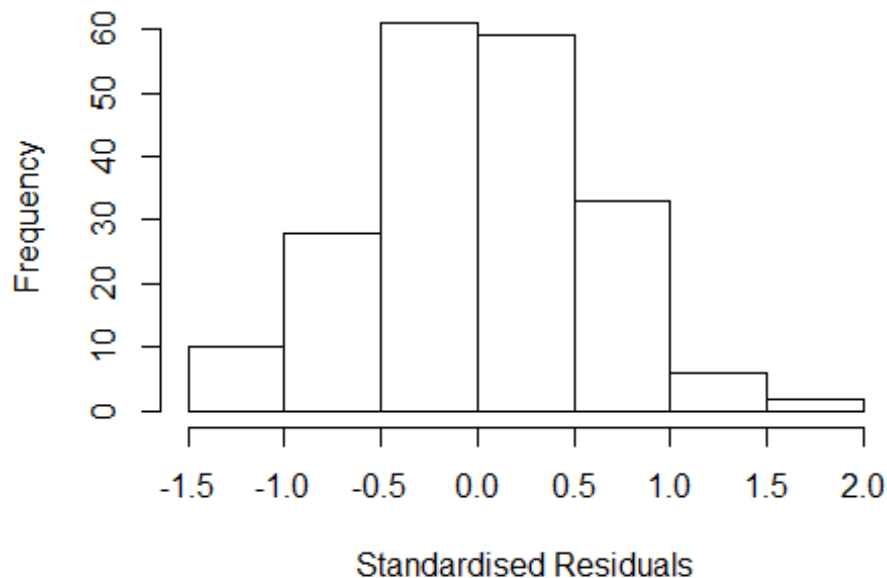
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 18, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.40224886796286"
## [1] "Male last author team size 2018 geometric mean: 2.90072293183791"

## Warning in wilcox.test.default(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 1.659 1          1.288
## LastAuthorFemale  1.948 1          1.396
## UniqueAuthors    3.008 4          1.148
## Year              3.972 16         1.044
```

Residuals from first and last author and team size



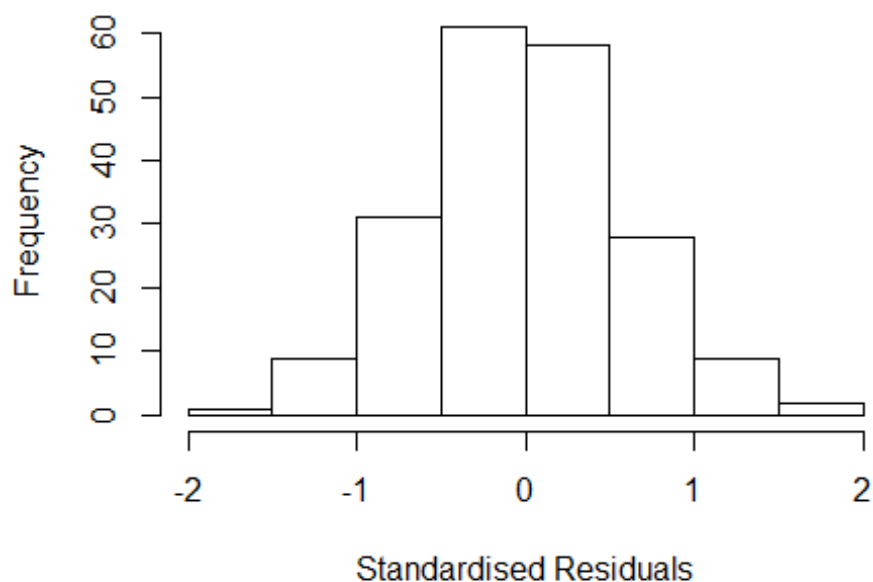
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.26289 -0.41243 0.00383 0.35193 1.54002
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.29563 0.51761 2.50 0.013 *
## FirstAuthorFemale1 -0.25769 0.12760 -2.02 0.045 *
## LastAuthorFemale1 -0.09404 0.13767 -0.68 0.495
## UniqueAuthors2 -0.07092 0.12308 -0.58 0.565
## UniqueAuthors3 -0.07975 0.15046 -0.53 0.597
## UniqueAuthors4 0.37503 0.17010 2.20 0.029 *
## UniqueAuthors5 -0.47255 0.31815 -1.49 0.139
## Year1997 0.00368 0.59152 0.01 0.995
## Year1998 0.12613 0.59561 0.21 0.833
## Year1999 -0.86112 0.54648 -1.58 0.117
```

```

## Year2000      -0.07113      0.60961      -0.12      0.907
## Year2001      0.06702      0.55952      0.12      0.905
## Year2002     -0.20654      0.59260     -0.35      0.728
## Year2003     -0.33407      0.55046     -0.61      0.545
## Year2004     -0.17671      0.56607     -0.31      0.755
## Year2005     -0.19554      0.54596     -0.36      0.721
## Year2006      0.25216      0.56500      0.45      0.656
## Year2007      0.05844      0.53789      0.11      0.914
## Year2008     -0.38445      0.54134     -0.71      0.479
## Year2009     -0.07166      0.54402     -0.13      0.895
## Year2010     -0.46537      0.54438     -0.85      0.394
## Year2011     -0.10374      0.57107     -0.18      0.856
## Year2012      0.03818      0.56371      0.07      0.946
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.65
## Multiple R-squared:  0.179, Adjusted R-squared:  0.0764
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 186 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.554  0.879  0.959   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      5.03e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.526 1      1.235
## LastAuthorFemale  1.740 1      1.319
## Year              1.620 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
## -1.5016 -0.4117 -0.0102 0.4021 1.6063
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2890 0.5295 2.43 0.016 *
## FirstAuthorFemale1 -0.2281 0.1300 -1.75 0.081 .
## LastAuthorFemale1 -0.0953 0.1325 -0.72 0.473
## Year1997 -0.0400 0.6182 -0.06 0.948
## Year1998 0.1104 0.6136 0.18 0.857
## Year1999 -0.9007 0.5549 -1.62 0.106
## Year2000 -0.0825 0.6225 -0.13 0.895
## Year2001 0.0149 0.5628 0.03 0.979
## Year2002 -0.2632 0.5897 -0.45 0.656
## Year2003 -0.3643 0.5601 -0.65 0.516
## Year2004 -0.2222 0.5687 -0.39 0.696
## Year2005 -0.2683 0.5530 -0.49 0.628
```

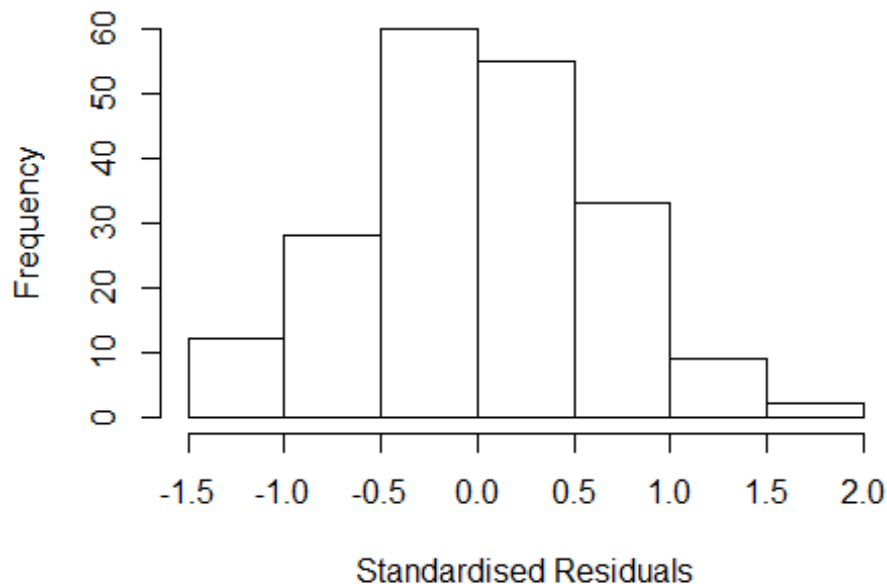


```

## Year2006          0.2126      0.5685      0.37      0.709
## Year2007          0.0375      0.5438      0.07      0.945
## Year2008         -0.4016      0.5506     -0.73      0.467
## Year2009         -0.0559      0.5493     -0.10      0.919
## Year2010         -0.4986      0.5476     -0.91      0.364
## Year2011         -0.1913      0.5567     -0.34      0.732
## Year2012          0.0201      0.5653      0.04      0.972
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.652
## Multiple R-squared:  0.148, Adjusted R-squared:  0.0625
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 183 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.523  0.871  0.956  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      5.03e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
##   nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##   trace.lev      mts      compute.rd
##      0          1000          0
##           psi          subsampling          cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.237 1          1.112
## Year              1.237 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
## -1.48450 -0.40597 -0.00148  0.40241  1.61200
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.27725     0.55634    2.30   0.023 *
## FirstAuthorFemale1 -0.26845     0.11722   -2.29   0.023 *
## Year1997        -0.00803     0.63481   -0.01   0.990
## Year1998         0.11392     0.64185    0.18   0.859
## Year1999        -0.88893     0.58051   -1.53   0.127
## Year2000        -0.06690     0.64452   -0.10   0.917
## Year2001         0.01323     0.58909    0.02   0.982
## Year2002        -0.25785     0.61683   -0.42   0.676
## Year2003        -0.36571     0.58528   -0.62   0.533
## Year2004        -0.22052     0.59351   -0.37   0.711
## Year2005        -0.26225     0.57861   -0.45   0.651
## Year2006         0.20725     0.59123    0.35   0.726
```

```

## Year2007          0.04769    0.56965    0.08    0.933
## Year2008         -0.38562    0.57752   -0.67    0.505
## Year2009         -0.05685    0.57478   -0.10    0.921
## Year2010         -0.49379    0.57341   -0.86    0.390
## Year2011         -0.19641    0.58025   -0.34    0.735
## Year2012          0.01761    0.58985    0.03    0.976
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.651
## Multiple R-squared:  0.146, Adjusted R-squared:  0.0658
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 183 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.520  0.872  0.957  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      5.03e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## 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
## -1.4957 -0.4581  0.0122  0.4280  1.6442

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2626    0.5847    2.16  0.032 *
## LastAuthorFemale1 -0.1984    0.1201   -1.65  0.100
## Year1997          -0.1276    0.7173   -0.18  0.859
## Year1998           0.1478    0.6561    0.23  0.822
## Year1999          -0.8743    0.6078   -1.44  0.152
## Year2000          -0.0721    0.6825   -0.11  0.916
## Year2001           0.0104    0.6174    0.02  0.987
## Year2002          -0.2916    0.6378   -0.46  0.648
## Year2003          -0.4003    0.6158   -0.65  0.516
## Year2004          -0.2640    0.6219   -0.42  0.672
## Year2005          -0.2797    0.6066   -0.46  0.645
## Year2006           0.2332    0.6217    0.38  0.708
## Year2007           0.0488    0.5964    0.08  0.935
## Year2008          -0.3995    0.6010   -0.66  0.507
## Year2009          -0.0706    0.6013   -0.12  0.907
## Year2010          -0.5017    0.6023   -0.83  0.406
## Year2011          -0.1757    0.6121   -0.29  0.774
## Year2012           0.0487    0.6176    0.08  0.937
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.658
## Multiple R-squared:  0.134, Adjusted R-squared:  0.0523
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 183 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.512  0.876  0.953  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      5.03e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 199"
## [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 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    3    2    4    3    4    4    4    5    4    2   13   10   10   18    7
## 2012
##    8
##
## 1996 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    2    3    1    0    4    0    5    3    2   11    9    8   11    3
## 2012
##    3
##
## 1996 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    1    3    1    0    3    0    5    2    2    9    8    8    6    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: 3.5035693237132"

## Warning in wilcox.test.default(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: 3.87298334620742"
## [1] "Male last author team size 2018 geometric mean: 3.30192724889463"

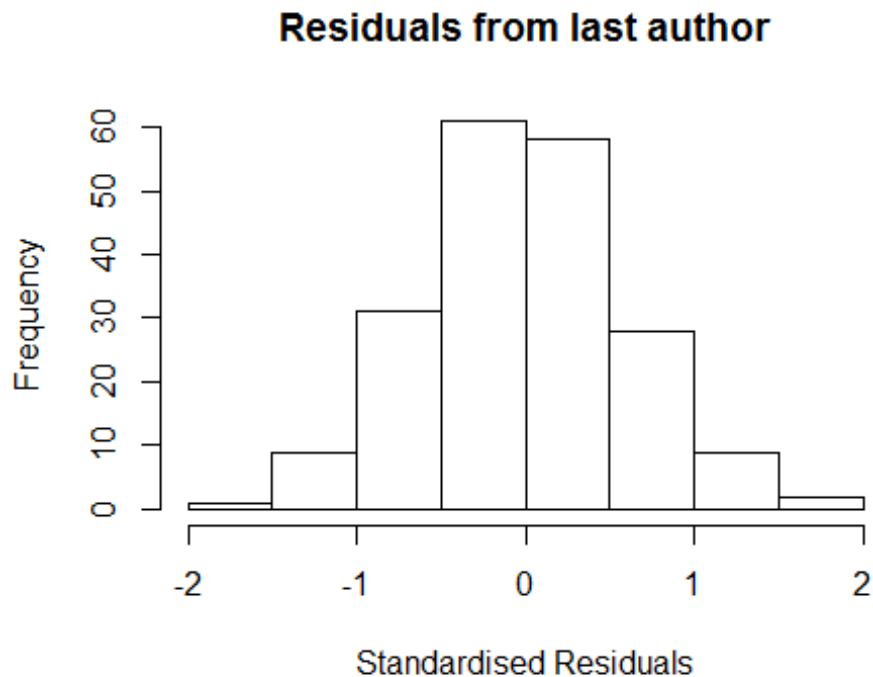
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7, 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"

```

```
## 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 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 +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.005 -0.322  0.051  0.287  0.841
##
```

```

## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3870    0.0000      Inf < 2e-16 ***
## FirstAuthorFemale1 -0.3949    0.2694    -1.47  0.15185
## LastAuthorFemale1 -0.0297    0.8005    -0.04  0.97059
## UniqueAuthors2      0.5413    0.2804     1.93  0.06186 .
## UniqueAuthors3      0.5945    0.2882     2.06  0.04687 *
## UniqueAuthors4      0.7604    0.3142     2.42  0.02102 *
## UniqueAuthors5      1.3687    0.3515     3.89  0.00044 ***
## Year1998           -1.0720    0.0000    -Inf < 2e-16 ***
## Year1999           -0.9628    0.4113    -2.34  0.02523 *
## Year2000           -0.8483    0.2804    -3.03  0.00470 **
## Year2002           -1.4646    0.3021    -4.85  2.7e-05 ***
## Year2004           -0.6303    0.3446    -1.83  0.07614 .
## Year2005           -0.9104    0.2720    -3.35  0.00200 **
## Year2006           -0.9568    0.3845    -2.49  0.01790 *
## Year2007           -0.5365    0.3380    -1.59  0.12166
## Year2008           -0.8143    0.2898    -2.81  0.00816 **
## Year2009           -0.3255    0.2404    -1.35  0.18473
## Year2010           -0.9311    0.3413    -2.73  0.01001 *
## Year2011           -0.8602    0.4568    -1.88  0.06827 .
## Year2012           -1.1399    0.4367    -2.61  0.01337 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.605
## Multiple R-squared:  0.421, Adjusted R-squared:  0.0976
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 47 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.764  0.927  0.965   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           1.85e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01           5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500           50           2           1           1000           200
## trace.lev      mts      compute.rd
##           0           1000           0
##           psi           subsampling           cov
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)

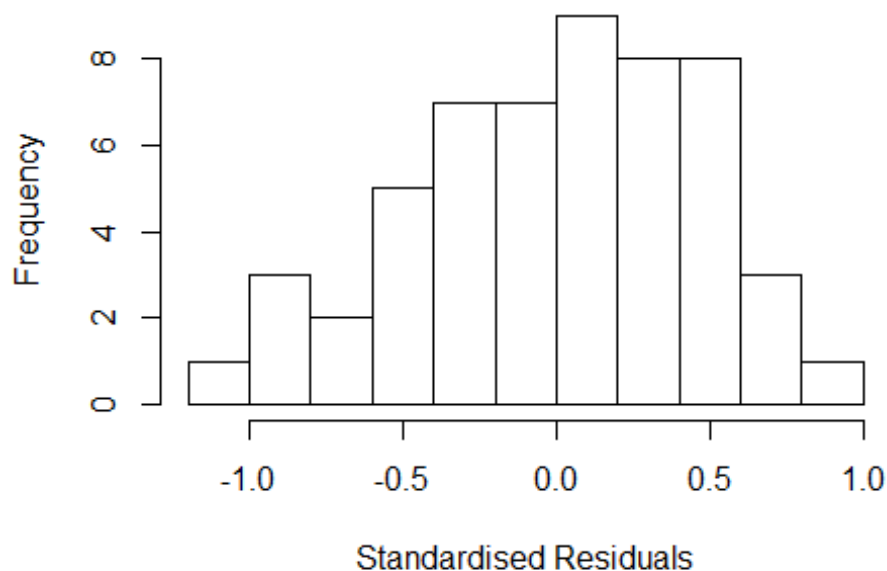
```

```
## [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 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
## -1.0135 -0.3544  0.0419  0.3524  0.9761
##
```



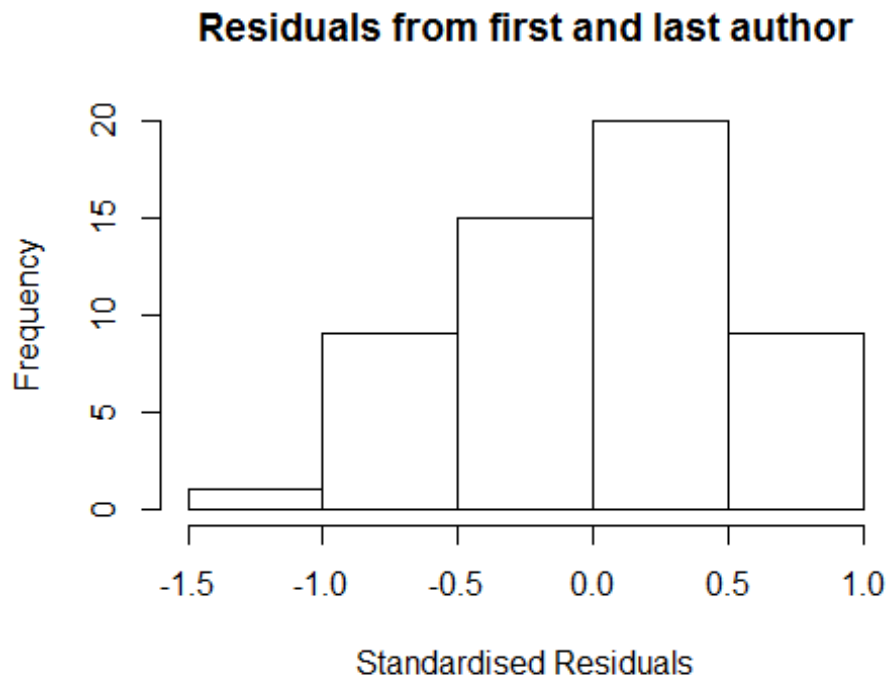
```

## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3870    0.0000      Inf <2e-16 ***
## FirstAuthorFemale1 -0.3002    0.2831    -1.06  0.2956
## LastAuthorFemale1 -0.0399    0.8137    -0.05  0.9611
## Year1998          -1.0720    0.0000    -Inf <2e-16 ***
## Year1999          -0.6530    0.3661    -1.78  0.0824 .
## Year2000          -0.3070    0.0000    -Inf <2e-16 ***
## Year2002          -1.0877    0.1355    -8.03  1e-09 ***
## Year2004          -0.0705    0.2180    -0.32  0.7482
## Year2005          -0.2595    0.0793    -3.27  0.0023 **
## Year2006          -0.6239    0.7550    -0.83  0.4137
## Year2007           0.0402    0.1997     0.20  0.8416
## Year2008          -0.1561    0.1890    -0.83  0.4142
## Year2009           0.1810    0.1821     0.99  0.3266
## Year2010          -0.2176    0.2083    -1.05  0.3026
## Year2011          -0.5895    0.8058    -0.73  0.4689
## Year2012          -0.5604    0.3916    -1.43  0.1606
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.644
## Multiple R-squared:  0.309, Adjusted R-squared:  0.0359
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 46 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.787  0.878  0.962  0.936  0.987  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.85e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          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
## -1.0129 -0.3548  0.0448  0.3525  0.9764
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   1.39e+00   0.00e+00    Inf < 2e-16 ***
## FirstAuthorFemale1 -2.97e-01   2.92e-01  -1.02e+00  0.3151
## Year1998       -1.07e+00   0.00e+00   -Inf < 2e-16 ***
## Year1999       -6.53e-01   3.65e-01  -1.79e+00  0.0815 .
## Year2000       -3.07e-01   3.23e-08  -9.50e+06 < 2e-16 ***
```

```

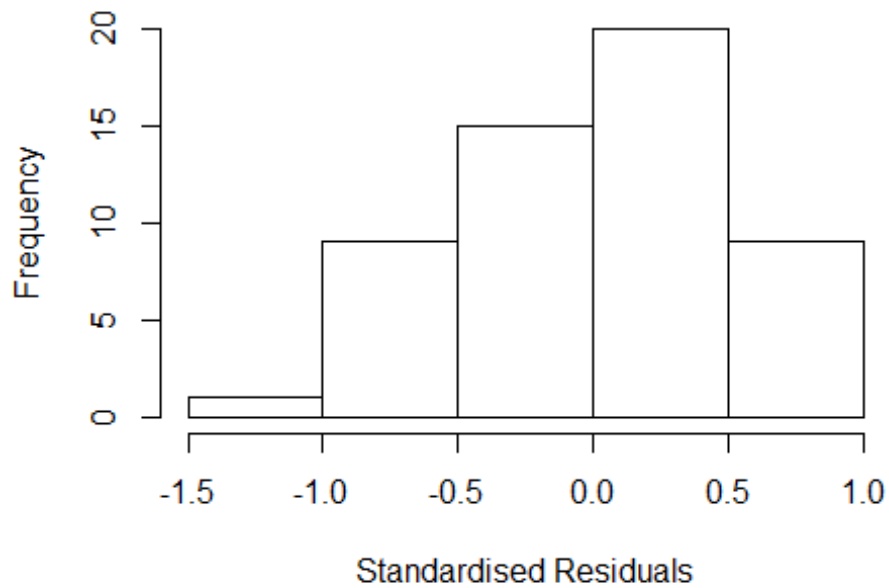
## Year2002      -1.09e+00  1.35e-01 -8.03e+00  8.6e-10 ***
## Year2004      -7.11e-02  2.17e-01 -3.30e-01  0.7447
## Year2005      -2.60e-01  7.93e-02 -3.27e+00  0.0022 **
## Year2006      -6.26e-01  7.49e-01 -8.30e-01  0.4089
## Year2007       3.59e-02  2.07e-01  1.70e-01  0.8630
## Year2008      -1.56e-01  1.89e-01 -8.30e-01  0.4128
## Year2009       1.80e-01  1.82e-01  9.90e-01  0.3277
## Year2010      -2.18e-01  2.08e-01 -1.05e+00  0.3016
## Year2011      -5.89e-01  8.01e-01 -7.40e-01  0.4664
## Year2012      -5.72e-01  4.23e-01 -1.35e+00  0.1844
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.648
## Multiple R-squared:  0.308, Adjusted R-squared:  0.0602
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 46 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.790  0.882  0.963   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      1.85e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          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

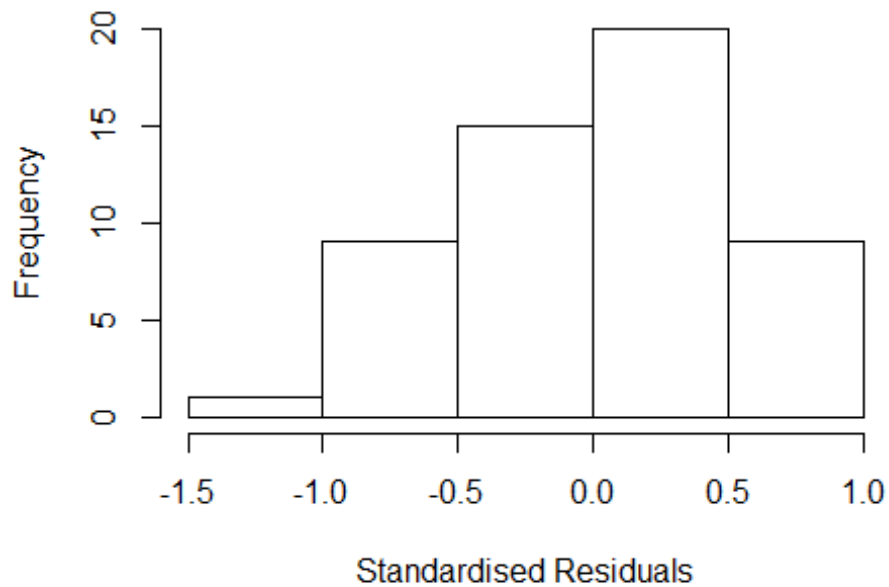
```

Residuals from first author



```
##          GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale  NaN  1          NaN
## Year              NaN 13          NaN
```

Residuals from last author



```

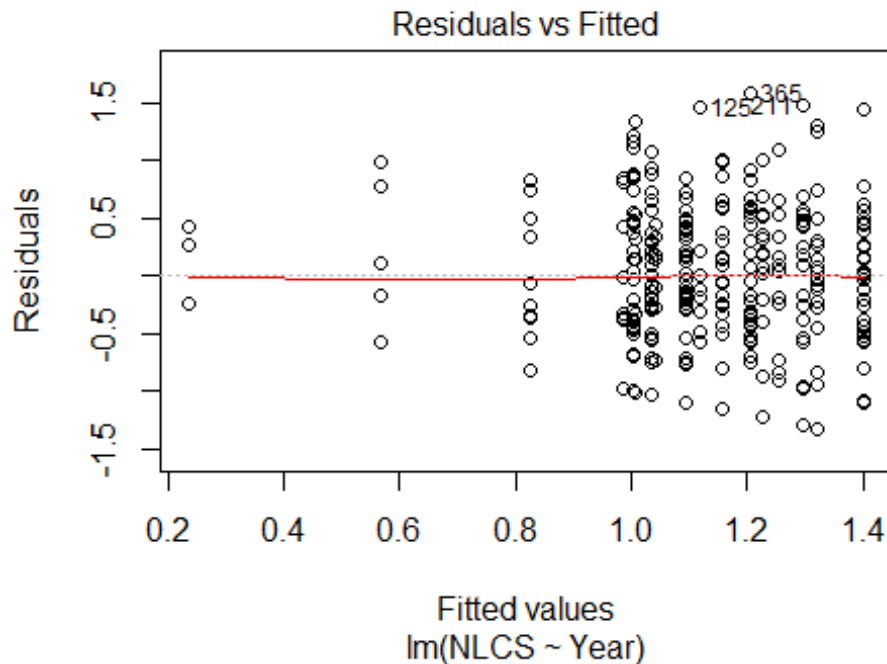
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0512 -0.3879 0.0513 0.3623 1.0186
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3870 0.0000 Inf < 2e-16 ***
## LastAuthorFemale1 0.0523 0.7438 0.07 0.9443
## Year1998 -1.0720 0.0000 -Inf < 2e-16 ***
## Year1999 -0.6538 0.3627 -1.80 0.0792 .
## Year2000 -0.3070 0.0000 -Inf < 2e-16 ***
## Year2002 -1.0877 0.1353 -8.04 8.4e-10 ***
## Year2004 -0.0736 0.2175 -0.34 0.7370
## Year2005 -0.2595 0.0792 -3.27 0.0022 **
## Year2006 -0.7740 0.5160 -1.50 0.1416
## Year2007 -0.0248 0.2400 -0.10 0.9183
## Year2008 -0.1986 0.1785 -1.11 0.2729
## Year2009 0.1429 0.1918 0.75 0.4606
## Year2010 -0.2178 0.2068 -1.05 0.2986
## Year2011 -0.5895 0.7815 -0.75 0.4552
## Year2012 -0.6797 0.3384 -2.01 0.0516 .
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.667
## Multiple R-squared: 0.292, Adjusted R-squared: 0.038
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~ = 1. The remaining 47 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.787 0.913 0.957 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 1.85e-03 1.82e-12
## warn.limit.reject warn.limit.meanrw
## 5.00e-01 5.00e-01
## nResample max.it best.r.s k.fast.s k.max maxit.scale
## 500 50 2 1 1000 200

```

```

## trace.lev      mts compute.rd
##           0      1000         0
##           psi      subsampling      cov
##           "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 54"
## [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
##    7    7    7   12   12   23   16   13   19   29   24   37   39   40   50
## 2011 2012
##   56   42
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7    5    7    9    7   12   10    8   15   22   18   29   27   31   33
## 2011 2012
##   40   28
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    5    5    7    7   12    8    7   15   19   17   25   21   24   27
## 2011 2012
##   34   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 = 11, df = 16, p-value = 0.8

```



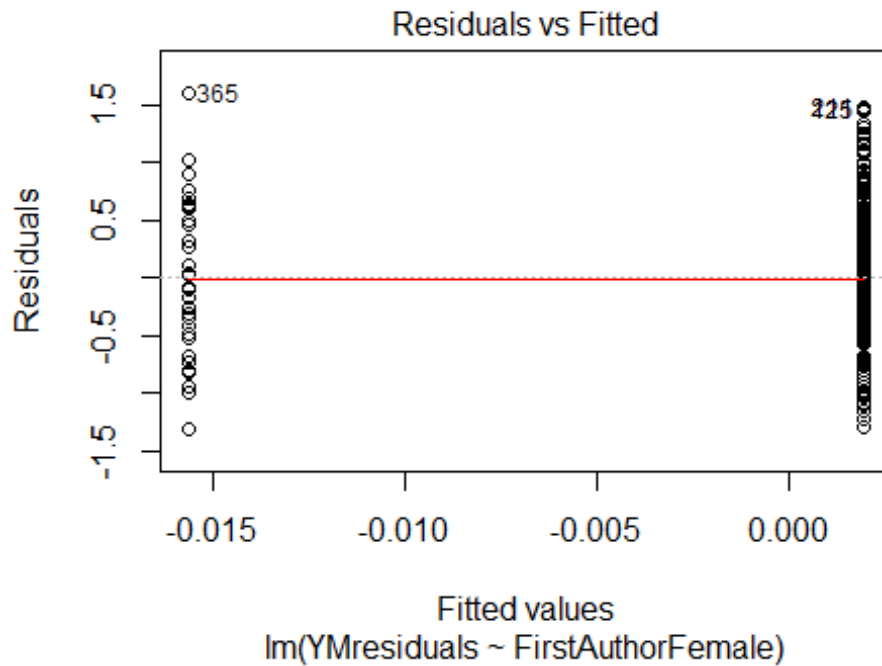
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.8, df = 1, p-value = 0.2

## [1] "Female first author team size 2018 geometric mean: 2.88449914061482"
## [1] "Male first author team size 2018 geometric mean: 2.24531414385029"

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

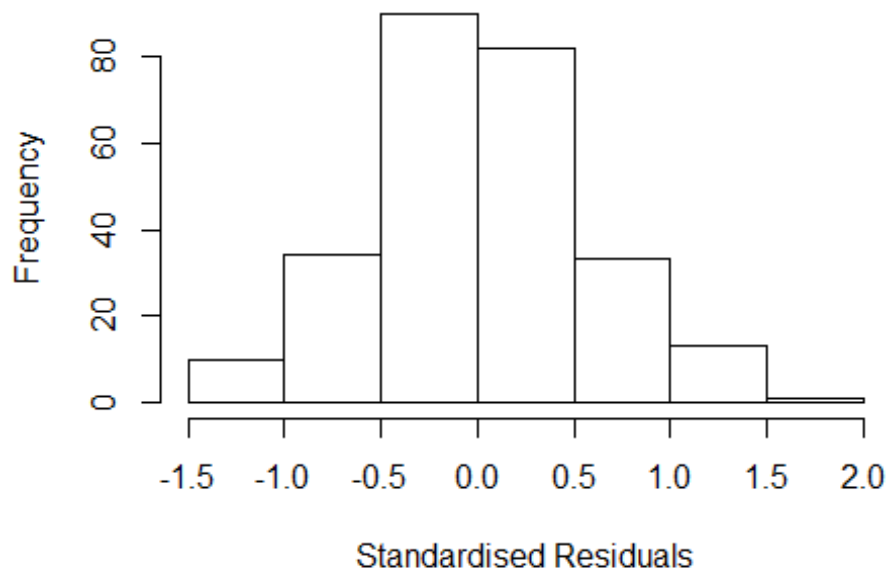
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 27, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.46410161513775"
## [1] "Male last author team size 2018 geometric mean: 2.22683496772043"

## Warning in wilcox.test.default(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 = 0.09
## 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 1.268 1          1.126
## LastAuthorFemale  1.567 1          1.252
## UniqueAuthors    2.865 4          1.141
## Year              4.464 16         1.048
```


Residuals from first and last author and team size



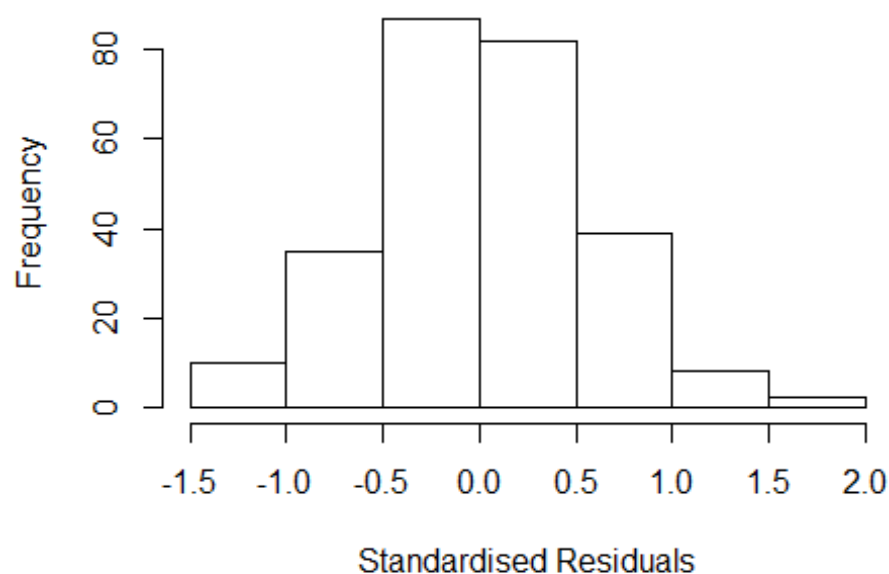
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.35146 -0.35574 -0.00799 0.38131 1.83067
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5648 0.3457 1.63 0.104
## FirstAuthorFemale1 -0.1408 0.1368 -1.03 0.304
## LastAuthorFemale1 0.0686 0.1535 0.45 0.656
## UniqueAuthors2 -0.0657 0.1184 -0.56 0.579
## UniqueAuthors3 0.1131 0.1228 0.92 0.358
## UniqueAuthors4 0.3147 0.1434 2.19 0.029 *
## UniqueAuthors5 -0.0973 0.2516 -0.39 0.699
## Year1997 -0.3077 0.3747 -0.82 0.412
## Year1998 0.0123 0.4729 0.03 0.979
## Year1999 0.2278 0.4049 0.56 0.574
```

```

## Year2000          0.5169      0.3839      1.35      0.179
## Year2001          0.7117      0.4028      1.77      0.078 .
## Year2002          0.2814      0.4162      0.68      0.500
## Year2003          0.3819      0.3894      0.98      0.328
## Year2004          0.6947      0.4008      1.73      0.084 .
## Year2005          0.7029      0.3855      1.82      0.069 .
## Year2006          0.7218      0.3995      1.81      0.072 .
## Year2007          0.7378      0.3776      1.95      0.052 .
## Year2008          0.3602      0.3864      0.93      0.352
## Year2009          0.5403      0.3733      1.45      0.149
## Year2010          0.2778      0.3774      0.74      0.462
## Year2011          0.5066      0.3695      1.37      0.172
## Year2012          0.6051      0.4054      1.49      0.137
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.533
## Multiple R-squared:  0.175, Adjusted R-squared:  0.0992
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 245 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.214  0.872  0.953   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      3.80e-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 1.166 1      1.080
## LastAuthorFemale  1.338 1      1.157
## Year              1.530 16      1.013

```

Residuals from first and last author



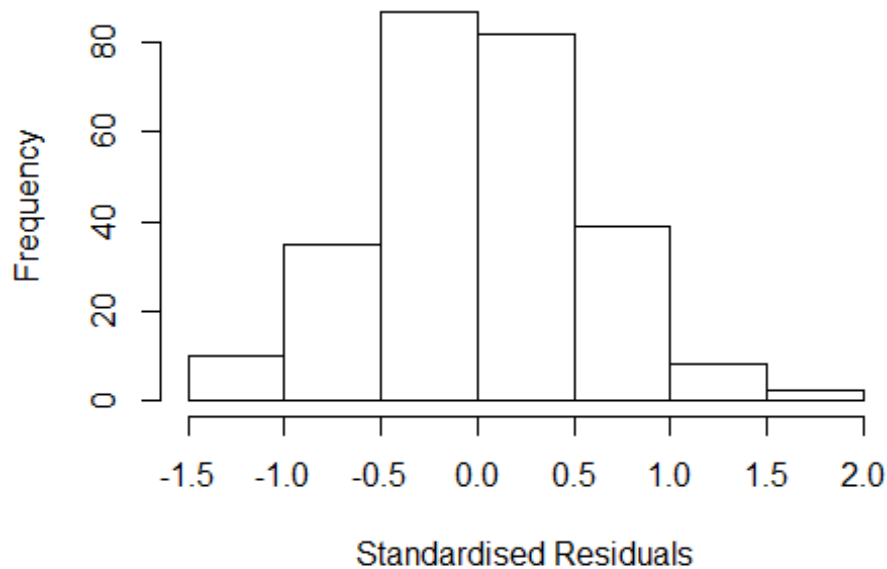
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3076 -0.3382 -0.0111 0.4008 1.6786
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.63590 0.29633 2.15 0.033 *
## FirstAuthorFemale1 -0.14534 0.13910 -1.04 0.297
## LastAuthorFemale1 0.00907 0.14371 0.06 0.950
## Year1997 -0.40561 0.32560 -1.25 0.214
## Year1998 0.02746 0.51110 0.05 0.957
## Year1999 0.14404 0.35169 0.41 0.682
## Year2000 0.44343 0.33084 1.34 0.181
## Year2001 0.62546 0.34905 1.79 0.074 .
## Year2002 0.24990 0.36826 0.68 0.498
## Year2003 0.29669 0.31430 0.94 0.346
## Year2004 0.65617 0.34787 1.89 0.060 .
## Year2005 0.70524 0.33221 2.12 0.035 *
```

```

## Year2006          0.67171      0.35249      1.91      0.058 .
## Year2007          0.67328      0.31688      2.12      0.035 *
## Year2008          0.27811      0.31828      0.87      0.383
## Year2009          0.52718      0.31514      1.67      0.096 .
## Year2010          0.30841      0.32268      0.96      0.340
## Year2011          0.47811      0.30695      1.56      0.121
## Year2012          0.55594      0.33975      1.64      0.103
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.563
## Multiple R-squared:  0.135, Adjusted R-squared:  0.0712
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 242 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.353  0.872  0.953  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      3.80e-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 1.15 1      1.073
## 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
## -1.3072 -0.3391 -0.0117 0.4004 1.6755
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6360 0.2961 2.15 0.033 *
## FirstAuthorFemale1 -0.1439 0.1385 -1.04 0.300
## Year1997 -0.4057 0.3254 -1.25 0.214
## Year1998 0.0305 0.5108 0.06 0.952
## Year1999 0.1440 0.3515 0.41 0.682
## Year2000 0.4431 0.3306 1.34 0.181
## Year2001 0.6262 0.3480 1.80 0.073 .
## Year2002 0.2509 0.3664 0.68 0.494
## Year2003 0.2964 0.3141 0.94 0.346
## Year2004 0.6558 0.3476 1.89 0.060 .
## Year2005 0.7070 0.3287 2.15 0.032 *
## Year2006 0.6712 0.3522 1.91 0.058 .
```

```

## Year2007          0.6733      0.3168      2.13      0.035 *
## Year2008          0.2789      0.3177      0.88      0.381
## Year2009          0.5277      0.3151      1.68      0.095 .
## Year2010          0.3095      0.3214      0.96      0.336
## Year2011          0.4791      0.3063      1.56      0.119
## Year2012          0.5558      0.3397      1.64      0.103
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.563
## Multiple R-squared:  0.135, Adjusted R-squared:  0.0748
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 242 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.357  0.872  0.953  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      3.80e-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 1.33  1      1.153
## Year              1.33 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
## -1.33868 -0.35463 -0.00943  0.41384  1.71340

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.6357    0.2967   2.14  0.033 *
## LastAuthorFemale1 -0.0191    0.1435  -0.13  0.894
## Year1997          -0.4054    0.3260  -1.24  0.215
## Year1998          -0.0071    0.5433  -0.01  0.990
## Year1999           0.1441    0.3520   0.41  0.683
## Year2000           0.4203    0.3308   1.27  0.205
## Year2001           0.6199    0.3532   1.76  0.080 .
## Year2002           0.2083    0.3735   0.56  0.578
## Year2003           0.2754    0.3121   0.88  0.378
## Year2004           0.6569    0.3482   1.89  0.060 .
## Year2005           0.7030    0.3330   2.11  0.036 *
## Year2006           0.6336    0.3514   1.80  0.073 .
## Year2007           0.6592    0.3159   2.09  0.038 *
## Year2008           0.2745    0.3191   0.86  0.391
## Year2009           0.5247    0.3149   1.67  0.097 .
## Year2010           0.2989    0.3235   0.92  0.356
## Year2011           0.4666    0.3075   1.52  0.131
## Year2012           0.5404    0.3374   1.60  0.111
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.561
## Multiple R-squared:  0.131, Adjusted R-squared:  0.0709
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 246 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.331  0.881  0.952  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      3.80e-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: 263"
## [1] ""

```

```

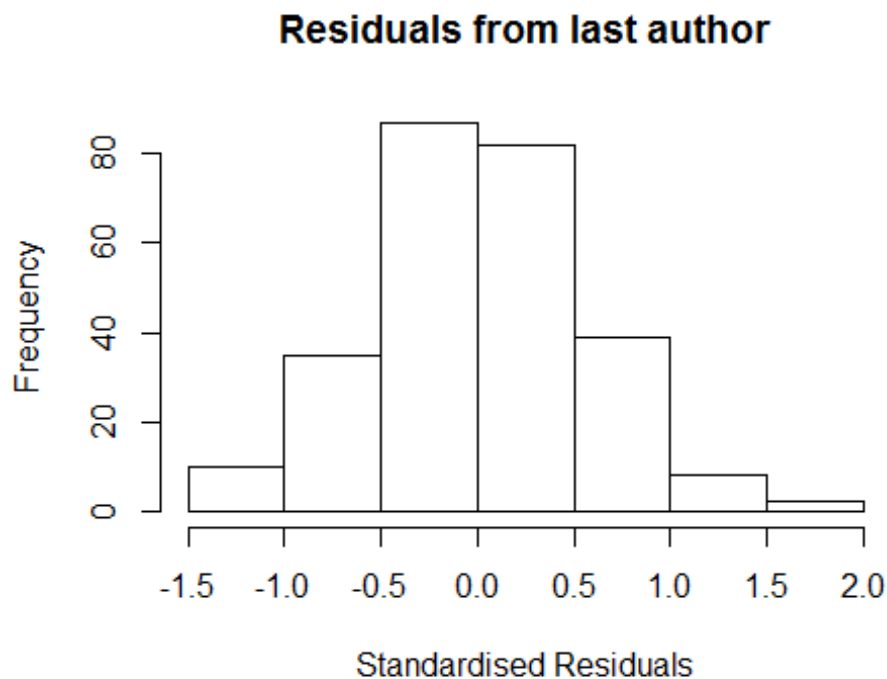
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1800"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2    1    2    2    1    2    1    2    2    4    1    3    3    3    6
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2    1    2    2    1    2    1    2    2    3    1    3    2    2    4
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2    1    2    1    1    2    1    2    2    3    1    3    1    2    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: 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: 27"
## [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 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    2   10    3    3    3    3    1    3    3    8    2   11    8    9    1
## 2012
##    4
##
## 1996 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    2   10    0    1    1    3    1    3    2    7    0    8    6    7    1
## 2012
##    3
##
## 1996 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    2    8    0    1    1    2    1    3    1    7    0    6    6    7    1
## 2012
##    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: 3"

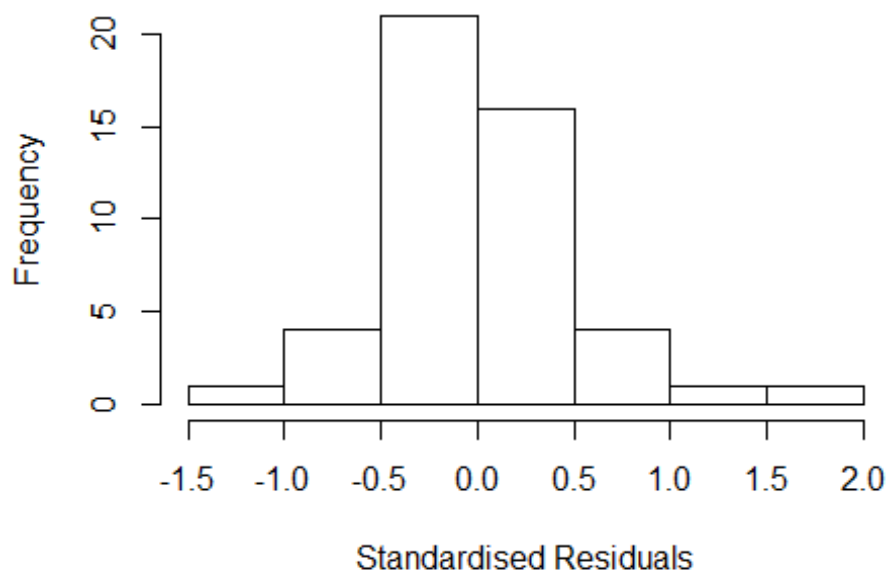
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## 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: NaN"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 5.773e+13  1      7.598e+06
## LastAuthorFemale  5.991e+00  1      2.448e+00
## UniqueAuthors    5.825e+14  3      2.890e+02
## Year              4.317e+15 13      3.993e+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:
```

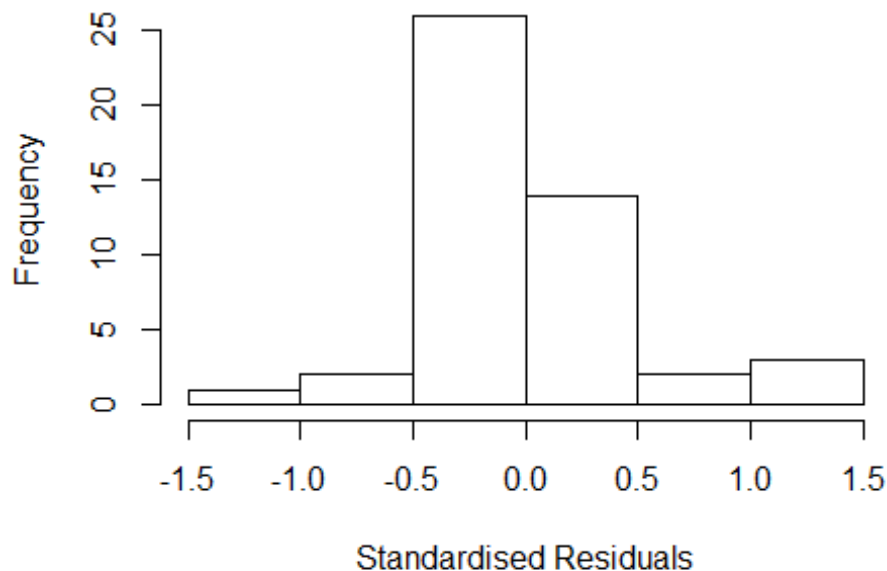
```

##           Min           1Q           Median           3Q           Max
## -1.25e+00 -2.60e-01 -3.05e-16  3.08e-01  1.50e+00
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.9936      0.4917   2.02   0.053 .
## FirstAuthorFemale1 -0.0946      0.2247  -0.42   0.677
## LastAuthorFemale1 -0.1305      0.2075  -0.63   0.534
## UniqueAuthors2      0.0217      0.2341   0.09   0.927
## UniqueAuthors3     -0.1963      0.2048  -0.96   0.346
## UniqueAuthors4      0.5517      0.7230   0.76   0.452
## Year1998            0.6485      0.4952   1.31   0.201
## Year2000            0.2564      0.4917   0.52   0.606
## Year2001           -0.4796      0.4917  -0.98   0.337
## Year2002           -0.1951      0.5470  -0.36   0.724
## Year2003           -0.6628      0.5497  -1.21   0.238
## Year2004           -0.1545      0.5530  -0.28   0.782
## Year2005           -0.5596      0.4917  -1.14   0.264
## Year2006            0.2500      0.5391   0.46   0.646
## Year2008           -0.3277      0.4954  -0.66   0.514
## Year2009            0.2669      0.6094   0.44   0.665
## Year2010            0.2583      0.5437   0.48   0.638
## Year2011            1.3234      0.4917   2.69   0.012 *
## Year2012           -0.0935      0.5697  -0.16   0.871
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.505
## Multiple R-squared:  0.464, Adjusted R-squared:  0.132
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 40 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.355  0.910  0.959  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           2.08e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01           5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500           50           2           1           1000           200
## trace.lev      mts      compute.rd
##           0           1000           0
##           psi           subsampling           cov
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"

```

```
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 30.635  1          5.535
## LastAuthorFemale  5.358  1          2.315
## Year              65.164 13          1.174
```

Residuals from first and last author



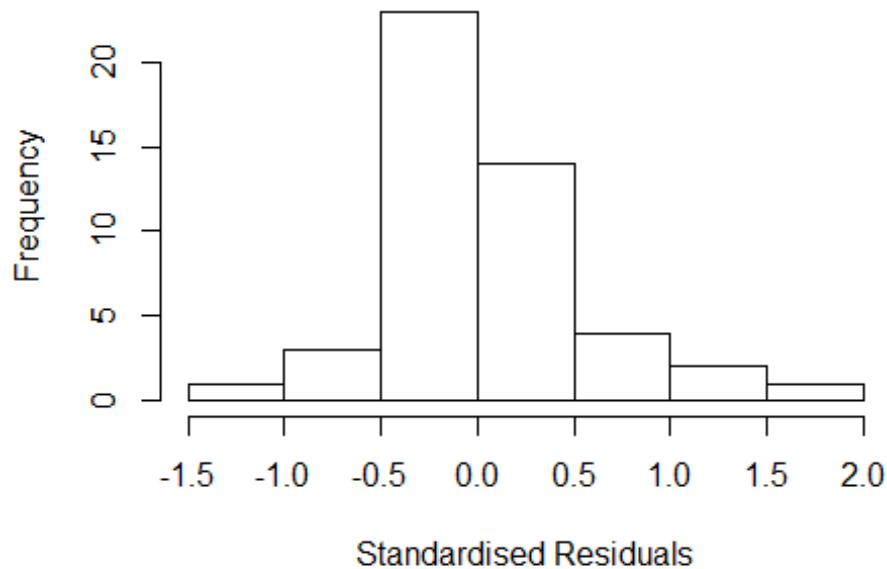
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.237315 -0.271994 -0.000289  0.321138  1.456236
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0045     0.4828   2.08    0.046 *
## FirstAuthorFemale1 -0.0801     0.2276  -0.35    0.727
## LastAuthorFemale1 -0.1918     0.2199  -0.87    0.390
## Year1998          0.6254     0.4983   1.26    0.219
```

```

## Year2000          0.2455      0.4828      0.51      0.615
## Year2001         -0.4905      0.4828     -1.02      0.317
## Year2002         -0.2060      0.5384     -0.38      0.705
## Year2003         -0.6664      0.5333     -1.25      0.220
## Year2004         -0.1600      0.5466     -0.29      0.772
## Year2005         -0.5705      0.4828     -1.18      0.246
## Year2006          0.3491      0.5243      0.67      0.510
## Year2008         -0.2919      0.5191     -0.56      0.578
## Year2009          0.3033      0.5475      0.55      0.584
## Year2010          0.2328      0.5470      0.43      0.673
## Year2011          1.3125      0.4828      2.72      0.011 *
## Year2012         -0.0935      0.5784     -0.16      0.873
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.524
## Multiple R-squared:  0.44,   Adjusted R-squared:  0.177
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 38 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.420  0.921  0.958  0.913  0.981  0.992
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      2.08e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -6.178e+14 1      NaN
## Year              -6.178e+14 13      NaN

```

Residuals from first author



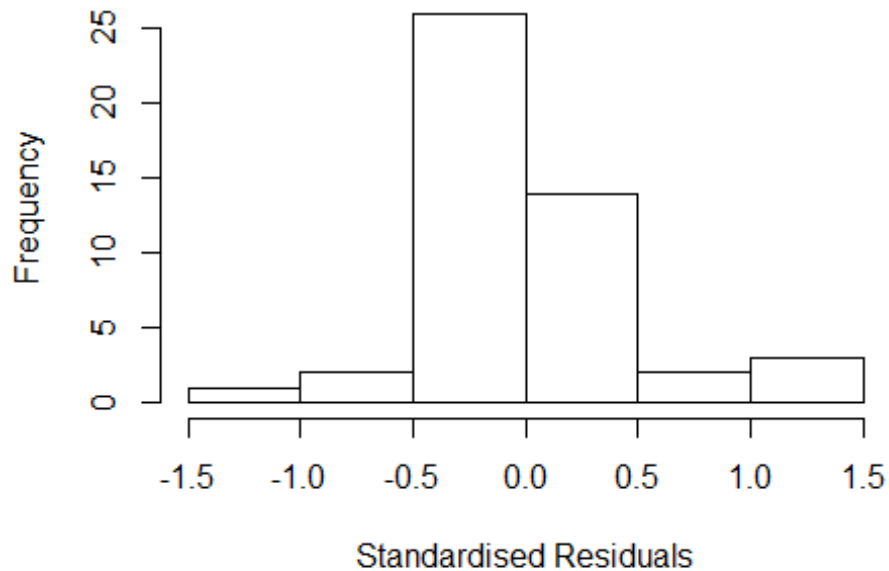
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.19e+00 -2.59e-01 5.55e-17 2.82e-01 1.50e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0045 0.4861 2.07 0.047 *
## FirstAuthorFemale1 -0.2203 0.1455 -1.51 0.140
## Year1998 0.6020 0.5039 1.19 0.241
## Year2000 0.2455 0.4861 0.51 0.617
## Year2001 -0.4905 0.4861 -1.01 0.320
## Year2002 -0.2060 0.5415 -0.38 0.706
## Year2003 -0.5262 0.5036 -1.04 0.304
## Year2004 -0.0640 0.5118 -0.13 0.901
## Year2005 -0.5705 0.4861 -1.17 0.249
## Year2006 0.3373 0.5259 0.64 0.526
## Year2008 -0.3050 0.5197 -0.59 0.561
## Year2009 0.2578 0.5585 0.46 0.647
```

```

## Year2010          0.1904      0.5379      0.35      0.726
## Year2011          1.3125      0.4861      2.70      0.011 *
## Year2012         -0.0935      0.5817     -0.16      0.873
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.517
## Multiple R-squared:  0.436, Adjusted R-squared:  0.197
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 41 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.380  0.911  0.955  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      2.08e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## 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 13      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
## -1.24e+00 -2.56e-01 -5.55e-17 3.36e-01 1.46e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0045 0.4911 2.05 0.049 *
## LastAuthorFemale1 -0.2311 0.1565 -1.48 0.149
## Year1998 0.6303 0.5057 1.25 0.221
## Year2000 0.2455 0.4911 0.50 0.620
## Year2001 -0.4905 0.4911 -1.00 0.325
## Year2002 -0.2060 0.5463 -0.38 0.709
## Year2003 -0.7465 0.4911 -1.52 0.138
## Year2004 -0.2160 0.5218 -0.41 0.682
## Year2005 -0.5705 0.4911 -1.16 0.254
## Year2006 0.3363 0.5285 0.64 0.529
## Year2008 -0.3058 0.5220 -0.59 0.562
## Year2009 0.3026 0.5528 0.55 0.588
```

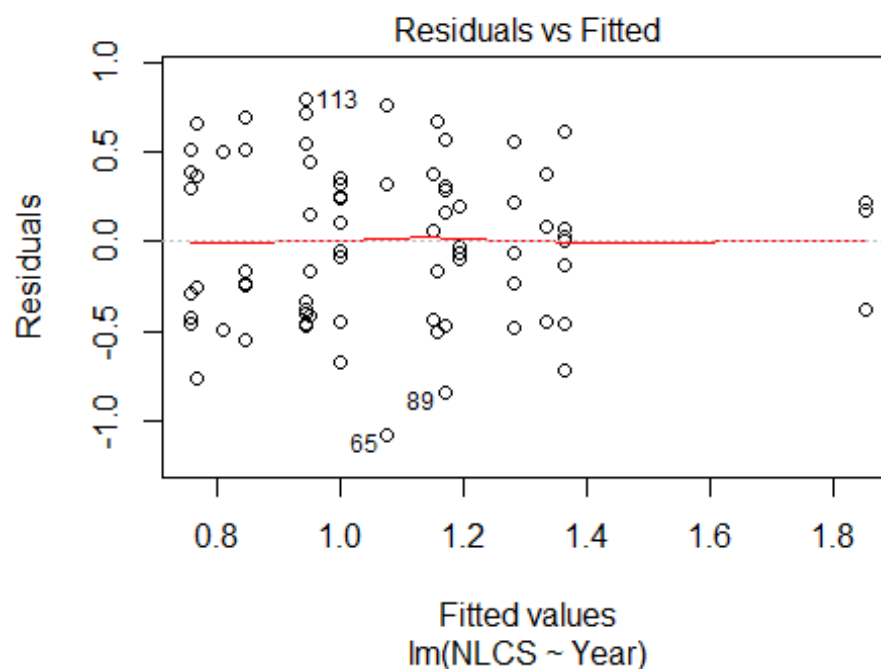


```

## Year2010          0.2375      0.5544      0.43      0.671
## Year2011          1.3125      0.4911      2.67      0.012 *
## Year2012         -0.0935      0.5867     -0.16      0.874
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.508
## Multiple R-squared:  0.445, Adjusted R-squared:  0.21
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 38 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.392  0.918  0.959  0.908  0.981  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      2.08e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 48"
## [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
##    7    3    9    4   11    5    2    4    5    4    7    7   10    9    8
## 2011 2012
##    6    8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    2    8    3    9    3    0    3    3    3    4    6    6    8    5
## 2011 2012
##    4    4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010

```

```
##      5      2      8      2      9      3      0      3      3      3      4      6      5      7      4
## 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 = 11, df = 15, p-value = 0.8
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.41, df = 1, p-value = 0.5

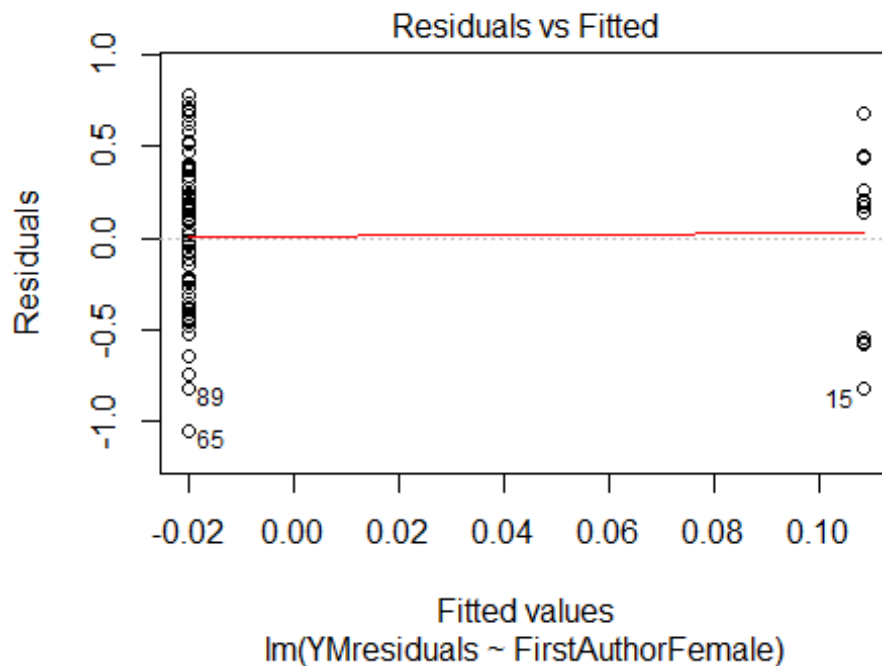
## [1] "Female first author team size 2018 geometric mean: 1"
## [1] "Male first author team size 2018 geometric mean: 1.41421356237309"

## 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 = 2, p-value = 0.8
## 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: 1.41421356237309"

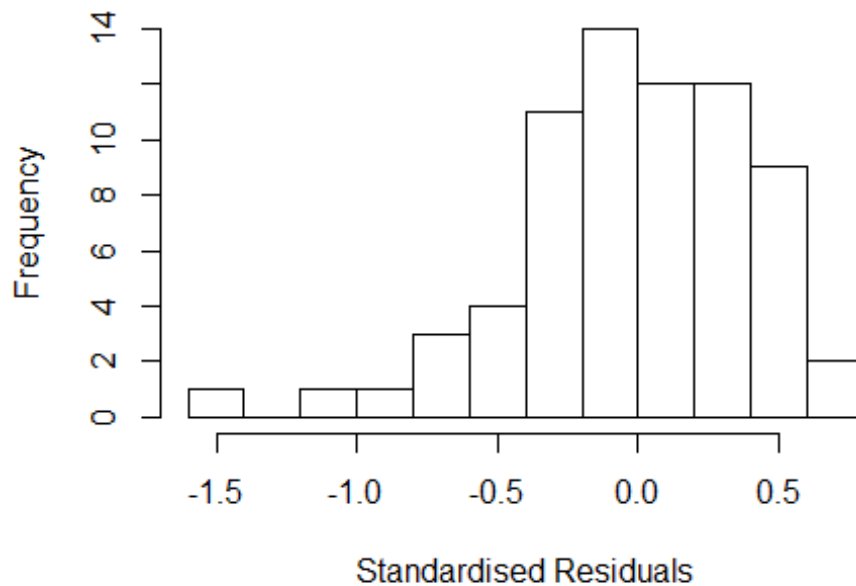
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2, 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 | 4.122 | 1 | 2.030 |
| LastAuthorFemale | 5.121 | 1 | 2.263 |
| UniqueAuthors | 14.480 | 3 | 1.561 |
| Year | 78.030 | 15 | 1.156 |

Residuals from first and last author and team size



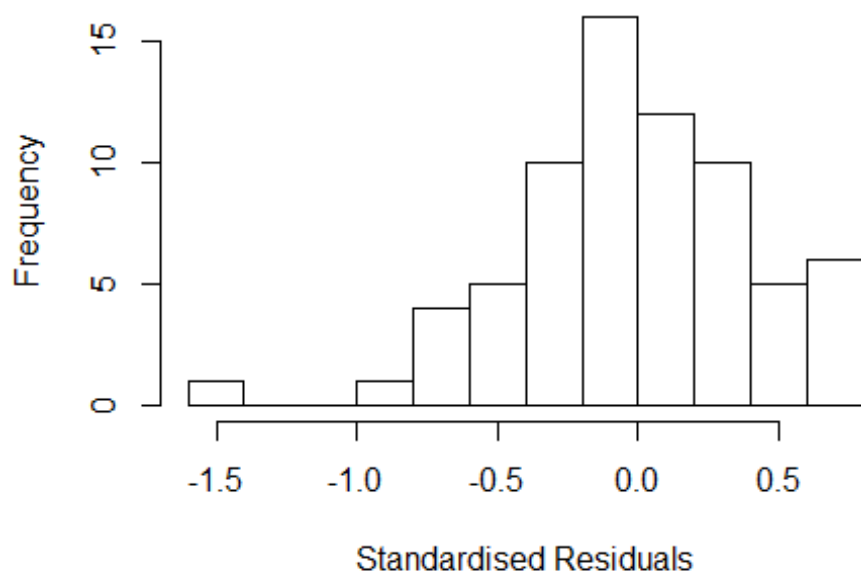
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4274 -0.2473 -0.0033 0.2994 0.6889
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8096 0.2194 3.69 0.00056 ***
## FirstAuthorFemale1 0.2641 0.2535 1.04 0.30262
## LastAuthorFemale1 0.4300 0.2761 1.56 0.12574
## UniqueAuthors2 0.0706 0.1389 0.51 0.61344
## UniqueAuthors3 -0.2366 0.1355 -1.75 0.08700 .
## UniqueAuthors4 0.2680 0.1952 1.37 0.17601
## Year1997 0.0814 0.3414 0.24 0.81259
## Year1998 0.5644 0.2586 2.18 0.03388 *
## Year1999 0.4123 0.4206 0.98 0.33175
## Year2000 0.1118 0.2575 0.43 0.66604
```

```

## Year2001          0.9115      0.4111      2.22  0.03127 *
## Year2003          0.1813      0.2750      0.66  0.51280
## Year2004          0.3377      0.2985      1.13  0.26339
## Year2005          0.5472      0.7083      0.77  0.44348
## Year2006          0.3601      0.2198      1.64  0.10787
## Year2007          0.0271      0.2543      0.11  0.91559
## Year2008          0.2385      0.4099      0.58  0.56330
## Year2009         -0.1208      0.2274     -0.53  0.59756
## Year2010          0.1227      0.2523      0.49  0.62895
## Year2011          0.2686      0.2290      1.17  0.24642
## Year2012         -0.0483      0.4434     -0.11  0.91364
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.426
## Multiple R-squared:  0.42, Adjusted R-squared:  0.183
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 61 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.239  0.908  0.953  0.911  0.987  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.43e-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 3.288 1      1.813
## LastAuthorFemale  6.098 1      2.469
## Year             11.533 15      1.085

```

Residuals from first and last author



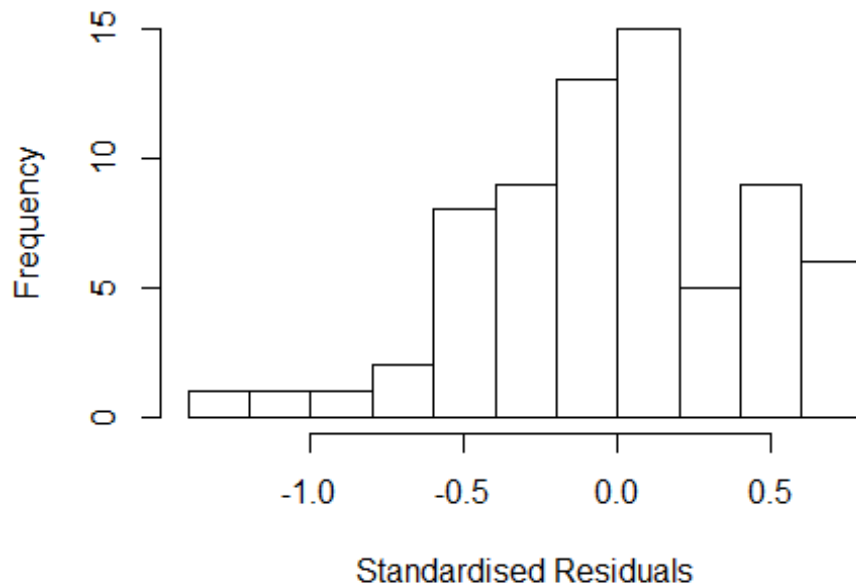
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5129 -0.2686 -0.0076 0.2805 0.7237
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.676072 0.231934 2.91 0.0052 **
## FirstAuthorFemale1 0.179391 0.224871 0.80 0.4286
## LastAuthorFemale1 0.492929 0.288340 1.71 0.0933 .
## Year1997 0.131928 0.558692 0.24 0.8143
## Year1998 0.688296 0.298965 2.30 0.0254 *
## Year1999 0.616428 0.429333 1.44 0.1571
## Year2000 0.286766 0.274122 1.05 0.3003
## Year2001 1.089459 0.447240 2.44 0.0183 *
## Year2003 0.317258 0.294318 1.08 0.2860
## Year2004 0.421809 0.298581 1.41 0.1637
## Year2005 0.836810 0.574467 1.46 0.1512
## Year2006 0.515536 0.238681 2.16 0.0354 *
```

```

## Year2007      0.136237    0.310419    0.44    0.6626
## Year2008      0.449711    0.392284    1.15    0.2569
## Year2009      0.000878    0.255516    0.00    0.9973
## Year2010      0.331730    0.269296    1.23    0.2235
## Year2011      0.571428    0.256702    2.23    0.0304 *
## Year2012      0.120741    0.467509    0.26    0.7972
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.401
## Multiple R-squared:  0.392, Adjusted R-squared:  0.193
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 63 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.123  0.859  0.941  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      1.43e-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 2.814 1      1.678
## Year              2.814 15      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
## -1.31608 -0.26270 -0.00238 0.30956 0.74758
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6828 0.2174 3.14 0.0028 **
## FirstAuthorFemale1 0.2915 0.2171 1.34 0.1851
## Year1997 0.1252 0.4976 0.25 0.8023
## Year1998 0.6738 0.2889 2.33 0.0235 *
## Year1999 0.6097 0.4058 1.50 0.1389
## Year2000 0.2371 0.2603 0.91 0.3666
## Year2001 1.1768 0.2754 4.27 8.1e-05 ***
## Year2003 0.4443 0.4005 1.11 0.2723
## Year2004 0.3746 0.2728 1.37 0.1755
## Year2005 0.6333 0.7486 0.85 0.4014
## Year2006 0.5090 0.2247 2.27 0.0276 *
## Year2007 0.1403 0.2951 0.48 0.6363
```

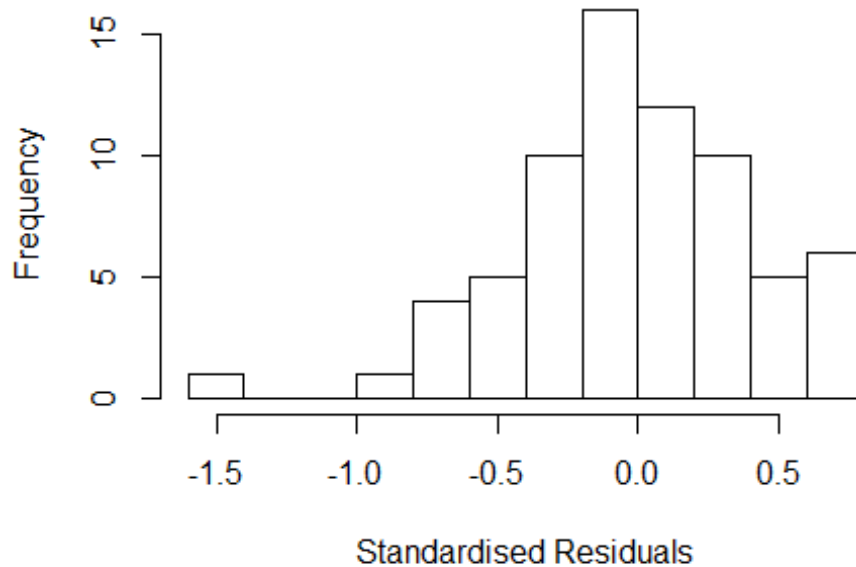


```

## Year2008          0.5006      0.3545      1.41      0.1638
## Year2009          0.0131      0.2567      0.05      0.9595
## Year2010          0.5457      0.2803      1.95      0.0568 .
## Year2011          0.5647      0.2433      2.32      0.0242 *
## Year2012          0.1027      0.4185      0.25      0.8070
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.469
## Multiple R-squared:  0.318, Adjusted R-squared:  0.112
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 62 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.411  0.888  0.949  0.918  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          1.43e-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 4.579 1          2.140
## Year            4.579 15          1.052

```

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
## -1.5245 -0.2453 -0.0466 0.2901 0.7640
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6756 0.2327 2.90 0.0054 **
## LastAuthorFemale1 0.5797 0.3021 1.92 0.0604 .
## Year1997 0.1324 0.5628 0.24 0.8149
## Year1998 0.6905 0.2881 2.40 0.0201 *
## Year1999 0.6169 0.4366 1.41 0.1635
## Year2000 0.3464 0.2671 1.30 0.2003
## Year2001 1.0922 0.5069 2.15 0.0358 *
## Year2003 0.2905 0.2848 1.02 0.3124
## Year2004 0.4816 0.3197 1.51 0.1379
## Year2005 0.8489 0.5456 1.56 0.1257
## Year2006 0.5160 0.2394 2.16 0.0357 *
## Year2007 0.1359 0.3115 0.44 0.6643
```

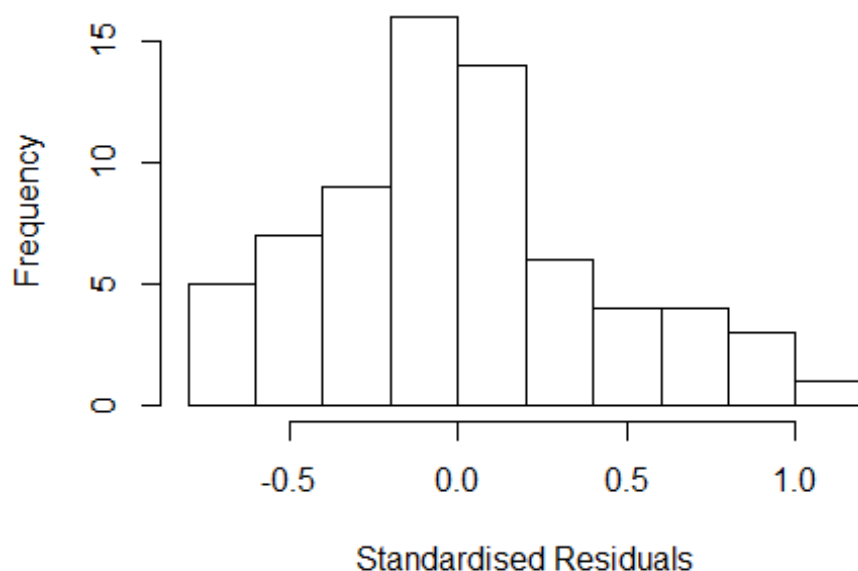
```

## Year2008          0.5278      0.3702      1.43      0.1598
## Year2009          0.0474      0.2671      0.18      0.8597
## Year2010          0.3320      0.2787      1.19      0.2389
## Year2011          0.5719      0.2574      2.22      0.0306 *
## Year2012          0.1221      0.4704      0.26      0.7962
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.397
## Multiple R-squared:  0.387, Adjusted R-squared:  0.201
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 65 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.108  0.862  0.950  0.898  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      1.43e-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 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
##    3    2    4    5    4    6    2    4    5    3    2    4    8    11   15
## 2011 2012
##    9    6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    3    5    2    5    1    4    5    3    2    2    6    9    13
## 2011 2012
##    7    4

```

```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    3    5    2    4    1    4    4    3    2    2    5    7   13
## 2011 2012
##    7    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: 1"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.931e+15 1      54134653
## LastAuthorFemale -4.437e+14 1           NaN
## UniqueAuthors    -8.265e+30 4           NaN
## Year              -4.578e+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))
```

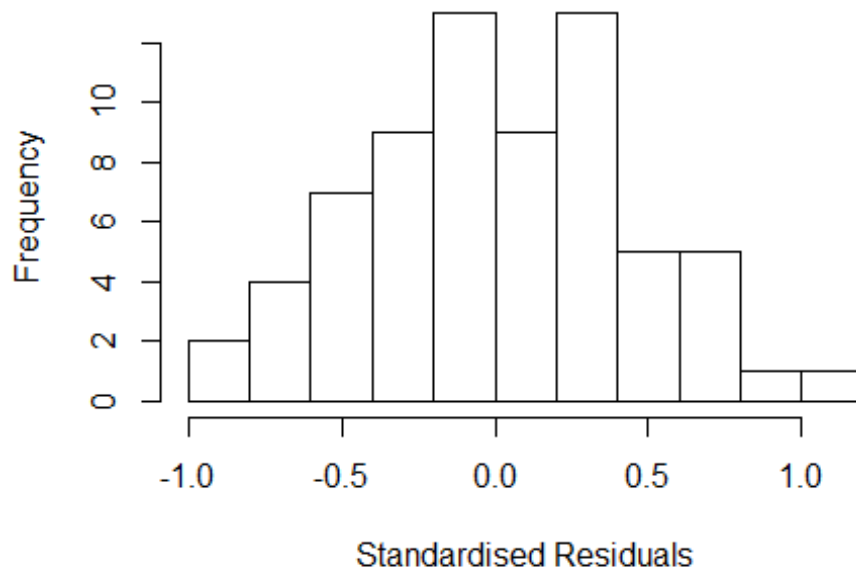
```

## \--> method = "MM"
## Residuals:
##      Min      1Q   Median      3Q      Max
## -7.70e-01 -2.48e-01 -4.44e-16  2.19e-01  1.06e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.1690      0.1229   1.37  0.17593
## FirstAuthorFemale1 -0.4360      0.2809  -1.55  0.12752
## LastAuthorFemale1   0.7815      0.2720   2.87  0.00613 **
## UniqueAuthors2      0.3378      0.1833   1.84  0.07182 .
## UniqueAuthors3      0.0272      0.2305   0.12  0.90670
## UniqueAuthors4     -0.3482      0.3330  -1.05  0.30120
## UniqueAuthors5      0.7956      0.2738   2.91  0.00561 **
## Year1997            0.3110      0.1229   2.53  0.01491 *
## Year1998            0.5667      0.3772   1.50  0.13984
## Year1999            0.7032      0.3134   2.24  0.02971 *
## Year2000           -0.0748      0.1639  -0.46  0.65053
## Year2001            1.0190      0.3809   2.68  0.01030 *
## Year2002            0.7710      0.1229   6.27  1.1e-07 ***
## Year2003            0.4879      0.3674   1.33  0.19077
## Year2004            0.7528      0.2609   2.89  0.00594 **
## Year2005            0.2041      0.2086   0.98  0.33281
## Year2006            1.3131      0.2086   6.29  1.0e-07 ***
## Year2007            0.8515      0.2239   3.80  0.00042 ***
## Year2008            0.6786      0.6040   1.12  0.26701
## Year2009            0.1272      0.1974   0.64  0.52235
## Year2010            0.3417      0.2057   1.66  0.10349
## Year2011            0.6674      0.2029   3.29  0.00193 **
## Year2012            0.6995      0.3829   1.83  0.07420 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.432
## Multiple R-squared:  0.424, Adjusted R-squared:  0.148
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 59 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.530  0.856  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      1.45e-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 2: First author gender, Last author gender, Year as
factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 7.586 1      2.754
## LastAuthorFemale 5.319 1      2.306
## Year              8.847 16      1.071
```

Residuals from first and last author



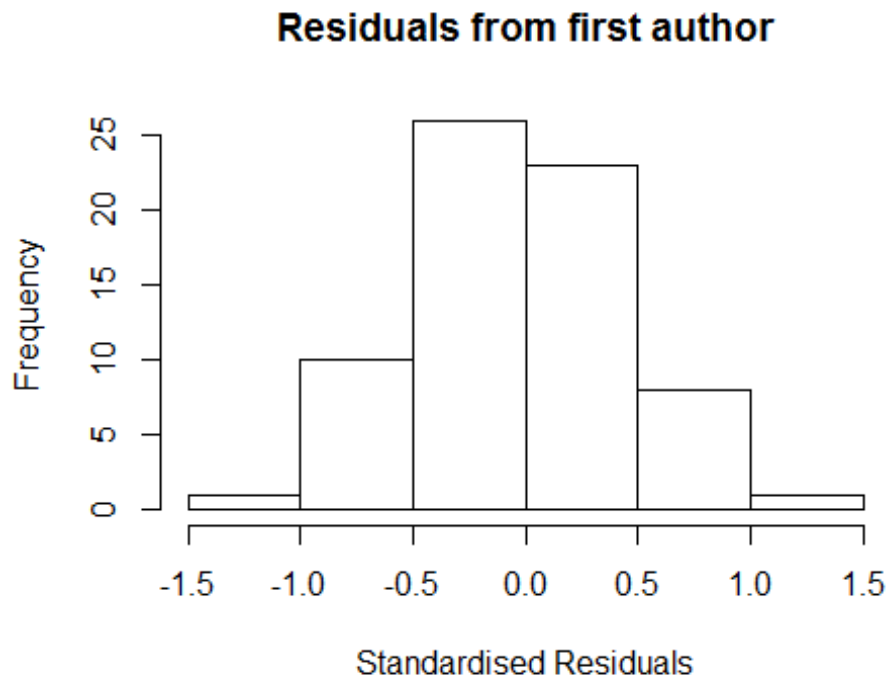
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.27e-01 -3.18e-01 -3.33e-16  2.54e-01  1.20e+00
##
```

```

## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.1690    0.1218   1.39  0.17154
## FirstAuthorFemale1 -0.1604    0.1913  -0.84  0.40578
## LastAuthorFemale1  0.5531    0.1600   3.46  0.00112 **
## Year1997         0.3110    0.1218   2.55  0.01379 *
## Year1998         0.8326    0.3515   2.37  0.02176 *
## Year1999         0.8185    0.2768   2.96  0.00473 **
## Year2000        -0.0984    0.1542  -0.64  0.52651
## Year2001         1.0450    0.3469   3.01  0.00406 **
## Year2002         0.7710    0.1218   6.33  6.7e-08 ***
## Year2003         0.6633    0.2938   2.26  0.02833 *
## Year2004         0.8867    0.2248   3.94  0.00025 ***
## Year2005         0.4279    0.2220   1.93  0.05957 .
## Year2006         1.3442    0.2152   6.25  9.0e-08 ***
## Year2007         0.8515    0.2201   3.87  0.00032 ***
## Year2008         0.8757    0.5598   1.56  0.12406
## Year2009         0.2661    0.1841   1.44  0.15471
## Year2010         0.3356    0.1755   1.91  0.06154 .
## Year2011         0.7658    0.2321   3.30  0.00179 **
## Year2012         0.9182    0.2991   3.07  0.00346 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.523
## Multiple R-squared:  0.355, Adjusted R-squared:  0.123
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~ = 1. The remaining 61 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.579  0.917  0.966  0.932  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      1.45e-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 4.225 1 2.055
## Year 4.225 16 1.046
```



```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1077 -0.2994 -0.0173 0.2928 1.2292
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.1690 0.1214 1.39 0.16994
## FirstAuthorFemale1 0.0499 0.1846 0.27 0.78800
## Year1997 0.3110 0.1214 2.56 0.01341 *
## Year1998 1.0464 0.2722 3.84 0.00034 ***
## Year1999 0.7694 0.2599 2.96 0.00465 **
## Year2000 0.0731 0.2335 0.31 0.75568
## Year2001 1.0535 0.3370 3.13 0.00293 **
## Year2002 0.7710 0.1214 6.35 5.8e-08 ***
## Year2003 0.7991 0.3599 2.22 0.03085 *
```

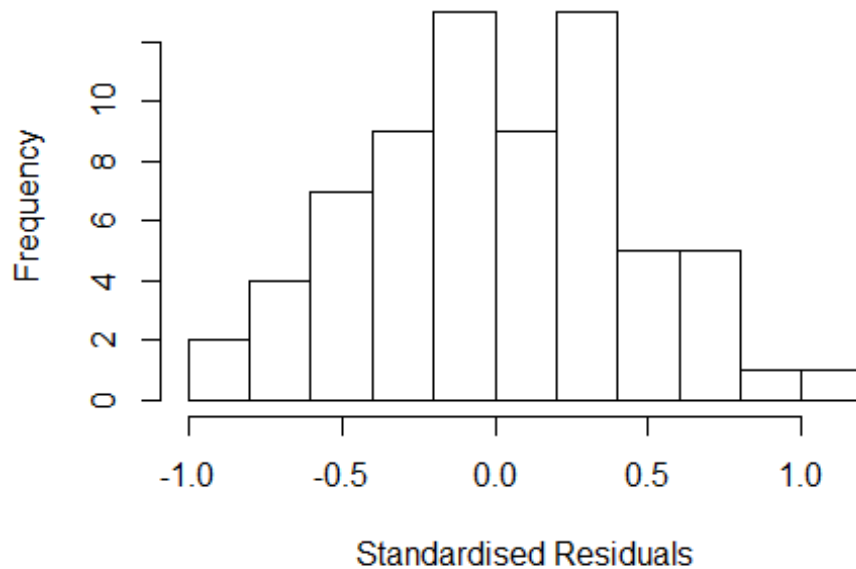


```

## Year2004          0.9710      0.2622      3.70  0.00052 ***
## Year2005          0.4287      0.2202      1.95  0.05707 .
## Year2006          1.2391      0.1683      7.36  1.5e-09 ***
## Year2007          0.8515      0.2187      3.89  0.00029 ***
## Year2008          0.8438      0.4553      1.85  0.06964 .
## Year2009          0.3111      0.2011      1.55  0.12798
## Year2010          0.4286      0.1803      2.38  0.02120 *
## Year2011          0.6843      0.2493      2.74  0.00834 **
## Year2012          0.8888      0.3421      2.60  0.01222 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.578
## Multiple R-squared:  0.282, Adjusted R-squared:  0.0426
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 62 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.630  0.920  0.970   0.938   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          1.45e-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 8.282 1          2.878
## Year            8.282 16          1.068

```

Residuals from last author



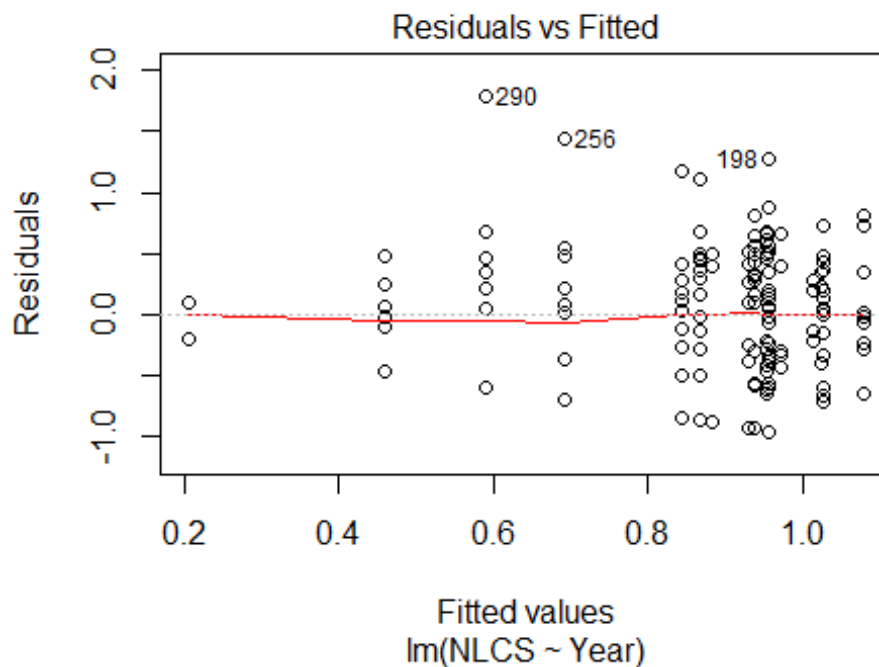
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.08e+00 -3.10e-01 -5.55e-17 2.62e-01 1.24e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.169 0.122 1.39 0.17164
## LastAuthorFemale1 0.464 0.126 3.67 0.00057 ***
## Year1997 0.311 0.122 2.55 0.01377 *
## Year1998 0.865 0.335 2.59 0.01257 *
## Year1999 0.780 0.257 3.04 0.00374 **
## Year2000 -0.134 0.140 -0.96 0.34201
## Year2001 1.044 0.348 3.00 0.00418 **
## Year2002 0.771 0.122 6.33 6.3e-08 ***
## Year2003 0.691 0.296 2.33 0.02370 *
## Year2004 0.868 0.218 3.98 0.00022 ***
## Year2005 0.428 0.222 1.93 0.05979 .
## Year2006 1.264 0.151 8.36 3.9e-11 ***
```

```

## Year2007          0.851      0.220      3.86  0.00032 ***
## Year2008          0.833      0.502      1.66  0.10302
## Year2009          0.256      0.181      1.41  0.16433
## Year2010          0.299      0.167      1.78  0.08029 .
## Year2011          0.705      0.249      2.83  0.00666 **
## Year2012          0.914      0.336      2.72  0.00900 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.516
## Multiple R-squared:  0.352, Adjusted R-squared:  0.136
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 60 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.543  0.911  0.965  0.928  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      1.45e-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: 69"
## [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
##   12   12   7   10   11   7   4   12   13   8   9   16   27   27   24
## 2011 2012
##   19   14
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    9   10    5    6    3    5    3   12   12    8    5   10   17   23   16
## 2011 2012

```

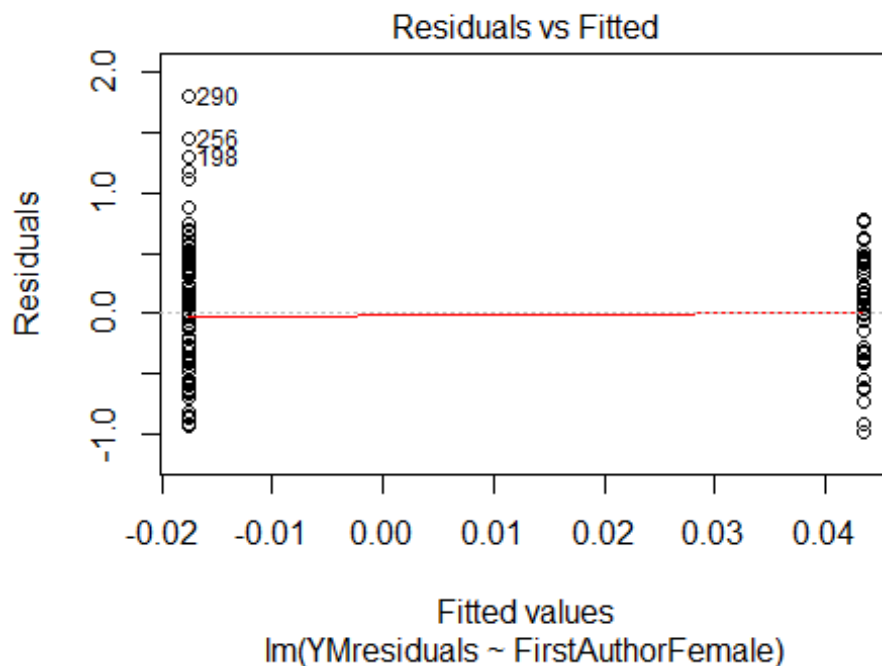
```
## 17 12
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 9 9 5 5 3 4 2 11 12 7 4 10 16 23 15
## 2011 2012
## 17 10
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 16, df = 16, p-value = 0.5
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.6, df = 1, p-value = 0.2
## [1] "Female first author team size 2018 geometric mean: 1.18920711500272"
## [1] "Male first author team size 2018 geometric mean: 2.31838774012131"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

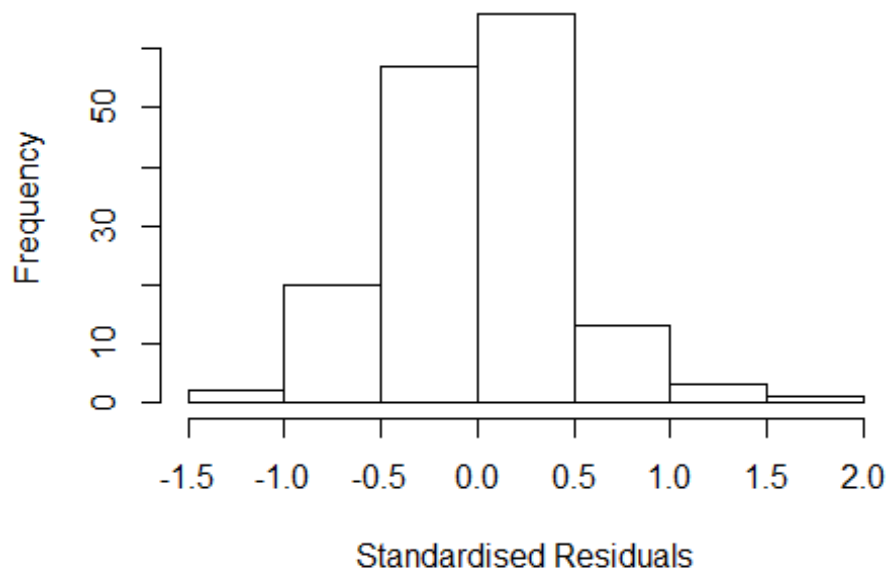
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4, p-value = 0.06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.43096908110526"
## [1] "Male last author team size 2018 geometric mean: 2.22090615485233"

## Warning in wilcox.test.default(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] "Regression 1: First author gender, last author gender, team size,
## Year as factors"
##
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.103  1      1.450
## LastAuthorFemale  1.647  1      1.283
## UniqueAuthors    5.027  4      1.224
## Year              9.928 16      1.074
```

Residuals from first and last author and team size



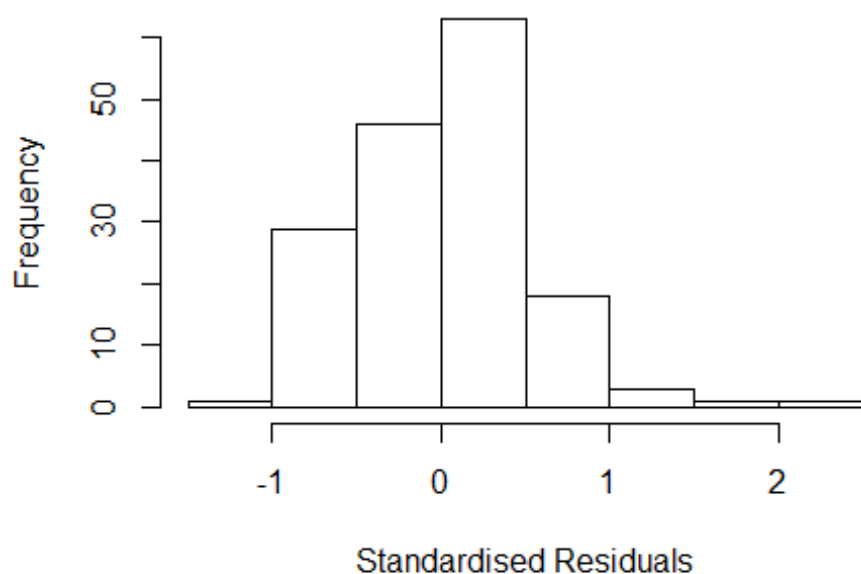
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.29269 -0.35353  0.00909  0.31089  1.87486
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.7983    0.1699   4.70 6.2e-06 ***
## FirstAuthorFemale1 0.1043    0.0932   1.12 0.26534
## LastAuthorFemale1 0.0375    0.0931   0.40 0.68790
## UniqueAuthors2    0.3862    0.1061   3.64 0.00038 ***
## UniqueAuthors3    0.3948    0.1595   2.48 0.01452 *
## UniqueAuthors4    0.1479    0.2202   0.67 0.50287
## UniqueAuthors5    0.1762    0.1748   1.01 0.31519
## Year1997          0.1584    0.2289   0.69 0.49011
## Year1998          0.1995    0.2238   0.89 0.37431
## Year1999         -0.0216    0.2040  -0.11 0.91601
```

```

## Year2000          0.1259      0.4575      0.28  0.78360
## Year2001          0.1709      0.2075      0.82  0.41164
## Year2002         -0.8900      0.1864     -4.78  4.5e-06 ***
## Year2003         -0.0945      0.2354     -0.40  0.68880
## Year2004         -0.1207      0.2127     -0.57  0.57141
## Year2005         -0.1712      0.1939     -0.88  0.37882
## Year2006         -0.0731      0.2201     -0.33  0.74032
## Year2007         -0.5822      0.1954     -2.98  0.00341 **
## Year2008         -0.0336      0.2271     -0.15  0.88259
## Year2009         -0.0740      0.2002     -0.37  0.71223
## Year2010         -0.0674      0.2244     -0.30  0.76447
## Year2011         -0.3806      0.2157     -1.76  0.07981 .
## Year2012         -0.6834      0.2749     -2.49  0.01412 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.476
## Multiple R-squared:  0.258, Adjusted R-squared:  0.141
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 149 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0854 0.9000 0.9510 0.9070 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          6.17e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.989 1          1.41
## LastAuthorFemale 1.903 1          1.38
## Year          2.554 16          1.03

```

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
## -1.0013 -0.3635 0.0319 0.3636 2.0406
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.02722 0.17655 5.82 3.7e-08 ***
## FirstAuthorFemale1 0.16776 0.11016 1.52 0.1300
## LastAuthorFemale1 -0.07330 0.11234 -0.65 0.5151
## Year1997 0.09462 0.23524 0.40 0.6881
## Year1998 -0.01940 0.23020 -0.08 0.9330
## Year1999 0.01979 0.24280 0.08 0.9351
## Year2000 -0.09445 0.46380 -0.20 0.8389
## Year2001 -0.00347 0.20067 -0.02 0.9862
## Year2002 -0.95760 0.18785 -5.10 1.1e-06 ***
## Year2003 -0.17168 0.24456 -0.70 0.4838
## Year2004 -0.27076 0.22567 -1.20 0.2322
## Year2005 -0.22430 0.23578 -0.95 0.3431
```

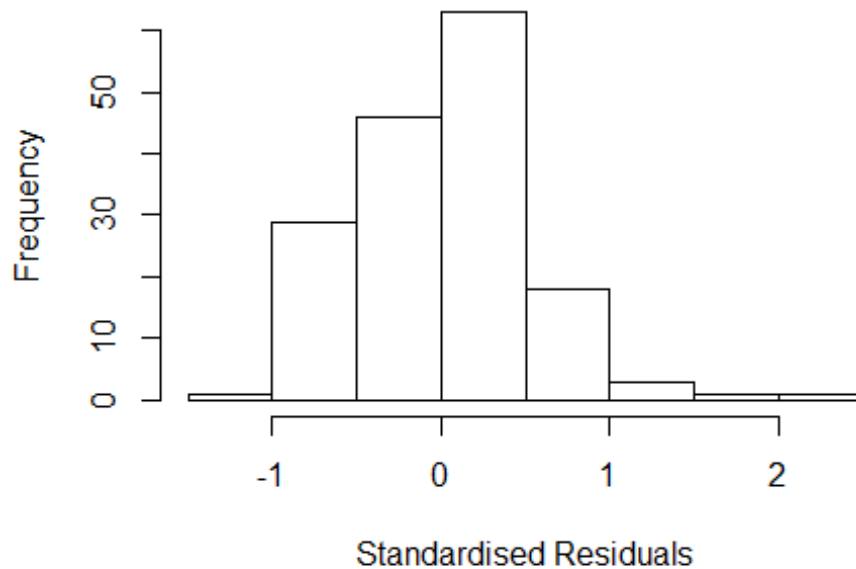


```

## Year2006          -0.27899    0.22890   -1.22    0.2249
## Year2007          -0.63044    0.20485   -3.08    0.0025 **
## Year2008          -0.12039    0.24018   -0.50    0.6170
## Year2009          -0.10783    0.20283   -0.53    0.5958
## Year2010          -0.23891    0.24922   -0.96    0.3394
## Year2011          -0.40923    0.22104   -1.85    0.0662 .
## Year2012          -0.69187    0.27264   -2.54    0.0122 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.503
## Multiple R-squared:  0.162, Adjusted R-squared:  0.0562
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 148 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0625 0.8820 0.9480 0.9060 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      6.17e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## 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
## -1.0316 -0.3593 0.0423 0.3637 2.0380
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.01446 0.16948 5.99 1.6e-08 ***
## FirstAuthorFemale1 0.12412 0.09900 1.25 0.2119
## Year1997 0.10166 0.23131 0.44 0.6609
## Year1998 -0.01311 0.22590 -0.06 0.9538
## Year1999 0.01629 0.23582 0.07 0.9450
## Year2000 -0.09364 0.45516 -0.21 0.8373
## Year2001 0.00161 0.19606 0.01 0.9935
## Year2002 -0.92302 0.18146 -5.09 1.1e-06 ***
## Year2003 -0.16048 0.24273 -0.66 0.5096
## Year2004 -0.26818 0.22487 -1.19 0.2350
## Year2005 -0.20746 0.23506 -0.88 0.3789
## Year2006 -0.28008 0.22449 -1.25 0.2142
```

```

## Year2007          -0.60632    0.19456   -3.12    0.0022 **
## Year2008          -0.10694    0.23504   -0.45    0.6498
## Year2009          -0.09617    0.19829   -0.49    0.6284
## Year2010          -0.24017    0.24800   -0.97    0.3344
## Year2011          -0.40585    0.21951   -1.85    0.0665 .
## Year2012          -0.67642    0.27792   -2.43    0.0162 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.5
## Multiple R-squared:  0.161, Adjusted R-squared:  0.062
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 149 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0586 0.8820 0.9460 0.9060 0.9840 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      6.17e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.356 1          1.165
## Year          1.356 16          1.010
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -0.9626 -0.3869  0.0566  0.3490  1.9787

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.020500   0.172996   5.90 2.5e-08 ***
## LastAuthorFemale1 0.041655   0.099078   0.42 0.6748
## Year1997        0.111398   0.233911   0.48 0.6346
## Year1998       -0.001230   0.229333  -0.01 0.9957
## Year1999        0.000384   0.233518   0.00 0.9987
## Year2000       -0.065672   0.468972  -0.14 0.8888
## Year2001        0.016864   0.199743   0.08 0.9328
## Year2002       -0.867000   0.205220  -4.22 4.2e-05 ***
## Year2003       -0.140141   0.250093  -0.56 0.5761
## Year2004       -0.222127   0.218533  -1.02 0.3111
## Year2005       -0.150783   0.241293  -0.62 0.5330
## Year2006       -0.247695   0.231294  -1.07 0.2860
## Year2007       -0.564823   0.198581  -2.84 0.0051 **
## Year2008       -0.099564   0.236835  -0.42 0.6748
## Year2009       -0.089616   0.200586  -0.45 0.6557
## Year2010       -0.207438   0.245663  -0.84 0.3998
## Year2011       -0.379849   0.219980  -1.73 0.0864 .
## Year2012       -0.623198   0.284028  -2.19 0.0298 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.506
## Multiple R-squared:  0.148, Adjusted R-squared:  0.0475
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 152 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0927 0.8850 0.9470 0.9090 0.9850 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      6.17e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 162"
## [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]"
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    5    4    4    3    3    6    8    6    6    6   14   12    5    8    6
## 2012
##    3
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    2    1    0    2    2    3    2    4    4    9    5    2    4    3
## 2012
##    1
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    1    1    0    2    2    3    2    3    4    9    5    2    4    2
## 2012
##    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: 2"

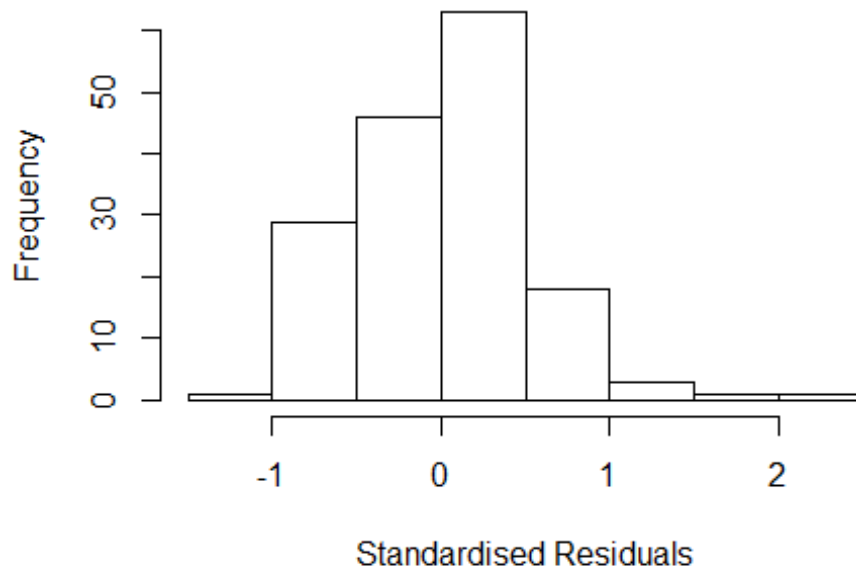
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0.5, p-value = NA
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2"
## [1] "Male last author team size 2018 geometric mean: 2"

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

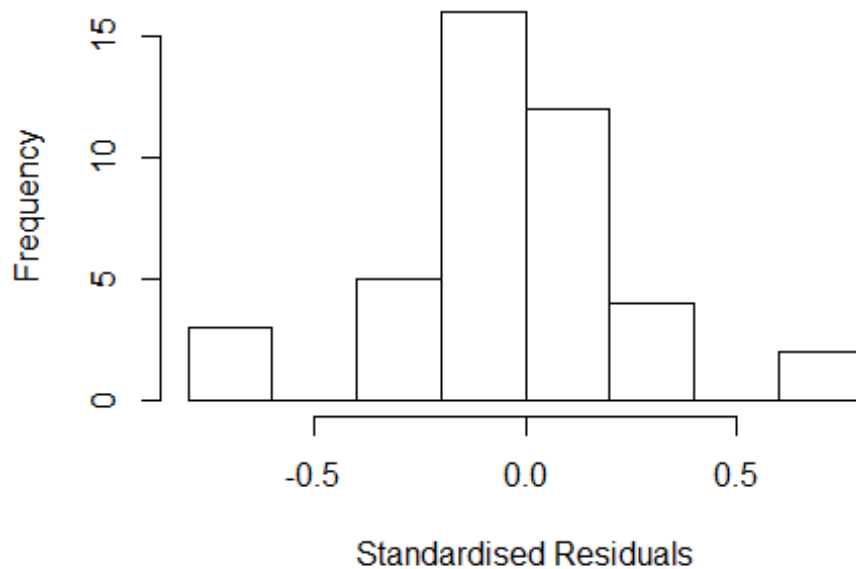
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0.5, p-value = NA
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##               GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.400e+15 1      3.741e+07
## LastAuthorFemale -2.019e+15 1           NaN
## UniqueAuthors    1.676e+60 4      3.373e+07
## Year              7.452e+59 14     1.375e+02
```

Residuals from first and last author and team size



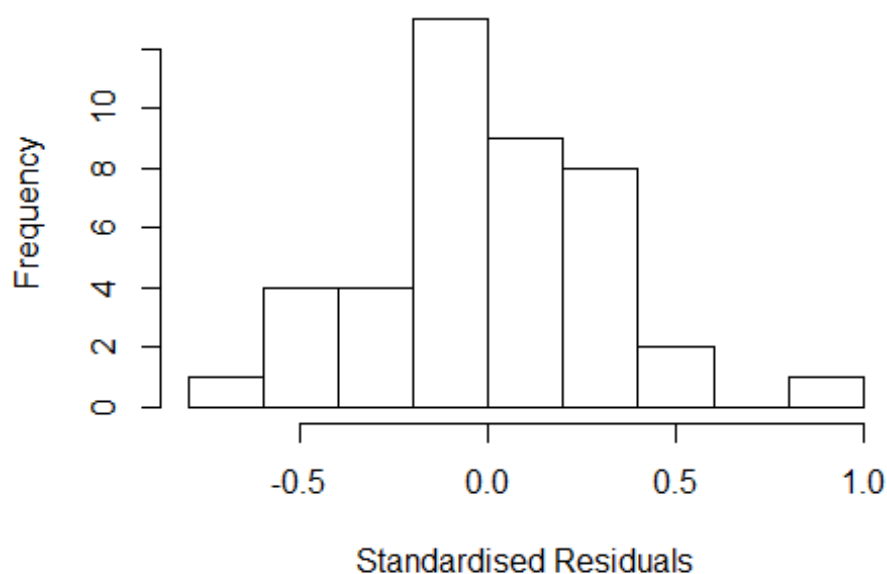
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -7.09e-01 -1.03e-01 -5.27e-16 1.37e-01 7.18e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.01e+00 1.96e-08 5.13e+07 < 2e-16 ***
## FirstAuthorFemale1 3.15e-01 1.66e-01 1.90e+00 0.0717 .
## LastAuthorFemale1 -9.56e-02 1.84e-01 -5.20e-01 0.6082
## UniqueAuthors2 1.69e-01 1.35e-01 1.25e+00 0.2251
## UniqueAuthors3 2.38e-01 2.33e-01 1.02e+00 0.3191
## UniqueAuthors4 -2.34e-01 1.46e-01 -1.60e+00 0.1242
## UniqueAuthors5 1.08e+00 1.66e-01 6.51e+00 1.9e-06 ***
## Year1998 -2.91e-01 1.52e-01 -1.92e+00 0.0691 .
## Year1999 -8.77e-01 2.36e-01 -3.71e+00 0.0013 **
## Year2001 -1.49e-01 1.76e-01 -8.50e-01 0.4044
```

```

## Year2002      -3.48e-01    2.14e-08 -1.63e+07 < 2e-16 ***
## Year2003      -2.84e-01    3.98e-01 -7.10e-01    0.4834
## Year2004       1.37e-01    1.52e-01  9.00e-01    0.3800
## Year2005      -5.89e-02    1.72e-01 -3.40e-01    0.7352
## Year2006      -5.19e-01    1.78e-01 -2.91e+00    0.0083 **
## Year2007      -5.90e-02    1.72e-01 -3.40e-01    0.7348
## Year2008      -5.05e-01    2.93e-01 -1.72e+00    0.0993 .
## Year2009      -4.63e-02    1.29e-01 -3.60e-01    0.7227
## Year2010       4.26e-04    1.13e-01  0.00e+00    0.9970
## Year2011       3.96e-01    1.57e-01  2.52e+00    0.0199 *
## Year2012      -1.65e-01    3.65e-01 -4.50e-01    0.6558
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.348
## Multiple R-squared:  0.552, Adjusted R-squared:  0.126
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 34 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.649  0.942  0.982  0.933  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           2.38e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01           5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500           50           2           1           1000           200
## trace.lev      mts      compute.rd
##           0           1000           0
##           psi           subsampling           cov
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.096e+15 1 3.310e+07
## LastAuthorFemale 1.666e+15 1 4.082e+07
## Year 2.631e+29 14 1.124e+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
## -7.06e-01 -1.72e-01 9.71e-17 2.09e-01 8.80e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.01e+00 2.66e-08 3.78e+07 < 2e-16 ***
## FirstAuthorFemale1 4.50e-01 2.21e-01 2.04e+00 0.05199 .
## LastAuthorFemale1 -2.78e-01 2.02e-01 -1.37e+00 0.18218
## Year1998 -2.44e-01 1.54e-01 -1.58e+00 0.12606
## Year1999 -8.43e-01 2.21e-01 -3.82e+00 0.00078 ***
## Year2001 -1.49e-01 1.74e-01 -8.60e-01 0.39850
## Year2002 1.25e-01 5.10e-01 2.50e-01 0.80802
## Year2003 -1.84e-01 2.64e-01 -7.00e-01 0.49224
## Year2004 3.05e-01 7.06e-02 4.33e+00 0.00021 ***
## Year2005 7.37e-02 2.42e-01 3.00e-01 0.76341
## Year2006 -4.56e-01 1.65e-01 -2.76e+00 0.01060 *
## Year2007 1.20e-02 1.54e-01 8.00e-02 0.93834
```

```

## Year2008          -4.29e-01    2.53e-01 -1.70e+00  0.10188
## Year2009          7.25e-02    1.46e-01  5.00e-01  0.62347
## Year2010          1.71e-01    7.50e-02  2.28e+00  0.03139 *
## Year2011          5.99e-01    8.65e-02  6.92e+00   3e-07 ***
## Year2012          2.55e-01    2.02e-01  1.26e+00  0.21979
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.368
## Multiple R-squared:  0.424, Adjusted R-squared:  0.0553
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 36 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.546  0.899  0.969  0.928  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      2.38e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##  nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##  trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -2.745e+15  1      NaN
## Year              -2.745e+15 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
## -0.687891 -0.220250  0.000235  0.218321  0.820661
##

```

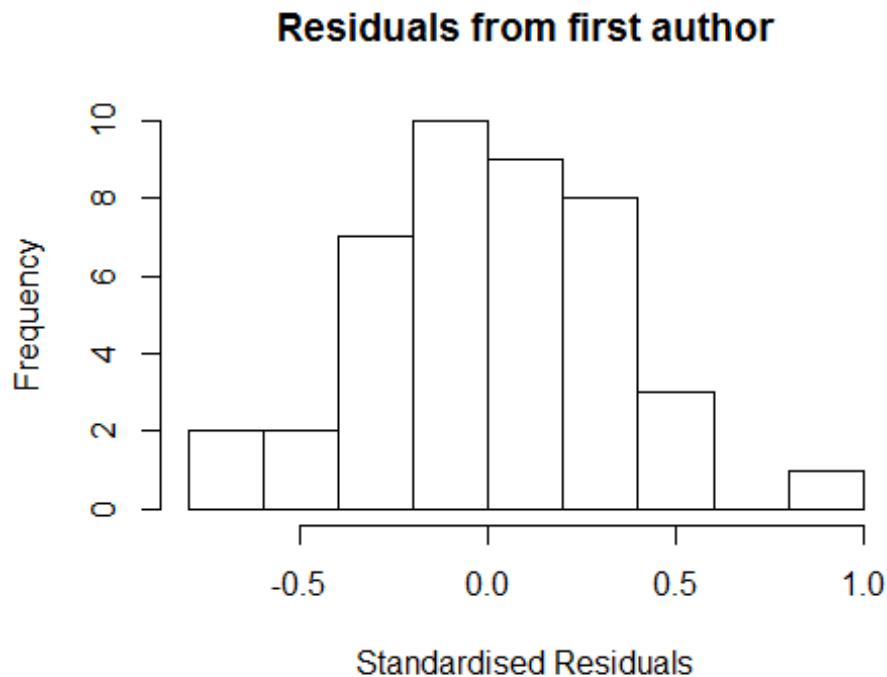
```

## Coefficients:
##               Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.01e+00   1.05e-08  9.61e+07 < 2e-16 ***
## FirstAuthorFemale1 2.81e-01   1.57e-01  1.79e+00  0.08517 .
## Year1998        -3.53e-01   1.57e-01 -2.25e+00  0.03328 *
## Year1999        -6.74e-01   1.57e-01 -4.29e+00  0.00022 ***
## Year2001        -1.49e-01   1.72e-01 -8.70e-01  0.39312
## Year2002         2.10e-01   6.59e-01  3.20e-01  0.75273
## Year2003        -1.90e-01   2.56e-01 -7.40e-01  0.46420
## Year2004         3.05e-01   7.04e-02  4.34e+00  0.00019 ***
## Year2005        -5.67e-02   1.40e-01 -4.10e-01  0.68764
## Year2006        -5.11e-01   1.90e-01 -2.69e+00  0.01235 *
## Year2007        -6.11e-03   1.46e-01 -4.00e-02  0.96694
## Year2008        -3.70e-01   2.38e-01 -1.55e+00  0.13255
## Year2009         7.25e-02   1.45e-01  5.00e-01  0.62062
## Year2010         1.03e-01   8.75e-02  1.17e+00  0.25182
## Year2011         5.99e-01   8.63e-02  6.94e+00  2.3e-07 ***
## Year2012        -2.30e-02   1.83e-08 -1.26e+06 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.396
## Multiple R-squared:  0.388, Adjusted R-squared:  0.035
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 32 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.648  0.904  0.967  0.927  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           2.38e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01           5.00e-01
##   nResample    max.it    best.r.s    k.fast.s    k.max maxit.scale
##           500         50         2         1         1000         200
##   trace.lev    mts    compute.rd
##           0         1000         0
##           psi           subsampling           cov
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"

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

```

```
## 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 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
## -7.26e-01 -1.66e-01  8.71e-16  2.12e-01  7.39e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.0080    0.0000    Inf < 2e-16 ***
## LastAuthorFemale1  0.0249    0.1499    0.17  0.8693
## Year1998        -0.0969    0.1499   -0.65  0.5236
## Year1999        -0.3930    0.0000   -Inf < 2e-16 ***
## Year2001        -0.1495    0.1737   -0.86  0.3972
```

```

## Year2002          0.3505      2.0755      0.17      0.8672
## Year2003         -0.1852      0.2617     -0.71      0.4854
## Year2004          0.3055      0.0706      4.33      0.0002 ***
## Year2005          0.0247      0.2103      0.12      0.9076
## Year2006         -0.3798      0.1335     -2.85      0.0085 **
## Year2007          0.0323      0.1553      0.21      0.8369
## Year2008         -0.2881      0.2383     -1.21      0.2375
## Year2009          0.0725      0.1456      0.50      0.6228
## Year2010          0.0968      0.0903      1.07      0.2935
## Year2011          0.5990      0.0865      6.93      2.4e-07 ***
## Year2012         -0.0479      0.1499     -0.32      0.7518
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.373
## Multiple R-squared:  0.348, Adjusted R-squared:  -0.0276
## Convergence in 49 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 35 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.674  0.927  0.966   0.920  0.984   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      2.38e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 42"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 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
##    1    6    1    2    6    4    6    8    4    6    9    8   11   11   10
## 2011 2012

```

```

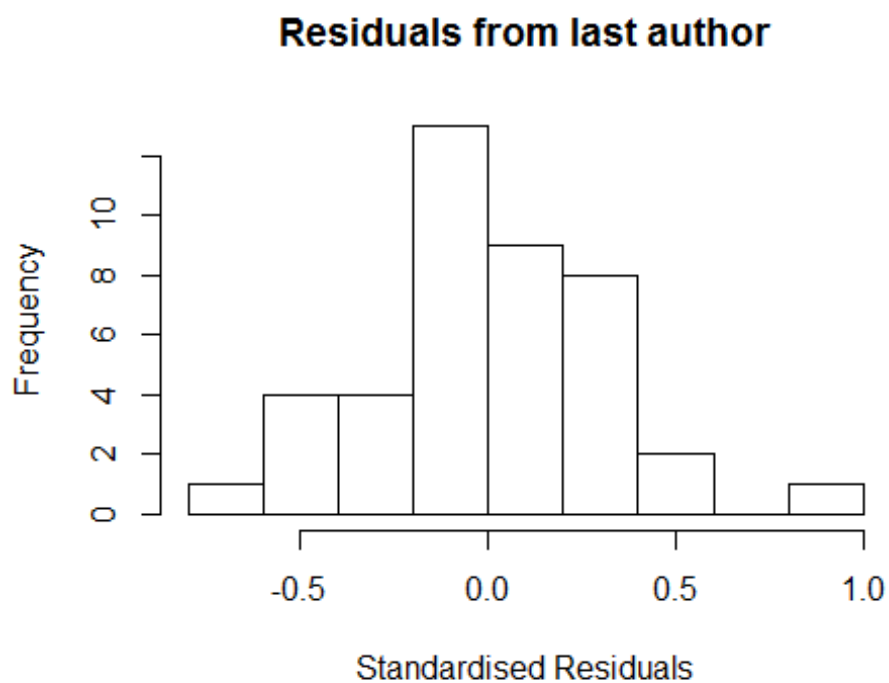
## 16 9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1 3 0 1 2 3 2 5 2 4 7 4 7 8 7
## 2011 2012
## 11 6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1 3 0 1 1 3 2 5 2 4 6 4 6 8 7
## 2011 2012
## 11 6
## [1] "Heteroscedasticity 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: 3.56762134500816"

## Warning in wilcox.test.default(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] "Female last author team size 2018 geometric mean: 2.6207413942089"
## [1] "Male last author team size 2018 geometric mean: 4.32674871092223"

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

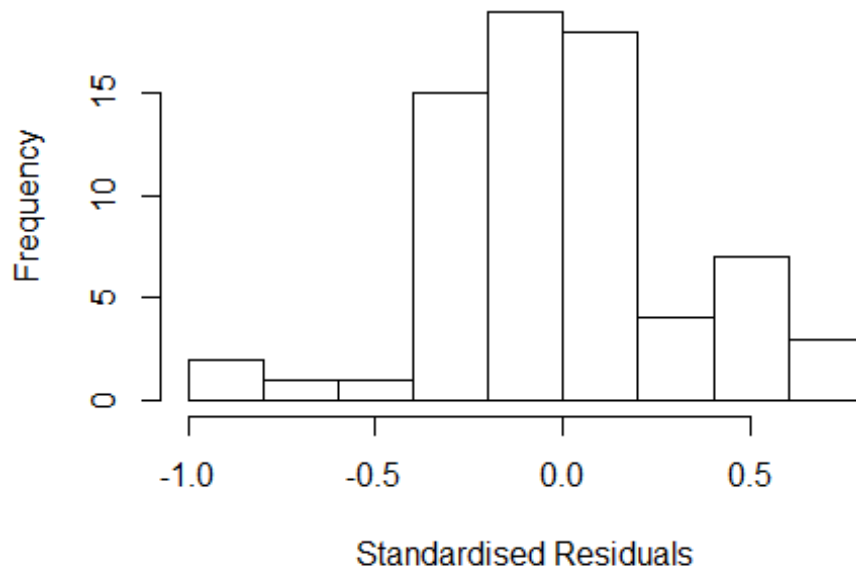
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3, p-value = 0.6
## 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 | -1.511e+02 | 1 | NaN |
| ## LastAuthorFemale | -1.323e+15 | 1 | NaN |
| ## UniqueAuthors | -2.274e+16 | 4 | NaN |
| ## Year | 2.550e+17 | 15 | 3.804 |

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
## -8.22e-01 -2.30e-01 2.22e-16 1.14e-01 7.80e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.427 0.357 -1.19 0.23807
## FirstAuthorFemale1 -0.142 0.209 -0.68 0.50015
## LastAuthorFemale1 0.979 0.258 3.80 0.00041 ***
## UniqueAuthors2 0.629 0.209 3.01 0.00420 **
## UniqueAuthors3 0.660 0.181 3.65 0.00064 ***
## UniqueAuthors4 0.864 0.185 4.67 2.4e-05 ***
## UniqueAuthors5 0.764 0.171 4.46 4.9e-05 ***
## Year1997 1.435 0.477 3.01 0.00415 **
## Year1999 0.665 0.357 1.86 0.06891 .
## Year2000 1.232 0.357 3.45 0.00119 **
```

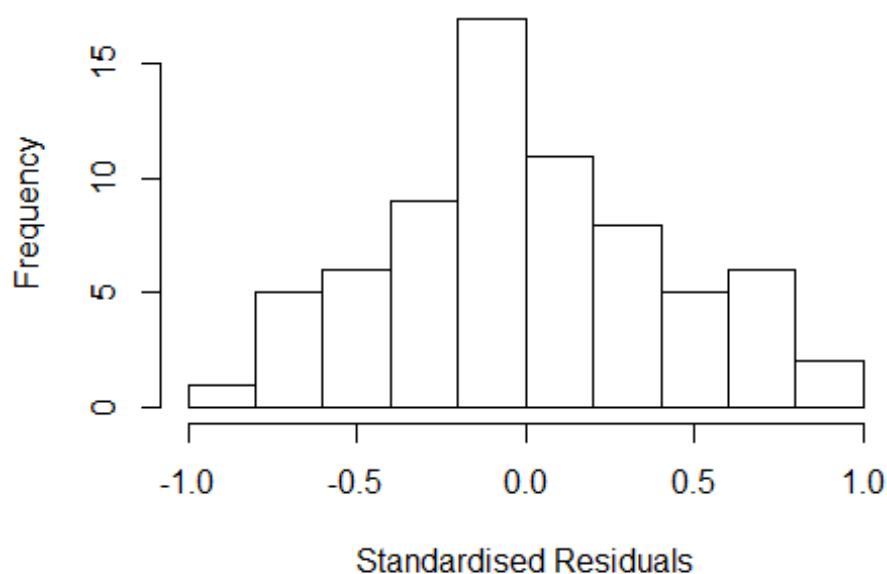


```

## Year2001          1.004      0.467      2.15  0.03680 *
## Year2002          0.742      0.361      2.06  0.04527 *
## Year2003          0.911      0.392      2.32  0.02449 *
## Year2004          0.803      0.357      2.25  0.02928 *
## Year2005          1.193      0.328      3.64  0.00067 ***
## Year2006          1.046      0.416      2.52  0.01522 *
## Year2007          0.991      0.300      3.30  0.00181 **
## Year2008          1.078      0.368      2.93  0.00517 **
## Year2009          1.063      0.388      2.74  0.00851 **
## Year2010          1.160      0.527      2.20  0.03268 *
## Year2011          0.722      0.410      1.76  0.08481 .
## Year2012          0.813      0.389      2.09  0.04208 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.356
## Multiple R-squared:  0.598, Adjusted R-squared:  0.423
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 60 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.574  0.889  0.958  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      1.43e-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 4.171e+00 1      2.042
## LastAuthorFemale -1.436e+15 1      NaN
## Year              -2.949e+15 15      NaN

```

Residuals from first and last author



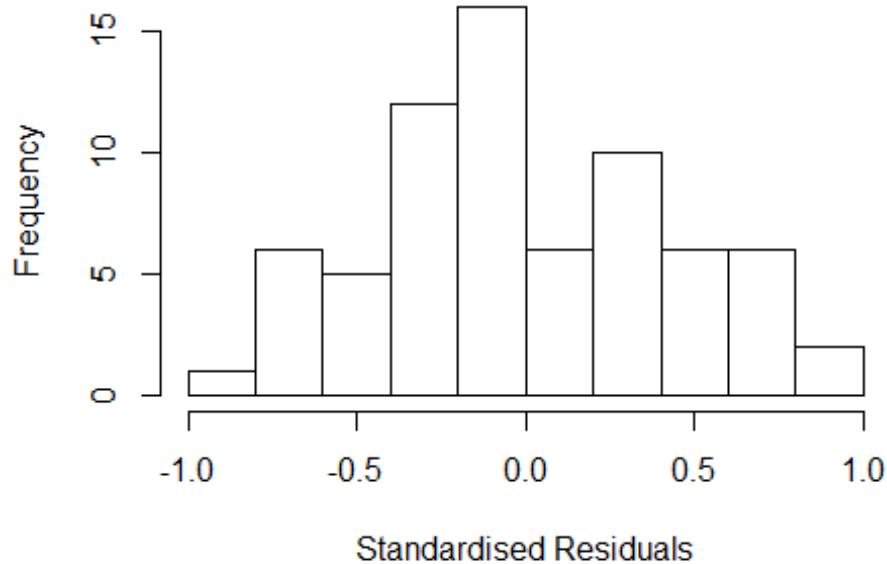
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.66e-01 -2.67e-01 -2.08e-16  2.50e-01  8.57e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.50462    0.25053     2.01  0.0492 *
## FirstAuthorFemale1 0.24138    0.17238     1.40  0.1674
## LastAuthorFemale1 0.70738    0.25053     2.82  0.0067 **
## Year1997         0.50308    0.36735     1.37  0.1767
## Year1999        -0.26662    0.25053    -1.06  0.2921
## Year2000         0.30038    0.25053     1.20  0.2360
## Year2001         0.07359    0.36743     0.20  0.8420
## Year2002        -0.18962    0.25554    -0.74  0.4614
## Year2003        -0.00225    0.29169    -0.01  0.9939
## Year2004        -0.12862    0.25053    -0.51  0.6098
## Year2005         0.42189    0.28083     1.50  0.1391
## Year2006         0.34899    0.32128     1.09  0.2824
```

```

## Year2007          0.38956    0.32699    1.19    0.2389
## Year2008          0.14962    0.21094    0.71    0.4813
## Year2009          0.55385    0.36583    1.51    0.1361
## Year2010          0.61732    0.35863    1.72    0.0911 .
## Year2011          0.19262    0.26896    0.72    0.4771
## Year2012          0.46141    0.37145    1.24    0.2197
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.561
## Multiple R-squared:  0.29,   Adjusted R-squared:  0.0584
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 59 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.748  0.911  0.971   0.943  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      1.43e-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 1.553 1      1.246
## Year              1.553 15      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
## -0.9650 -0.2724 -0.0219 0.2849 0.8494
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.21e+00 2.47e-08 4.91e+07 < 2e-16 ***
## FirstAuthorFemale1 4.11e-01 1.75e-01 2.35e+00 0.0226 *
## Year1997 -2.04e-01 2.67e-01 -7.70e-01 0.4476
## Year1999 -9.74e-01 4.27e-08 -2.28e+07 < 2e-16 ***
## Year2000 -4.07e-01 3.44e-08 -1.18e+07 < 2e-16 ***
## Year2001 -6.81e-01 3.30e-01 -2.06e+00 0.0441 *
## Year2002 -8.97e-01 5.03e-02 -1.78e+01 < 2e-16 ***
## Year2003 -7.09e-01 1.51e-01 -4.69e+00 1.9e-05 ***
## Year2004 -8.36e-01 2.41e-08 -3.47e+07 < 2e-16 ***
## Year2005 -2.85e-01 1.28e-01 -2.23e+00 0.0300 *
## Year2006 -3.89e-01 1.48e-01 -2.64e+00 0.0110 *
## Year2007 -3.18e-01 2.09e-01 -1.52e+00 0.1347
```

```

## Year2008          -4.79e-01   1.92e-01 -2.50e+00   0.0155 *
## Year2009          -1.80e-01   2.14e-01 -8.40e-01   0.4036
## Year2010          -1.19e-01   2.00e-01 -6.00e-01   0.5533
## Year2011          -4.84e-01   1.44e-01 -3.37e+00   0.0014 **
## Year2012          -2.47e-01   2.72e-01 -9.10e-01   0.3673
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.575
## Multiple R-squared:  0.248, Adjusted R-squared:  0.0213
## Convergence in 9 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 61 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.760  0.911  0.971  0.945  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      1.43e-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 3.45e+13  1      5.874e+06
## Year              3.45e+13 15      2.827e+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
## -9.67e-01 -2.74e-01  7.91e-16  2.36e-01  9.16e-01
##

```

```

## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.2912    0.1927    1.51  0.1367
## LastAuthorFemale1 0.9208    0.1927    4.78 1.4e-05 ***
## Year1997        0.7165    0.3316    2.16  0.0352 *
## Year1999       -0.0532    0.1927   -0.28  0.7835
## Year2000        0.5138    0.1927    2.67  0.0102 *
## Year2001        0.3604    0.2856    1.26  0.2124
## Year2002        0.0238    0.1992    0.12  0.9055
## Year2003        0.2107    0.2441    0.86  0.3919
## Year2004        0.0848    0.1927    0.44  0.6618
## Year2005        0.6351    0.2309    2.75  0.0081 **
## Year2006        0.6058    0.2444    2.48  0.0164 *
## Year2007        0.6029    0.2850    2.12  0.0391 *
## Year2008        0.3911    0.1801    2.17  0.0344 *
## Year2009        0.8064    0.2822    2.86  0.0061 **
## Year2010        0.8727    0.2781    3.14  0.0028 **
## Year2011        0.4339    0.1801    2.41  0.0195 *
## Year2012        0.6755    0.3384    2.00  0.0510 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.551
## Multiple R-squared:  0.279, Adjusted R-squared:  0.0613
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~ = 1. The remaining 61 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.739  0.920  0.969  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      1.43e-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 1903"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1997 1998 2001 2003 2005 2008 2009 2010 2011 2012
##    1    1    1    2    3    1    2    3    2    6
##
## 1997 1998 2001 2003 2005 2008 2009 2010 2011 2012
##    0    0    1    2    3    1    2    2    1    4
##
## 1997 1998 2001 2003 2005 2008 2009 2010 2011 2012
##    0    0    1    2    3    1    1    2    0    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: 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"

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

## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 14"
## [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 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    2    4    4    4    4    3    6    5    8    9    7    7    7    8    7
## 2012
##    8
##
## 1996 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    4    3    1    4    2    5    1    6    6    3    5    4    4    7
## 2012
##    6
##
## 1996 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    0    4    3    1    4    1    4    1    5    6    3    5    3    4    7
## 2012
##    6

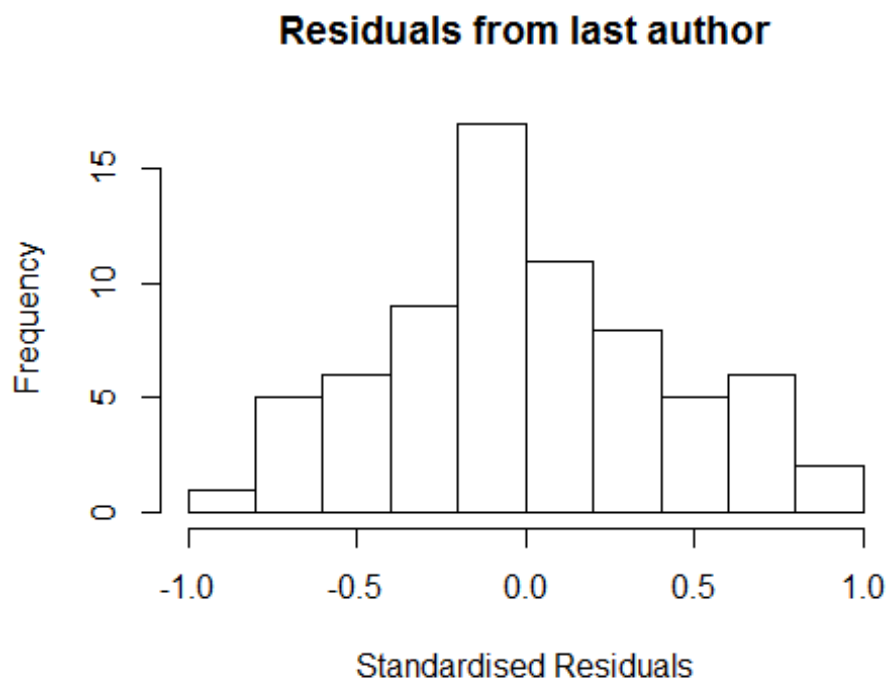
```

```
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.90636858599387"
## [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 = 15, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.56508458007329"
## [1] "Male last author team size 2018 geometric mean: 1.5157165665104"

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

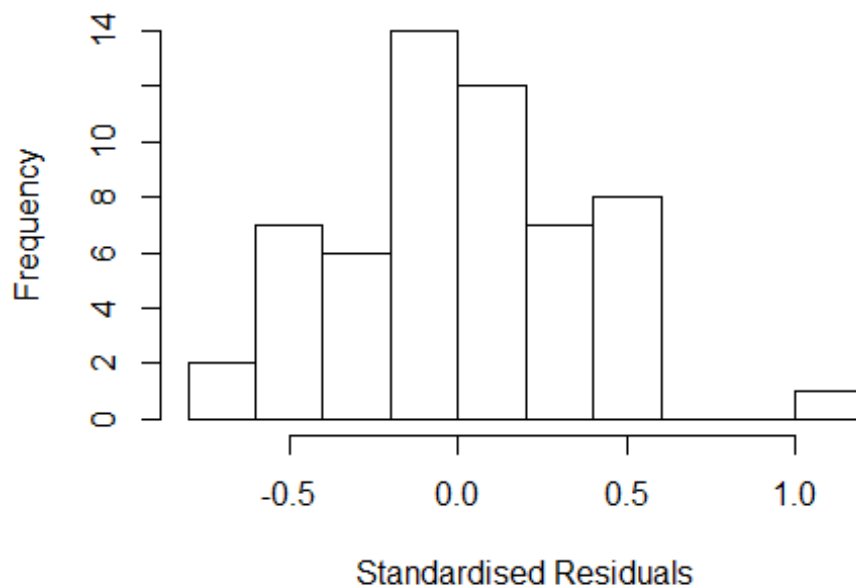


```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 10, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
```



```
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -1.748e+15  1      NaN
## LastAuthorFemale  -2.128e+01  1      NaN
## UniqueAuthors     -2.467e+31  4      NaN
## Year              -3.313e+45 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
## -7.73e-01 -2.79e-01  6.66e-16  2.68e-01  1.02e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.9565      0.4845   1.97   0.056 .
```

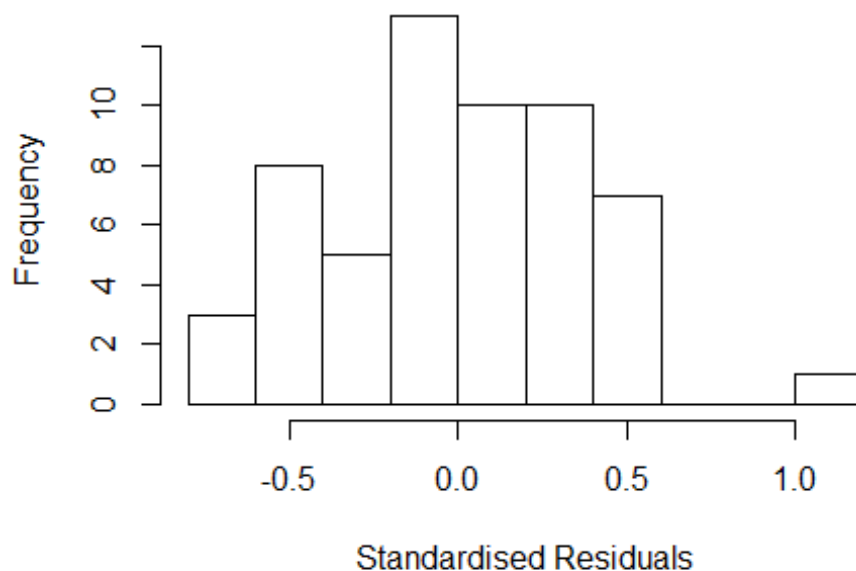
```

## FirstAuthorFemale1    0.2144    0.1476    1.45    0.155
## LastAuthorFemale1    -0.4615    0.1892    -2.44    0.020 *
## UniqueAuthors2       -0.1255    0.1542    -0.81    0.421
## UniqueAuthors3        0.0337    0.1888    0.18    0.860
## UniqueAuthors4        0.0987    0.2573    0.38    0.704
## UniqueAuthors5        0.6449    0.3081    2.09    0.043 *
## Year1999              -0.0352    0.5413    -0.07    0.949
## Year2000              -0.2795    0.4845    -0.58    0.568
## Year2001              -0.1666    0.5029    -0.33    0.742
## Year2002               0.5050    0.4197    1.20    0.237
## Year2003               0.1034    0.4463    0.23    0.818
## Year2004               0.2376    0.3568    0.67    0.510
## Year2005              -0.1063    0.4737    -0.22    0.824
## Year2006               0.4887    0.4295    1.14    0.263
## Year2007              -0.1238    0.3924    -0.32    0.754
## Year2008               0.2503    0.4586    0.55    0.589
## Year2009               0.4889    0.5002    0.98    0.335
## Year2010               0.0623    0.5060    0.12    0.903
## Year2011              -0.0292    0.4885    -0.06    0.953
## Year2012               0.0639    0.5168    0.12    0.902
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.45
## Multiple R-squared:  0.331, Adjusted R-squared:  -0.0411
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 47 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.586  0.893  0.957  0.933  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 2.263e+14 1 1.504e+07
## LastAuthorFemale 4.621e+00 1 2.150e+00
## Year 3.924e+14 14 3.321e+00
```

Residuals from first and last author



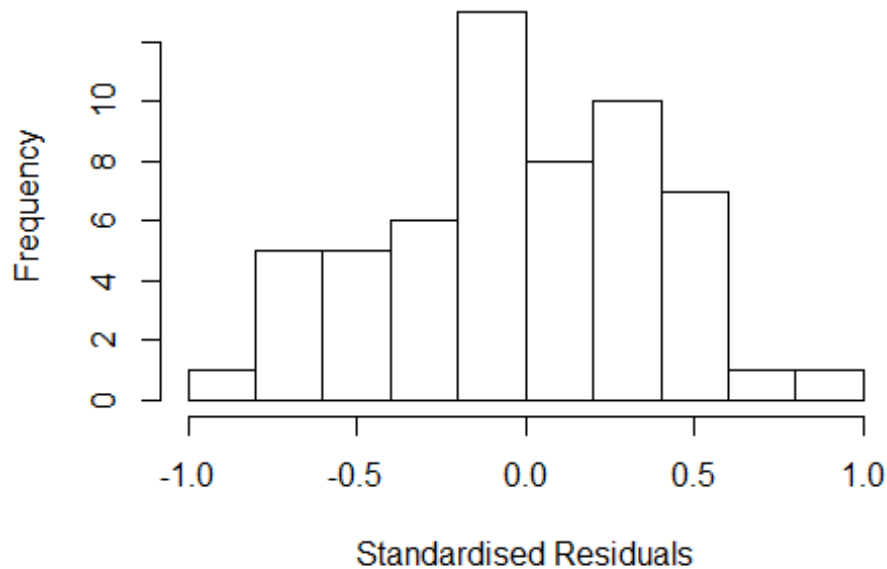
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -7.70e-01 -2.15e-01 -1.94e-16 2.89e-01 1.05e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9222 0.3840 2.40 0.0211 *
## FirstAuthorFemale1 0.2090 0.1334 1.57 0.1251
## LastAuthorFemale1 -0.4838 0.1672 -2.89 0.0062 **
## Year1999 0.0153 0.4850 0.03 0.9750
## Year2000 -0.2452 0.3840 -0.64 0.5268
## Year2001 -0.1243 0.4112 -0.30 0.7640
## Year2002 0.4138 0.3840 1.08 0.2877
## Year2003 0.0764 0.3865 0.20 0.8442
```

```

## Year2004          0.1518      0.3243      0.47      0.6423
## Year2005          0.0805      0.4517      0.18      0.8595
## Year2006          0.5444      0.3969      1.37      0.1778
## Year2007         -0.1178      0.3518     -0.33      0.7395
## Year2008          0.2342      0.3934      0.60      0.5549
## Year2009          0.4880      0.4204      1.16      0.2526
## Year2010          0.1107      0.4392      0.25      0.8024
## Year2011         -0.0487      0.4311     -0.11      0.9106
## Year2012          0.1225      0.4713      0.26      0.7962
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.455
## Multiple R-squared:  0.289, Adjusted R-squared:  0.00448
## 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.
##  0.570  0.899  0.957   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      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 -8.591e+14 1      NaN
## Year              -8.591e+14 14      NaN

```

Residuals from first author

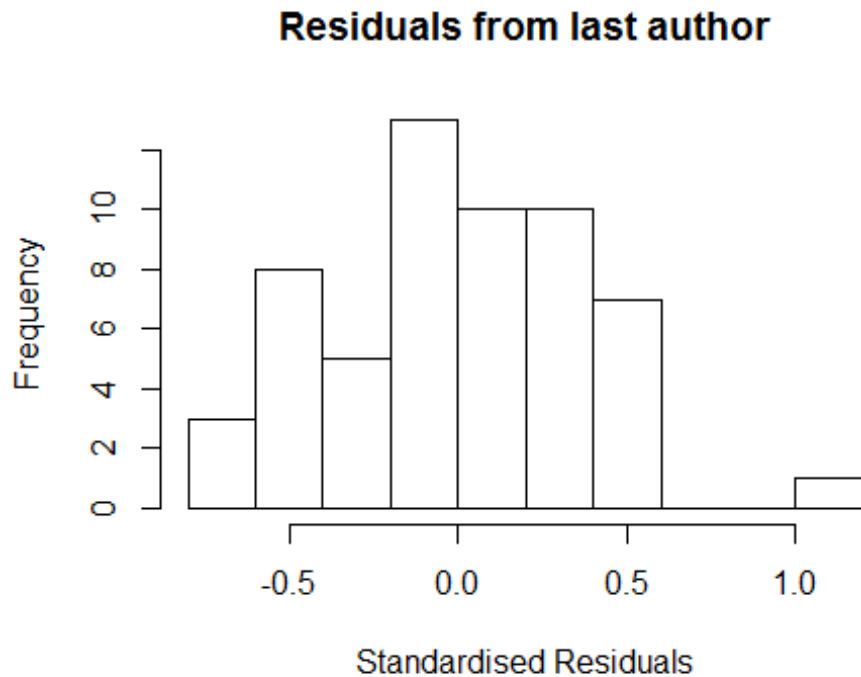


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -8.53e-01 -2.91e-01 5.55e-17 2.75e-01 9.47e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.02964 0.43757 2.35 0.023 *
## FirstAuthorFemale1 -0.06762 0.13364 -0.51 0.616
## Year1999 -0.21831 0.56215 -0.39 0.700
## Year2000 -0.35264 0.43757 -0.81 0.425
## Year2001 -0.16726 0.44355 -0.38 0.708
## Year2002 0.30636 0.43757 0.70 0.488
## Year2003 0.00604 0.46416 0.01 0.990
## Year2004 0.32098 0.40280 0.80 0.430
## Year2005 0.00758 0.55728 0.01 0.989
## Year2006 0.40122 0.43844 0.92 0.365
## Year2007 -0.04193 0.43855 -0.10 0.924
## Year2008 0.09403 0.43683 0.22 0.831
```

```

## Year2009          0.35467    0.54676    0.65    0.520
## Year2010          -0.10339    0.44105   -0.23    0.816
## Year2011          -0.18999    0.46104   -0.41    0.682
## Year2012          -0.10869    0.47847   -0.23    0.821
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.439
## Multiple R-squared:  0.212, Adjusted R-squared:  -0.0765
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 48 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.620  0.904  0.938  0.918  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      1.75e-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 0.8531 1          0.9236
## Year            0.8531 14          0.9943

```



```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -7.66e-01 -2.45e-01 1.11e-15 2.68e-01 9.75e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0017 0.4015 2.49 0.017 *
## LastAuthorFemale1 -0.3189 0.1614 -1.98 0.055 .
## Year1999 -0.0342 0.5046 -0.07 0.946
## Year2000 -0.3247 0.4015 -0.81 0.423
## Year2001 -0.1559 0.4245 -0.37 0.715
## Year2002 0.3343 0.4015 0.83 0.410
## Year2003 0.0638 0.4236 0.15 0.881
## Year2004 0.2813 0.4015 0.70 0.487
## Year2005 0.0641 0.4978 0.13 0.898
## Year2006 0.4730 0.4167 1.13 0.263
## Year2007 -0.0593 0.4239 -0.14 0.889
## Year2008 0.1595 0.4141 0.39 0.702
```

```

## Year2009          0.4335      0.4549      0.95      0.346
## Year2010          0.0548      0.4675      0.12      0.907
## Year2011         -0.1207      0.4476     -0.27      0.789
## Year2012          0.0722      0.4970      0.15      0.885
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.44
## Multiple R-squared:  0.271, Adjusted R-squared:  0.004
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 50 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.602  0.902  0.958  0.928  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      1.75e-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 1905"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1998 1999 2002 2003 2004 2005 2006 2010 2012
##    1    2    2    2    1    3    1    2    1
##
## 1998 1999 2002 2003 2004 2005 2006 2010 2012
##    1    1    0    1    1    3    0    2    0
##
## 1998 1999 2002 2003 2004 2005 2006 2010 2012
##    1    1    0    1    0    2    0    2    0
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"

```

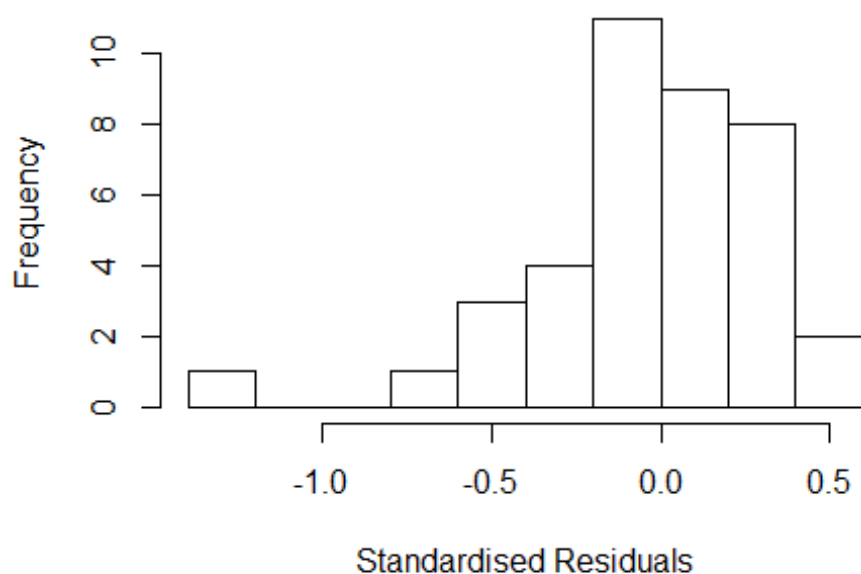


```

## [1] "Female first author team size 2018 geometric mean: NaN"
## [1] "Male first author team size 2018 geometric mean: 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 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
##    1    5    5    4    2    4   10   10    2    5    5    5    6    7    8
## 2011 2012
##    4    4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    2    2    2    1    3    4    5    0    4    4    3    3    5    4
## 2011 2012
##    1    2
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    1    2    2    1    2    4    4    0    3    4    2    3    4    4
## 2011 2012
##    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: 3.16227766016838"
## [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 1.687e+14  1      1.299e+07
## LastAuthorFemale  1.260e+00  1      1.122e+00
## Year              1.942e+14 14      3.238e+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
## -1.37e+00 -1.56e-01 8.88e-16 2.04e-01 5.04e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 8.89e-01 8.02e-09 1.11e+08 < 2e-16 ***
## FirstAuthorFemale1 1.78e-01 1.86e-01 9.60e-01 0.34868
## LastAuthorFemale1 1.47e-01 2.57e-01 5.70e-01 0.57164
## Year1998 5.46e-01 3.69e-02 1.48e+01 6.4e-13 ***
## Year1999 -4.32e-01 1.87e-01 -2.31e+00 0.03052 *
## Year2000 3.74e-01 1.86e-01 2.01e+00 0.05718 .
## Year2001 -3.05e-02 1.72e-01 -1.80e-01 0.86075
## Year2002 4.86e-01 5.40e-01 9.00e-01 0.37803
## Year2003 -7.71e-02 2.05e-01 -3.80e-01 0.71056
## Year2005 -1.40e-01 2.09e-01 -6.70e-01 0.50940
## Year2006 -1.45e-01 2.95e-01 -4.90e-01 0.62667
## Year2007 3.88e-01 3.01e-02 1.29e+01 1.0e-11 ***
```

```

## Year2008          2.56e-01   1.97e-01   1.30e+00   0.20790
## Year2009          1.04e-01   7.67e-02   1.36e+00   0.18861
## Year2010          4.03e-01   1.03e-01   3.90e+00   0.00077 ***
## Year2011         -8.80e-02   4.82e-09  -1.83e+07   < 2e-16 ***
## Year2012          3.10e-01   1.09e-01   2.84e+00   0.00952 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.441, Adjusted R-squared:  0.034
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 34 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.218  0.946  0.971  0.932  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      2.56e-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 -2.426e+16  1          NaN
## Year              -2.426e+16 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
## -1.45e+00 -1.82e-01  2.11e-15  2.02e-01  4.57e-01
##

```

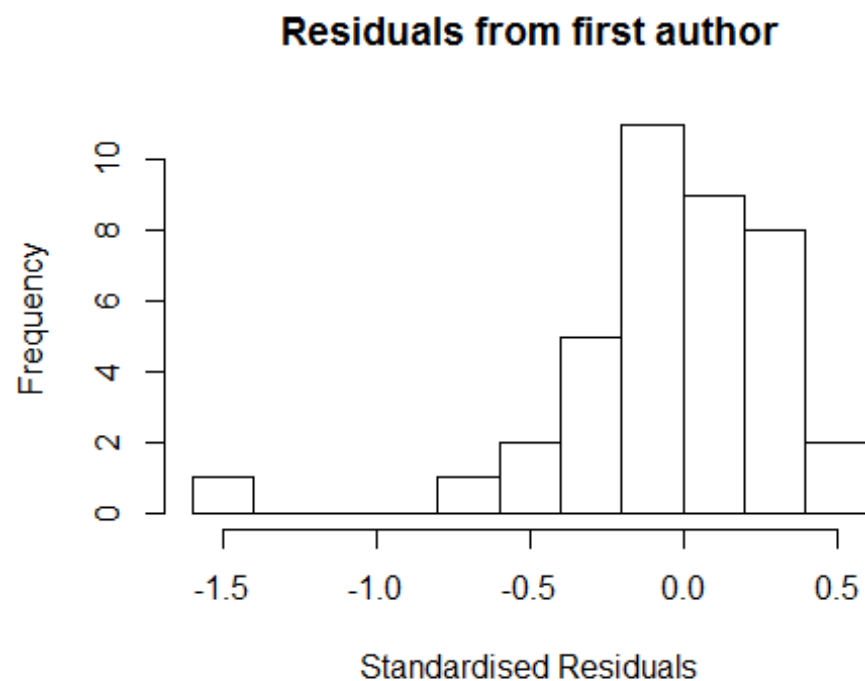
```

## Coefficients:
##               Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    8.89e-01   1.29e-08  6.87e+07 < 2e-16 ***
## FirstAuthorFemale1 1.48e-01   1.49e-01   9.90e-01  0.33333
## Year1998        5.46e-01   3.69e-02   1.48e+01  3.0e-13 ***
## Year1999       -4.01e-01   1.50e-01  -2.67e+00  0.01359 *
## Year2000        4.04e-01   1.49e-01   2.71e+00  0.01257 *
## Year2001       -3.05e-02   1.73e-01  -1.80e-01  0.86190
## Year2002        5.63e-01   4.27e-01   1.32e+00  0.19984
## Year2003       -6.27e-02   1.97e-01  -3.20e-01  0.75277
## Year2005       -1.37e-01   2.12e-01  -6.50e-01  0.52232
## Year2006       -1.28e-01   3.01e-01  -4.20e-01  0.67489
## Year2007        3.88e-01   3.01e-02   1.29e+01  5.4e-12 ***
## Year2008        2.66e-01   1.85e-01   1.43e+00  0.16496
## Year2009        1.71e-01   1.28e-01   1.34e+00  0.19453
## Year2010        4.03e-01   1.04e-01   3.88e+00  0.00076 ***
## Year2011       -8.80e-02   9.14e-09  -9.63e+06 < 2e-16 ***
## Year2012        3.25e-01   1.01e-01   3.22e+00  0.00382 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.377
## Multiple R-squared:  0.468, Adjusted R-squared:  0.121
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 34 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.105  0.918  0.967  0.923  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      2.56e-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"

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

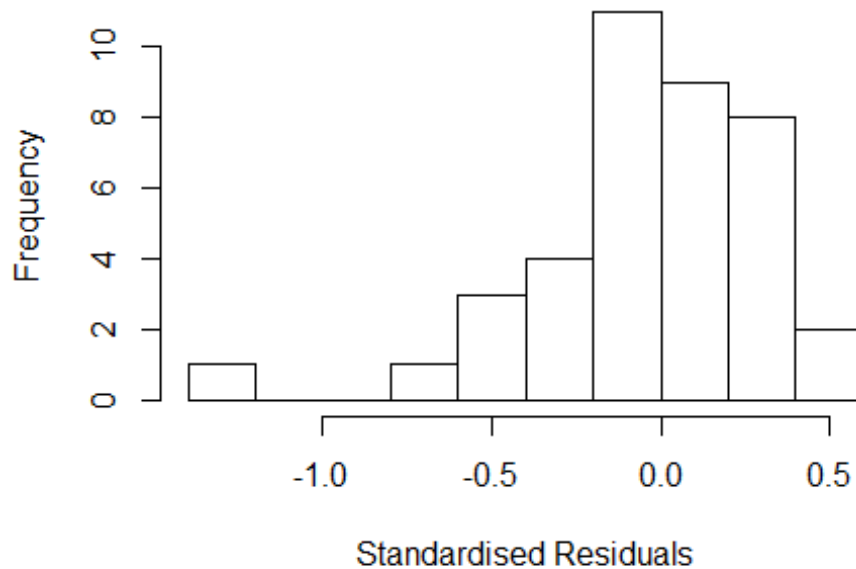
```

```
## 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 14          NaN
```

Residuals from last author



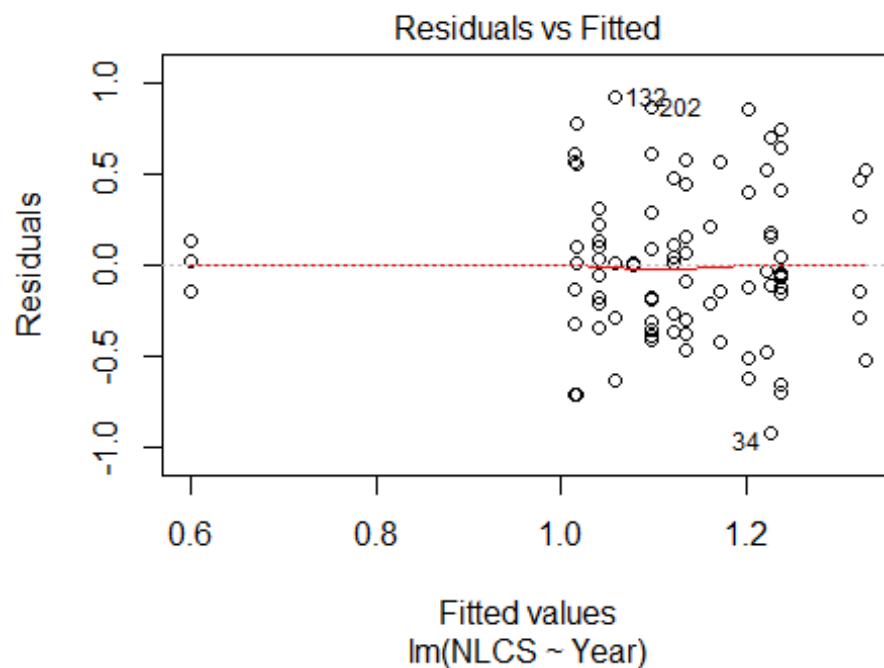
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.51e+00 -1.94e-01 2.08e-16 1.96e-01 5.45e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.88900 0.00000 Inf < 2e-16 ***
## LastAuthorFemale1 0.14750 0.25534 0.58 0.5691
## Year1998 0.54600 0.03688 14.81 3.0e-13 ***
## Year1999 -0.25350 0.01450 -17.48 8.9e-15 ***
## Year2000 0.55200 0.00000 Inf < 2e-16 ***
## Year2001 -0.03050 0.17127 -0.18 0.8602
## Year2002 0.62266 0.29077 2.14 0.0431 *
## Year2003 0.00101 0.19399 0.01 0.9959
## Year2005 -0.14082 0.20669 -0.68 0.5025
## Year2006 -0.09281 0.31128 -0.30 0.7683
## Year2007 0.38750 0.03011 12.87 5.4e-12 ***
## Year2008 0.30944 0.13545 2.28 0.0319 *
```

```

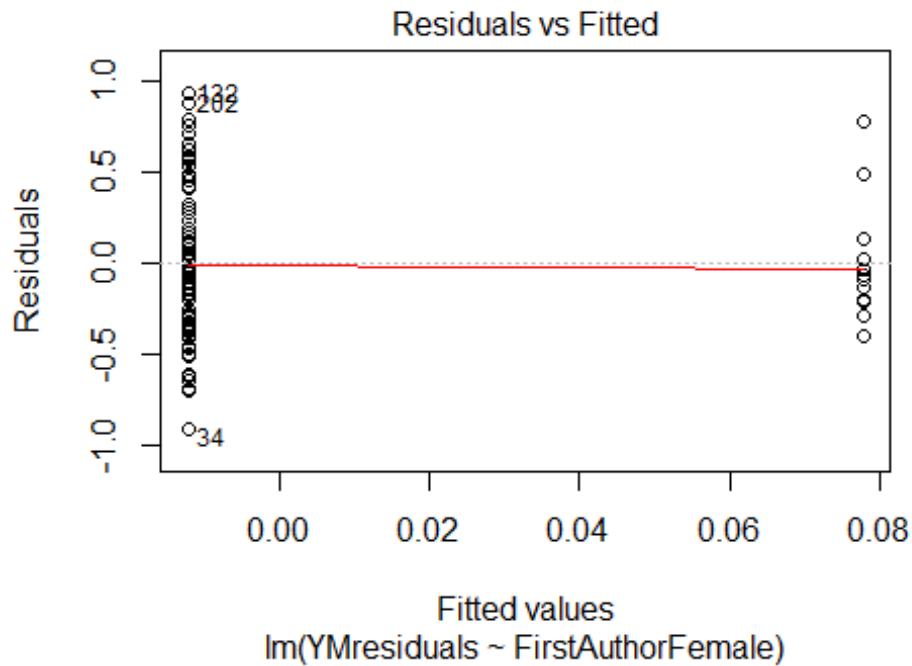
## Year2009          0.10400    0.07660    1.36    0.1877
## Year2010          0.40316    0.10307    3.91    0.0007 ***
## Year2011         -0.08800    0.00000   -Inf    < 2e-16 ***
## Year2012          0.39900    0.12330    3.24    0.0036 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.413
## Multiple R-squared:  0.441, Adjusted R-squared:  0.0766
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 30 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.151  0.929  0.967  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      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 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
##   11  12   9  11  11  10  12  11   5  16  15  14   9   3   6
## 2011 2012
##    3   6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5   6   5   3   3   2   5   8   4   9  11  10   5   2   6
## 2011 2012
##    3   2
##

```

```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    5    5    3    3    2    5    7    2    6   11    9    5    2    6
## 2011 2012
##    3    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 = 20, df = 16, p-value = 0.2
```

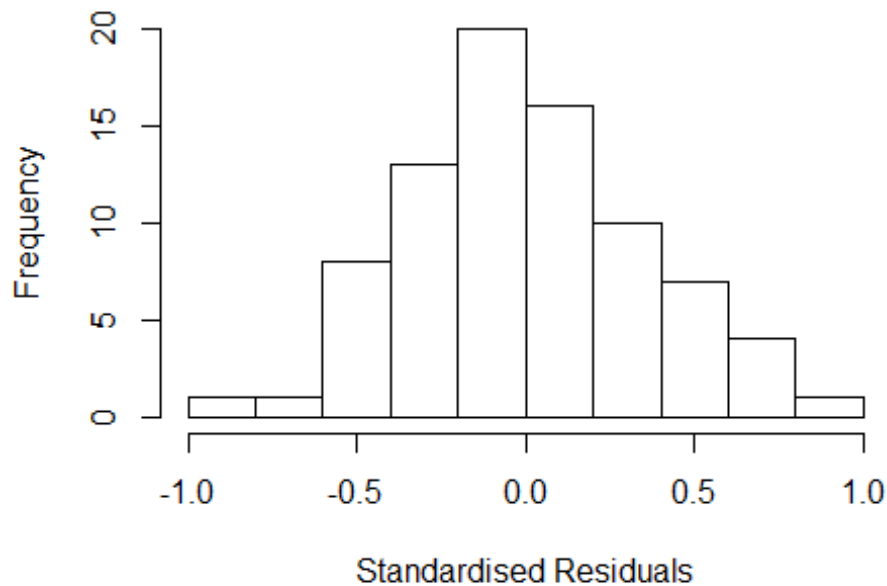


```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.1, df = 1, p-value = 0.3
```

```
## [1] "Female first author team size 2018 geometric mean: NaN"
## [1] "Male first author team size 2018 geometric mean: 3.66284150148471"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale  40.082  1         6.331
## LastAuthorFemale   5.944  1         2.438
## UniqueAuthors    105.374  4         1.790
## Year              3290.618 16         1.288
```

Residuals from first and last author and team size



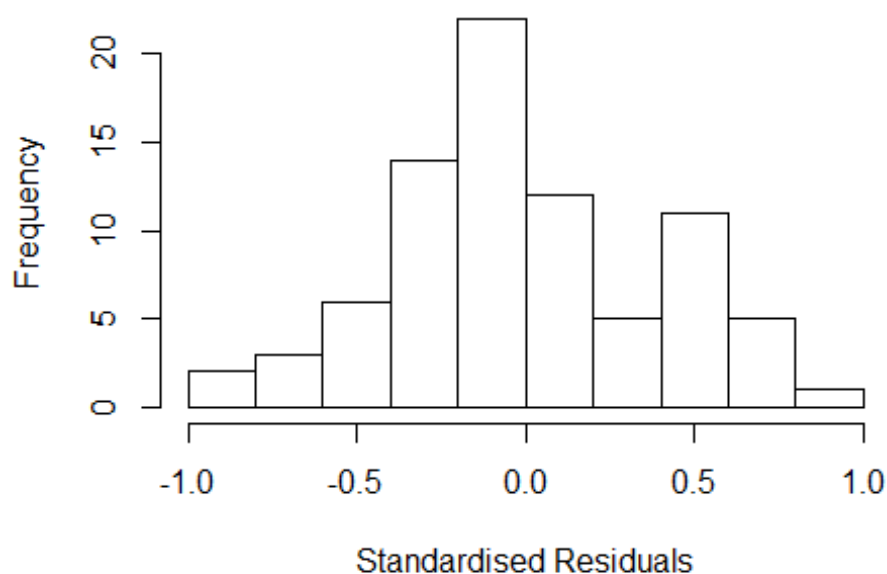
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8755 -0.2285 -0.0271 0.2205 0.9222
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3122 0.1512 8.68 4.5e-12 ***
## FirstAuthorFemale1 -0.1688 0.1923 -0.88 0.38370
## LastAuthorFemale1 -0.0980 0.1225 -0.80 0.42685
## UniqueAuthors2 -0.0414 0.1338 -0.31 0.75789
## UniqueAuthors3 0.5215 0.1859 2.80 0.00684 **
## UniqueAuthors4 0.3928 0.1624 2.42 0.01873 *
## UniqueAuthors5 0.5238 0.3973 1.32 0.19253
## Year1997 -0.3220 0.3446 -0.93 0.35401
## Year1998 -0.1297 0.3563 -0.36 0.71723
## Year1999 -0.6413 0.1927 -3.33 0.00152 **
```

```

## Year2000          -0.1485      0.3365    -0.44    0.66076
## Year2001           0.0153      0.5744     0.03    0.97886
## Year2002          -0.3106      0.2441    -1.27    0.20819
## Year2003          -0.2949      0.2165    -1.36    0.17832
## Year2004          -0.7197      0.1981    -3.63    0.00059 ***
## Year2005          -0.4359      0.1662    -2.62    0.01111 *
## Year2006          -0.2554      0.2468    -1.03    0.30502
## Year2007          -0.3411      0.1786    -1.91    0.06118 .
## Year2008          -0.3820      0.4277    -0.89    0.37549
## Year2009          -0.1003      0.2068    -0.48    0.62955
## Year2010          -0.3368      0.2201    -1.53    0.13147
## Year2011          -0.1894      0.2487    -0.76    0.44934
## Year2012          -0.1591      0.2774    -0.57    0.56856
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.397
## Multiple R-squared:  0.325, Adjusted R-squared:  0.0696
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 75 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.569  0.889  0.970  0.925  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          1.23e-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 2: First author gender, Last author gender, Year as
factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 5.328 1          2.308
## LastAuthorFemale  7.423 1          2.725
## Year              24.151 16          1.105

```

Residuals from first and last author



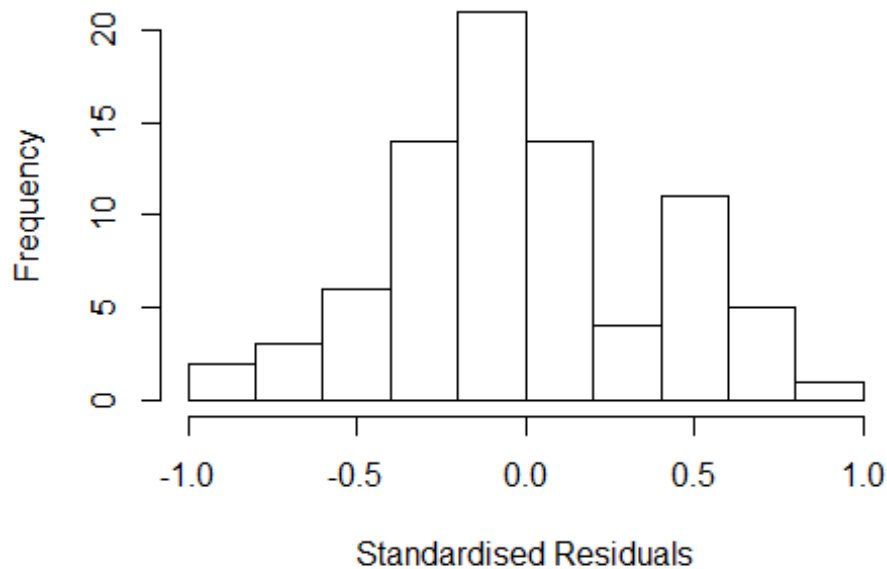
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9677 -0.2390 -0.0731 0.2434 0.9759
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3145 0.1470 8.94 9.4e-13 ***
## FirstAuthorFemale1 0.1501 0.1689 0.89 0.37758
## LastAuthorFemale1 -0.1541 0.1646 -0.94 0.35271
## Year1997 -0.1182 0.2984 -0.40 0.69337
## Year1998 -0.0398 0.2924 -0.14 0.89229
## Year1999 -0.7655 0.1757 -4.36 5.0e-05 ***
## Year2000 -0.1620 0.3077 -0.53 0.60037
## Year2001 0.0130 0.5116 0.03 0.97974
## Year2002 -0.1609 0.3155 -0.51 0.61195
## Year2003 -0.1709 0.2158 -0.79 0.43124
## Year2004 -0.7220 0.1941 -3.72 0.00043 ***
## Year2005 -0.4107 0.1584 -2.59 0.01188 *
```

```

## Year2006          -0.1166      0.2158   -0.54  0.59101
## Year2007          -0.3263      0.1915   -1.70  0.09341 .
## Year2008          -0.4091      0.3826   -1.07  0.28906
## Year2009          -0.2340      0.1834   -1.28  0.20677
## Year2010          -0.1728      0.1983   -0.87  0.38681
## Year2011          -0.0983      0.3023   -0.33  0.74622
## Year2012          -0.3045      0.2743   -1.11  0.27114
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.456
## Multiple R-squared:  0.165, Adjusted R-squared:  -0.0779
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 74 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.627  0.890  0.965  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      1.23e-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"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.63 1      1.905
## Year              3.63 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
## -0.9682 -0.2374 -0.0728 0.2130 0.9762
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3144 0.1471 8.93 8.4e-13 ***
## FirstAuthorFemale1 0.1451 0.1683 0.86 0.39191
## Year1997 -0.1178 0.2987 -0.39 0.69469
## Year1998 -0.0392 0.2926 -0.13 0.89388
## Year1999 -0.7637 0.1755 -4.35 5.0e-05 ***
## Year2000 -0.1621 0.3080 -0.53 0.60042
## Year2001 0.0131 0.5130 0.03 0.97971
## Year2002 -0.1605 0.3170 -0.51 0.61437
## Year2003 -0.1709 0.2162 -0.79 0.43207
## Year2004 -0.7219 0.1942 -3.72 0.00043 ***
## Year2005 -0.4098 0.1586 -2.58 0.01211 *
## Year2006 -0.1155 0.2159 -0.54 0.59448
```

```

## Year2007          -0.3266      0.1915   -1.71  0.09307 .
## Year2008          -0.4057      0.3831   -1.06  0.29371
## Year2009          -0.3084      0.1793   -1.72  0.09027 .
## Year2010          -0.2240      0.1911   -1.17  0.24559
## Year2011          -0.0982      0.3026   -0.32  0.74646
## Year2012          -0.2995      0.2741   -1.09  0.27872
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.454
## Multiple R-squared:  0.162, Adjusted R-squared:  -0.0641
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 76 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.624  0.889  0.967  0.929  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      1.23e-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 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 3.191 1      1.786
## Year      3.191 16      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
## -0.9688 -0.2454 -0.0633  0.2130  0.9766

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3143    0.1472   8.93 8.6e-13 ***
## LastAuthorFemale1 -0.1397    0.1640  -0.85 0.39756
## Year1997          -0.1173    0.2989  -0.39 0.69615
## Year1998          -0.0385    0.2927  -0.13 0.89568
## Year1999          -0.7145    0.1611  -4.43 3.8e-05 ***
## Year2000          -0.1623    0.3083  -0.53 0.60049
## Year2001           0.0132    0.5147   0.03 0.97967
## Year2002          -0.1517    0.3483  -0.44 0.66476
## Year2003          -0.1709    0.2161  -0.79 0.43204
## Year2004          -0.7218    0.1943  -3.71 0.00043 ***
## Year2005          -0.3857    0.1597  -2.42 0.01862 *
## Year2006          -0.0820    0.2028  -0.40 0.68742
## Year2007          -0.3270    0.1915  -1.71 0.09268 .
## Year2008          -0.3006    0.3237  -0.93 0.35667
## Year2009          -0.1660    0.1740  -0.95 0.34380
## Year2010          -0.1504    0.1929  -0.78 0.43853
## Year2011          -0.0982    0.3029  -0.32 0.74677
## Year2012          -0.1543    0.2152  -0.72 0.47600
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.452
## Multiple R-squared:  0.156, Adjusted R-squared:  -0.0711
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 75 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.620  0.893  0.965  0.927  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      1.23e-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: 81"
## [1] ""

```

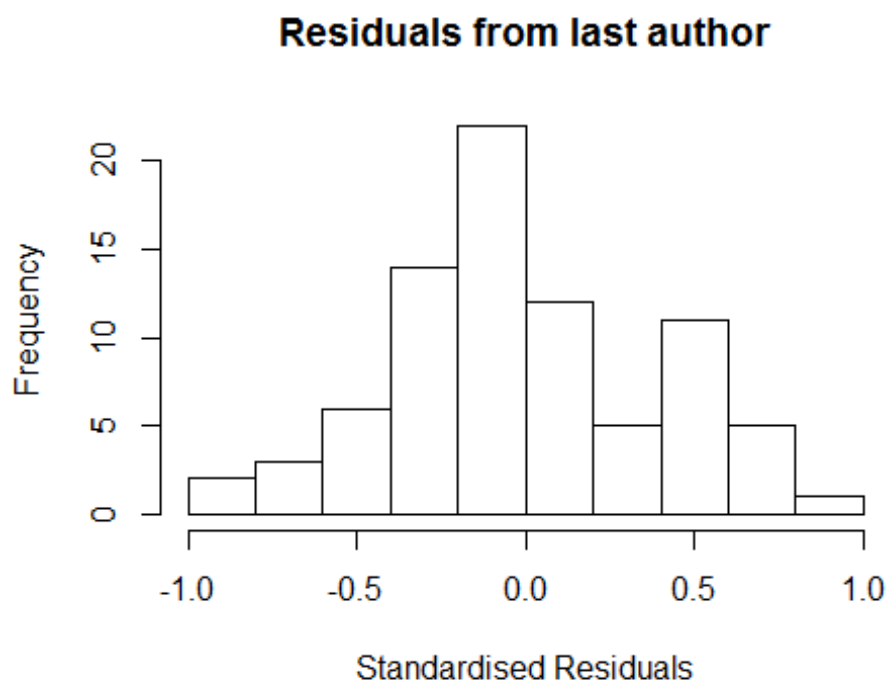


```

## [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
##    4    4    6    5    2    5    5    9    4    2    4    4    10    6    10
## 2011 2012
##    6    6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    2    2    3    1    4    3    5    2    2    2    2    8    3    6
## 2011 2012
##    3    4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    2    2    3    1    3    3    4    1    1    2    2    7    2    6
## 2011 2012
##    2    4
## [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: 3.16227766016838"
##
## 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: 3.41995189335339"
## [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

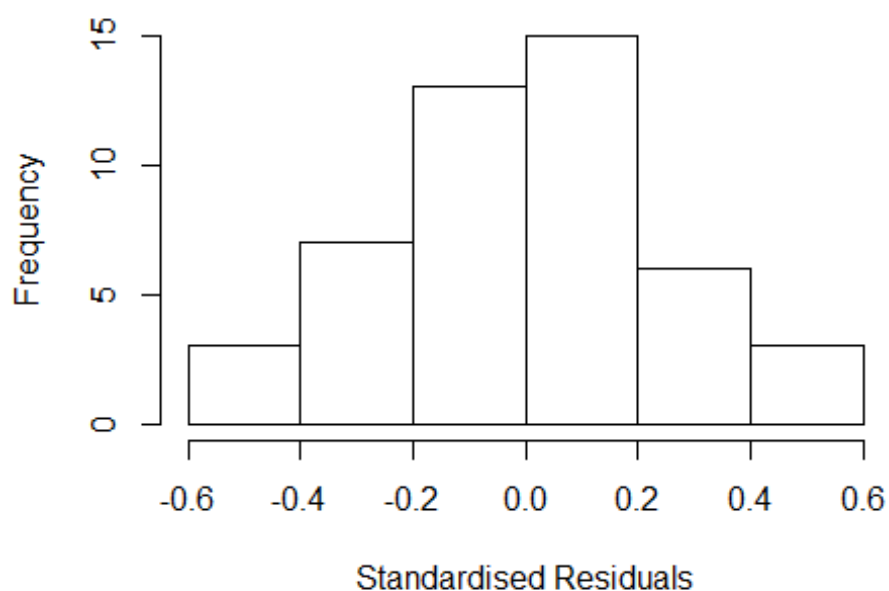
```



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

| | GVIF | Df | GVIF^(1/(2*Df)) |
|----------------------|------------|----|-----------------|
| ## FirstAuthorFemale | -2.077e+03 | 1 | NaN |
| ## LastAuthorFemale | 8.724e+01 | 1 | 9.34 |
| ## Year | -7.906e+14 | 16 | NaN |

Residuals from first and last author



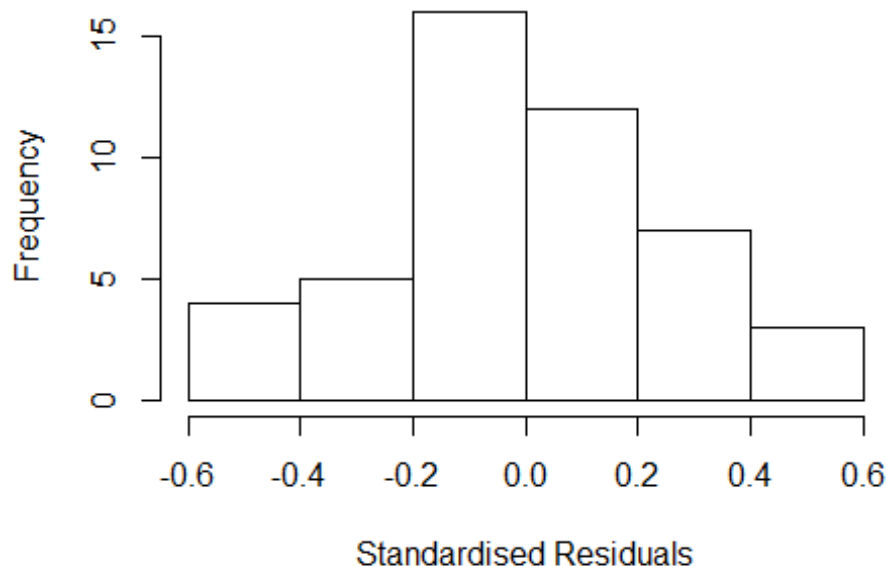
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.59800 -0.17970 0.00328 0.18514 0.59800
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9080 0.9616 0.94 0.353
## FirstAuthorFemale1 -0.3207 0.1705 -1.88 0.070 .
## LastAuthorFemale1 -0.2221 0.1245 -1.78 0.085 .
## Year1997 0.2515 0.9850 0.26 0.800
## Year1998 0.3925 0.9716 0.40 0.689
## Year1999 0.3722 0.9749 0.38 0.706
## Year2000 0.8537 0.9802 0.87 0.391
## Year2001 -0.0702 0.9728 -0.07 0.943
## Year2002 0.4713 0.9811 0.48 0.635
## Year2003 0.2201 0.9785 0.22 0.824
## Year2004 -0.9080 0.9616 -0.94 0.353
## Year2005 0.0870 0.9616 0.09 0.929
```

```

## Year2006          0.0535      1.0037      0.05      0.958
## Year2007          0.3685      0.9621      0.38      0.705
## Year2008          0.3117      0.9639      0.32      0.749
## Year2009          0.1690      0.9639      0.18      0.862
## Year2010          0.3507      0.9656      0.36      0.719
## Year2011          0.1180      0.9767      0.12      0.905
## Year2012          0.2830      0.9892      0.29      0.777
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.364
## Multiple R-squared:  0.433, Adjusted R-squared:  0.069
## Convergence in 29 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 37 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.769  0.938  0.971  0.943  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          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 9.571e+12 1      3.094e+06
## Year              9.571e+12 16      2.545e+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
## -5.98e-01 -1.67e-01 2.22e-16 1.74e-01 5.98e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9080 0.9918 0.92 0.367
## FirstAuthorFemale1 -0.2985 0.1610 -1.85 0.074 .
## Year1997 0.2515 1.0146 0.25 0.806
## Year1998 0.3925 1.0015 0.39 0.698
## Year1999 0.2847 0.9997 0.28 0.778
## Year2000 0.8315 1.0012 0.83 0.413
## Year2001 -0.0782 1.0004 -0.08 0.938
## Year2002 0.4575 1.0041 0.46 0.652
## Year2003 0.2203 1.0061 0.22 0.828
## Year2004 -0.9080 0.9918 -0.92 0.367
## Year2005 0.0870 0.9918 0.09 0.931
## Year2006 0.0535 1.0330 0.05 0.959
```

```

## Year2007          0.3685      0.9923      0.37      0.713
## Year2008          0.3084      0.9930      0.31      0.758
## Year2009          0.0580      0.9966      0.06      0.954
## Year2010          0.2808      0.9966      0.28      0.780
## Year2011          0.1180      1.0065      0.12      0.907
## Year2012          0.2800      1.0164      0.28      0.785
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.36
## Multiple R-squared:  0.413, Adjusted R-squared:  0.0696
## Convergence in 30 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 38 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.765  0.921  0.969  0.941  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-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 9.791 1      3.129
## Year              9.791 16      1.074

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -0.598 -0.177  0.000  0.178  0.598

```

```

##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9080    0.9124   1.00   0.33
## LastAuthorFemale1 -0.1864    0.1547  -1.21   0.24
## Year1997         0.2515    0.9368   0.27   0.79
## Year1998         0.3925    0.9228   0.43   0.67
## Year1999         0.2733    0.9493   0.29   0.78
## Year2000         0.5330    0.9124   0.58   0.56
## Year2001        -0.1824    0.9273  -0.20   0.85
## Year2002         0.2624    0.9393   0.28   0.78
## Year2003         0.2198    0.9279   0.24   0.81
## Year2004        -0.9080    0.9124  -1.00   0.33
## Year2005         0.0870    0.9124   0.10   0.92
## Year2006         0.0535    0.9559   0.06   0.96
## Year2007         0.3685    0.9129   0.40   0.69
## Year2008         0.2661    0.9147   0.29   0.77
## Year2009         0.1512    0.9156   0.17   0.87
## Year2010         0.3393    0.9167   0.37   0.71
## Year2011         0.1180    0.9282   0.13   0.90
## Year2012         0.2194    0.9256   0.24   0.81
##
## Robust residual standard error: 0.371
## Multiple R-squared:  0.373, Adjusted R-squared:  0.0052
## Convergence in 27 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 42 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.777  0.926  0.967  0.947  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] "Sample size for the above analysis: 47"
## [1] ""
## [1] ""
## [1] "#####"

```

```

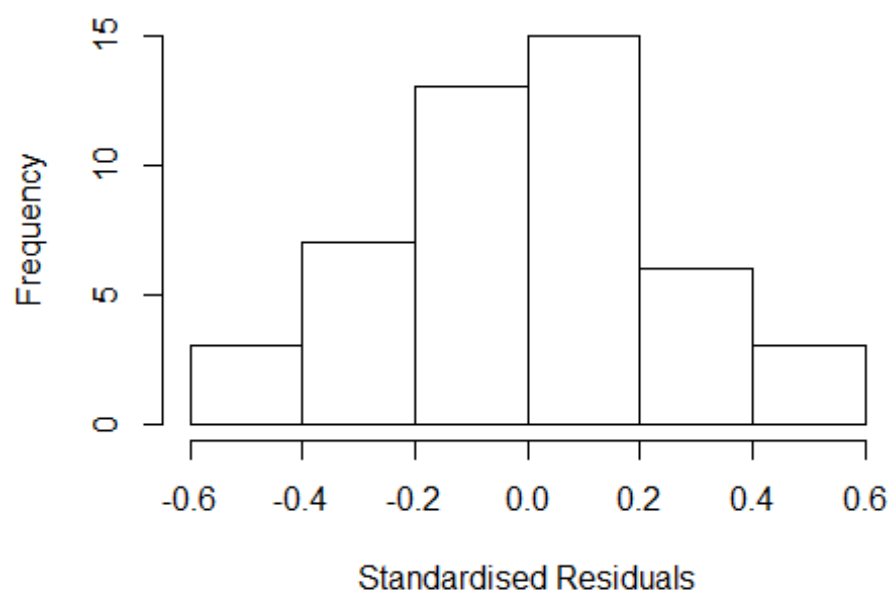
## [1] "Analysis of AJSC 1909"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    1    1    6    5    3    6    2    5    4    7   12    4    8
## 2011 2012
##    6    9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    0    0    0    2    3    3    4    1    4    3    4    6    3    7
## 2011 2012
##    5    4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    0    0    0    2    3    3    4    1    4    3    4    5    3    7
## 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"
## [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

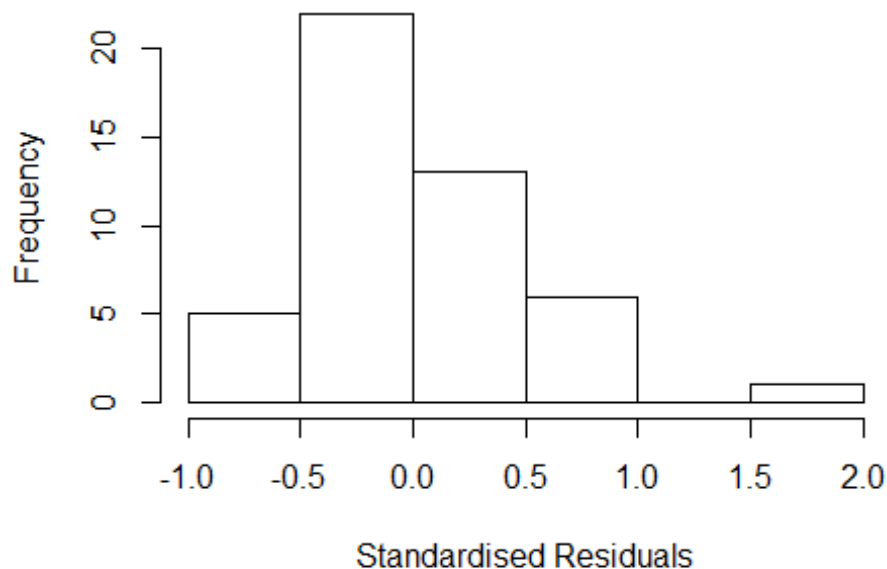
```


Residuals from last author



| | | | |
|----------------------|------|----|-------------------------|
| ## | GVIF | Df | $GVIF^{1/(2 \cdot Df)}$ |
| ## FirstAuthorFemale | NaN | 1 | NaN |
| ## LastAuthorFemale | NaN | 1 | NaN |
| ## UniqueAuthors | NaN | 3 | NaN |
| ## Year | NaN | 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
## -7.78e-01 -2.66e-01 1.11e-16 2.42e-01 1.70e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.52e-15 0.00e+00 -Inf < 2e-16 ***
## FirstAuthorFemale1 -1.23e-01 2.90e-01 -0.42 0.6763
## LastAuthorFemale1 5.63e-01 3.02e-01 1.86 0.0733 .
## UniqueAuthors2 8.99e-02 3.06e-01 0.29 0.7709
## UniqueAuthors3 3.54e-02 2.17e-01 0.16 0.8714
## UniqueAuthors4 7.48e-01 4.97e-01 1.51 0.1434
## Year2000 -8.99e-02 3.06e-01 -0.29 0.7709
## Year2001 9.51e-01 1.28e-01 7.43 4.4e-08 ***
## Year2002 8.97e-01 2.99e-01 3.00 0.0057 **
## Year2003 1.37e+00 5.76e-01 2.38 0.0243 *
```

```

## Year2004          6.38e-01  0.00e+00  Inf < 2e-16 ***
## Year2005          1.58e+00  2.88e-01  5.49 7.3e-06 ***
## Year2006          1.53e+00  2.91e-01  5.26 1.4e-05 ***
## Year2007          1.29e+00  2.73e-01  4.73 5.7e-05 ***
## Year2008          1.06e+00  3.26e-01  3.26 0.0030 **
## Year2009          1.34e+00  2.95e-01  4.54 9.8e-05 ***
## Year2010          1.13e+00  2.16e-01  5.24 1.4e-05 ***
## Year2011          1.39e+00  2.08e-01  6.66 3.2e-07 ***
## Year2012          5.78e-01  3.61e-01  1.60 0.1209
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.501
## Multiple R-squared:  0.541, Adjusted R-squared:  0.245
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.226  0.893  0.967   0.926   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          2.13e-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 5.867e+02 1 24.22
## LastAuthorFemale -4.608e+15 1 NaN
## Year -1.813e+18 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,

```

```

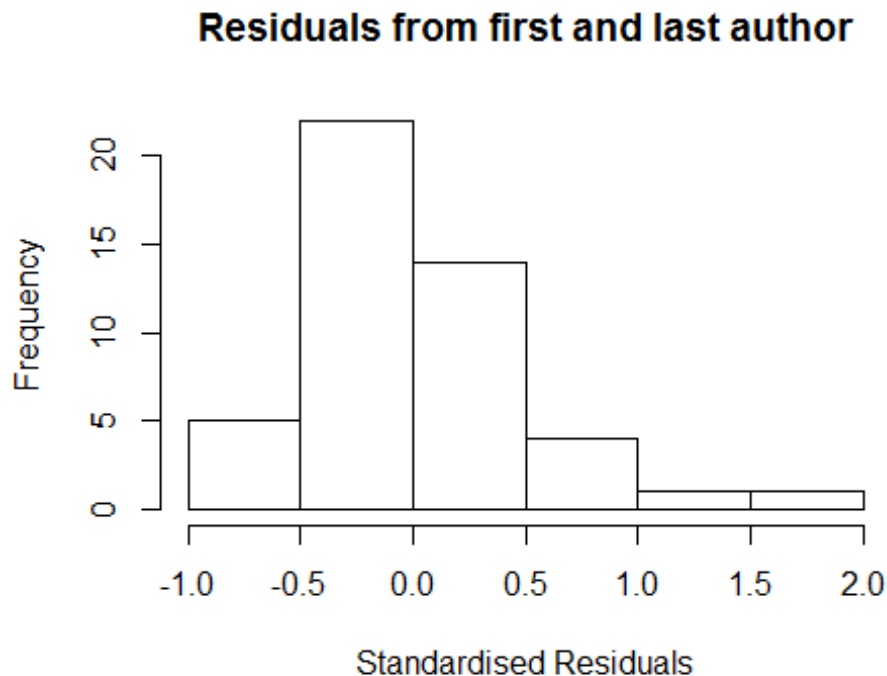
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q   Median      3Q      Max
## -7.80e-01 -2.37e-01 -2.22e-16  2.61e-01  1.78e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   -2.87e-16   1.61e-08   0.00e+00  1.00000
## FirstAuthorFemale1 -1.32e-01   1.91e-01  -6.90e-01  0.49545
## LastAuthorFemale1  5.92e-01   2.74e-01   2.16e+00  0.03866 *
## Year2000         4.43e-16   2.01e-08   0.00e+00  1.00000
## Year2001         9.82e-01   9.52e-02   1.03e+01  1.5e-11 ***
## Year2002         8.99e-01   2.92e-01   3.08e+00  0.00433 **
## Year2003         1.38e+00   4.52e-01   3.06e+00  0.00454 **
## Year2004         6.38e-01   2.04e-08   3.12e+07  < 2e-16 ***
## Year2005         1.62e+00   2.41e-01   6.74e+00  1.5e-07 ***
## Year2006         1.53e+00   2.86e-01   5.34e+00  8.0e-06 ***
## Year2007         1.35e+00   2.74e-01   4.94e+00  2.6e-05 ***
## Year2008         1.14e+00   1.45e-01   7.85e+00  7.3e-09 ***
## Year2009         1.50e+00   5.82e-01   2.57e+00  0.01512 *
## Year2010         1.30e+00   1.20e-01   1.08e+01  4.6e-12 ***
## Year2011         1.42e+00   2.04e-01   6.97e+00  8.1e-08 ***
## Year2012         9.97e-01   2.66e-01   3.75e+00  0.00073 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.519
## Multiple R-squared:  0.484, Adjusted R-squared:  0.235
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 40 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.216  0.915   0.962   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      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"

## 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

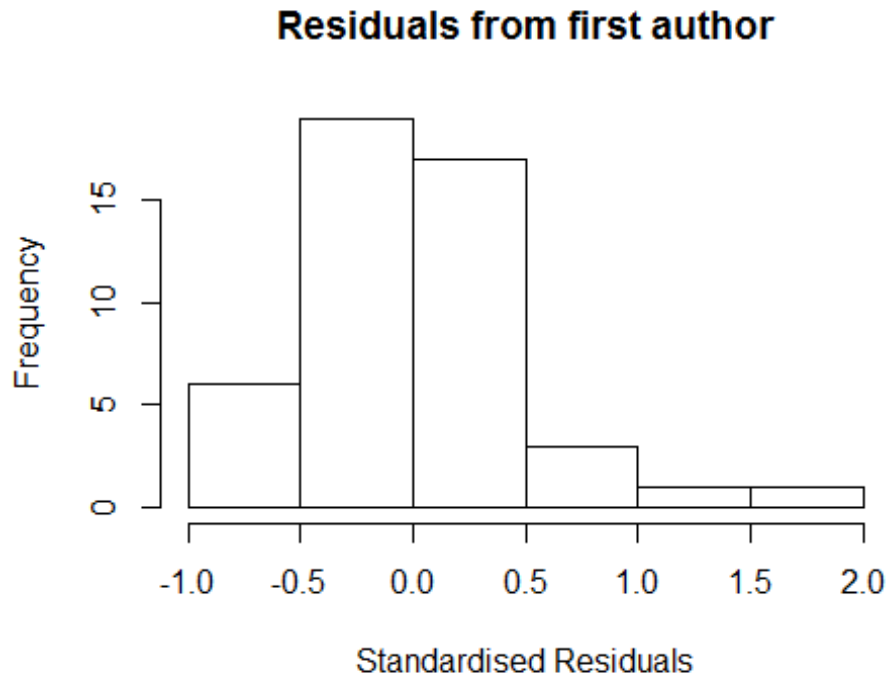
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q   Median      3Q      Max
## -7.79e-01 -2.58e-01 -1.73e-16  2.98e-01  1.76e+00
##
## Coefficients:
```

```

##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      4.55e-16   0.00e+00    Inf < 2e-16 ***
## FirstAuthorFemale1 -1.32e-01   1.91e-01   -0.69  0.49459
## Year2000          -2.82e-16   0.00e+00   -Inf < 2e-16 ***
## Year2001           9.82e-01   9.52e-02   10.32  1.0e-11 ***
## Year2002           9.01e-01   2.91e-01    3.10  0.00404 **
## Year2003           1.40e+00   4.69e-01    2.99  0.00535 **
## Year2004           6.38e-01   0.00e+00    Inf < 2e-16 ***
## Year2005           1.62e+00   2.40e-01    6.77  1.2e-07 ***
## Year2006           1.53e+00   2.84e-01    5.37  6.8e-06 ***
## Year2007           1.52e+00   2.55e-01    5.96  1.2e-06 ***
## Year2008           1.13e+00   1.44e-01    7.86  5.8e-09 ***
## Year2009           1.51e+00   5.77e-01    2.61  0.01369 *
## Year2010           1.30e+00   1.20e-01   10.84  3.0e-12 ***
## Year2011           1.42e+00   2.04e-01    6.98  6.6e-08 ***
## Year2012           9.97e-01   2.65e-01    3.76  0.00068 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.53
## Multiple R-squared:  0.463, Adjusted R-squared:  0.228
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 41 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.248  0.920  0.962  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           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"
## 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
```



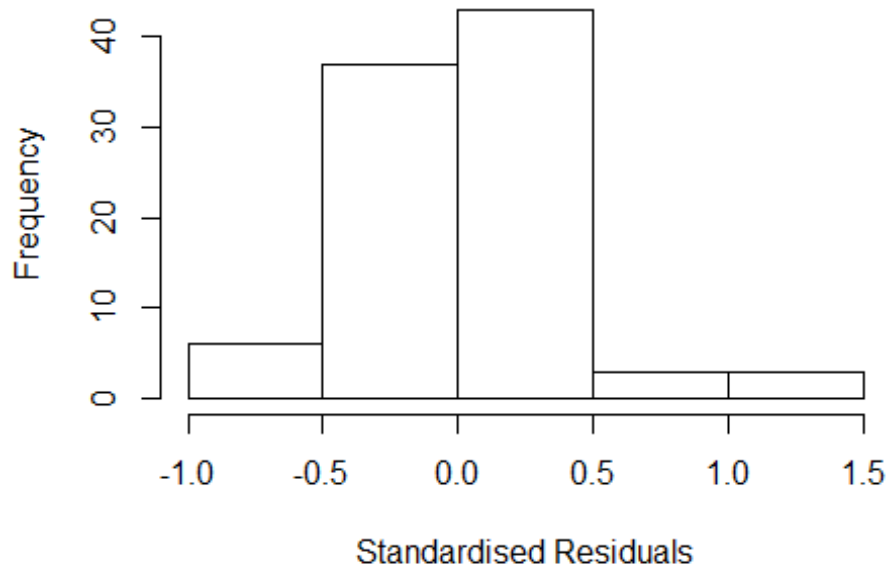
```
## [1] "Sample size for the above analysis: 47"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1910"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    9    5    6    4    6    6    6   11   14   15   10    9   13   16
## 2011 2012
##   16   19
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    5    3    2    1    4    2    3    5    6    9    8    4   11   12
## 2011 2012
##   10    9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    5    3    2    1    4    2    3    4    6    9    8    4   10   11
## 2011 2012
##   10    7
```

```

## [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: 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: 1"
## [1] "Male last 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] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale    5.048  1         2.247
## LastAuthorFemale     3.529  1         1.879
## UniqueAuthors       75.463  4         1.717
## Year                 308.549 16         1.196

```


Residuals from first and last author and team size



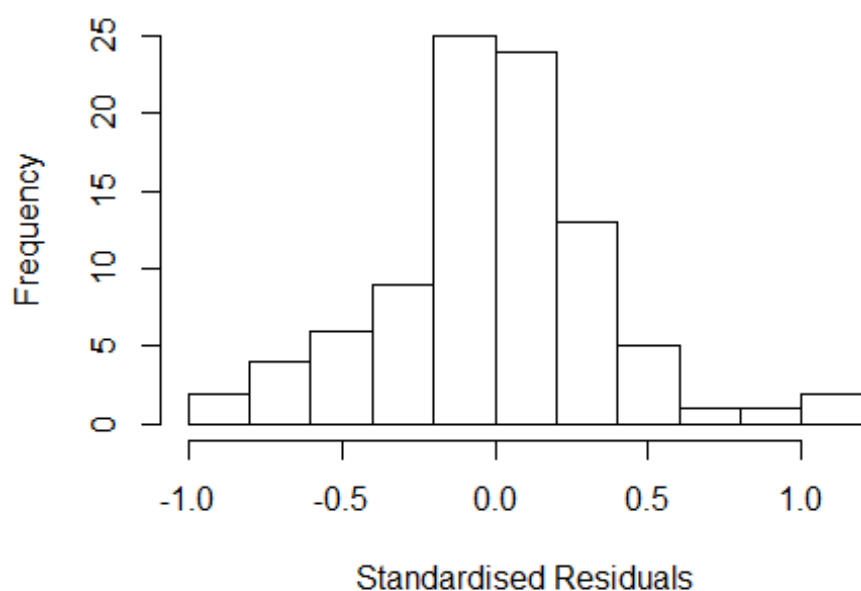
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9403 -0.2108 0.0192 0.1860 1.3901
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.0949 0.3791 0.25 0.8031
## FirstAuthorFemale1 0.1884 0.0934 2.02 0.0475 *
## LastAuthorFemale1 0.0615 0.1222 0.50 0.6165
## UniqueAuthors2 0.0292 0.1087 0.27 0.7886
## UniqueAuthors3 0.2365 0.1175 2.01 0.0481 *
## UniqueAuthors4 0.1841 0.1249 1.47 0.1450
## UniqueAuthors5 0.2610 0.2196 1.19 0.2388
## Year1997 0.6367 0.4164 1.53 0.1308
## Year1998 0.9131 0.4928 1.85 0.0682 .
## Year1999 0.5708 0.3971 1.44 0.1551
```

```

## Year2000      0.9671      0.3791      2.55      0.0130 *
## Year2001      0.7337      0.3825      1.92      0.0592 .
## Year2002      1.0993      0.4012      2.74      0.0078 **
## Year2003      0.7952      0.3810      2.09      0.0406 *
## Year2004      0.7884      0.3799      2.08      0.0417 *
## Year2005      0.6635      0.4149      1.60      0.1144
## Year2006      0.6873      0.3787      1.81      0.0739 .
## Year2007      0.8880      0.4137      2.15      0.0353 *
## Year2008      0.7047      0.5195      1.36      0.1793
## Year2009      0.5359      0.3988      1.34      0.1835
## Year2010      0.7001      0.3821      1.83      0.0712 .
## Year2011      0.6613      0.4014      1.65      0.1040
## Year2012      0.7339      0.4104      1.79      0.0781 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.35
## Multiple R-squared:  0.308, Adjusted R-squared:  0.0874
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 82 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0797 0.9080 0.9650 0.9000 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      1.09e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500      50      2      1      1000      200
## trace.lev      mts      compute.rd
##      0      1000      0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 6.673 1      2.583
## LastAuthorFemale  2.831 1      1.683
## Year             14.183 16      1.086

```

Residuals from first and last author



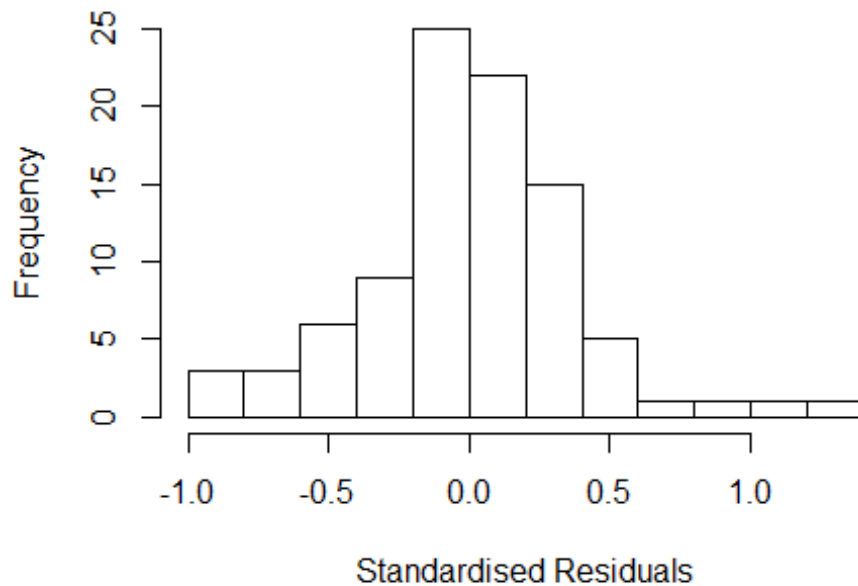
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.91638 -0.17795 0.00951 0.19607 1.17605
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.30895 0.45201 0.68 0.496
## FirstAuthorFemale1 0.17527 0.09408 1.86 0.066 .
## LastAuthorFemale1 0.00682 0.13268 0.05 0.959
## Year1997 0.48771 0.50492 0.97 0.337
## Year1998 0.70942 0.52603 1.35 0.182
## Year1999 0.37792 0.45801 0.83 0.412
## Year2000 0.75305 0.45201 1.67 0.100 .
## Year2001 0.58667 0.46566 1.26 0.212
## Year2002 0.90642 0.46064 1.97 0.053 .
## Year2003 0.65468 0.46308 1.41 0.162
## Year2004 0.69286 0.46159 1.50 0.138
## Year2005 0.44606 0.48202 0.93 0.358
```

```

## Year2006          0.61244      0.47447      1.29      0.201
## Year2007          0.84498      0.46688      1.81      0.074 .
## Year2008          0.54407      0.57934      0.94      0.351
## Year2009          0.42103      0.47884      0.88      0.382
## Year2010          0.61208      0.46971      1.30      0.197
## Year2011          0.60744      0.48826      1.24      0.217
## Year2012          0.66964      0.46018      1.46      0.150
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.372
## Multiple R-squared:  0.221, Adjusted R-squared:  0.0291
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 83 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.296  0.896  0.970  0.908  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      1.09e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 6.073 1      2.464
## Year              6.073 16      1.058

```

Residuals from first author



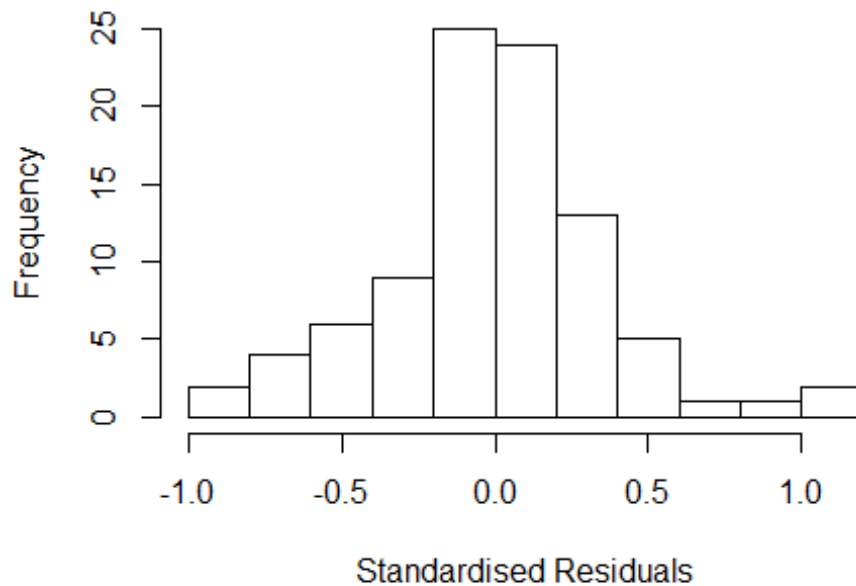
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.91665 -0.17560 0.00805 0.20111 1.22854
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.257 0.373 0.69 0.494
## FirstAuthorFemale1 0.176 0.090 1.96 0.054 .
## Year1997 0.549 0.429 1.28 0.205
## Year1998 0.764 0.463 1.65 0.103
## Year1999 0.430 0.380 1.13 0.261
## Year2000 0.805 0.373 2.16 0.034 *
## Year2001 0.639 0.390 1.64 0.105
## Year2002 0.959 0.383 2.50 0.015 *
## Year2003 0.707 0.385 1.83 0.071 .
## Year2004 0.745 0.384 1.94 0.056 .
## Year2005 0.501 0.408 1.23 0.223
## Year2006 0.659 0.400 1.65 0.103
```

```

## Year2007          0.897      0.425      2.11      0.038 *
## Year2008          0.606      0.525      1.16      0.252
## Year2009          0.476      0.393      1.21      0.230
## Year2010          0.667      0.387      1.72      0.089 .
## Year2011          0.660      0.415      1.59      0.116
## Year2012          0.721      0.382      1.89      0.063 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.354
## Multiple R-squared:  0.235, Adjusted R-squared:  0.0591
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 82 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.204  0.883  0.966  0.898  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      1.09e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 2.394 1      1.547
## Year              2.394 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
## -0.9846 -0.2099 0.0183 0.1946 1.1949
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.2901 0.4440 0.65 0.516
## LastAuthorFemale1 0.0819 0.1190 0.69 0.494
## Year1997 0.5336 0.5024 1.06 0.292
## Year1998 0.7290 0.5241 1.39 0.168
## Year1999 0.4844 0.4452 1.09 0.280
## Year2000 0.7719 0.4440 1.74 0.086 .
## Year2001 0.6056 0.4586 1.32 0.191
## Year2002 1.0129 0.4441 2.28 0.025 *
## Year2003 0.7332 0.4491 1.63 0.107
## Year2004 0.7988 0.4555 1.75 0.084 .
## Year2005 0.4659 0.4759 0.98 0.331
## Year2006 0.6864 0.4391 1.56 0.122
```

```

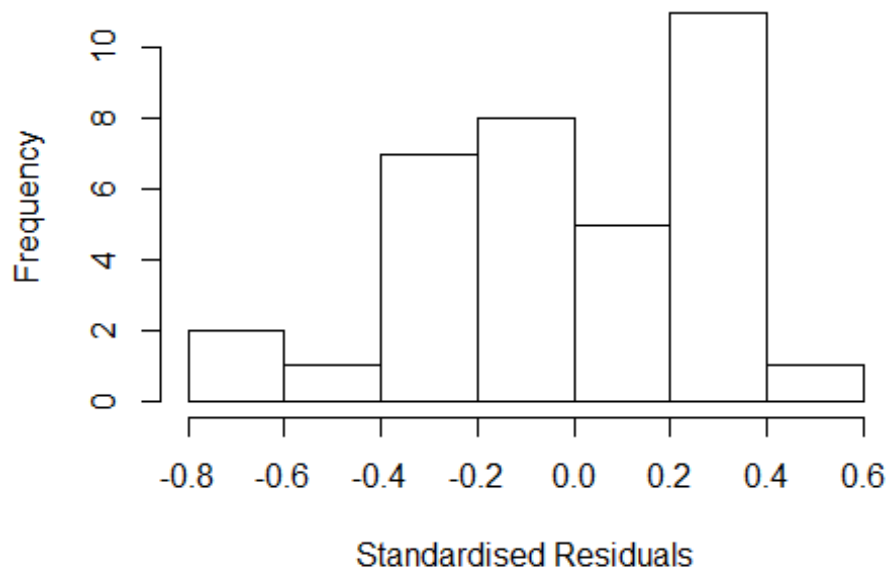
## Year2007          0.8782      0.4801      1.83      0.071 .
## Year2008          0.6840      0.5858      1.17      0.247
## Year2009          0.4463      0.4543      0.98      0.329
## Year2010          0.6856      0.4591      1.49      0.140
## Year2011          0.6945      0.4788      1.45      0.151
## Year2012          0.7154      0.4498      1.59      0.116
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.365
## Multiple R-squared:  0.193, Adjusted R-squared:  0.00785
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 83 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.263  0.888  0.962  0.902  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      1.09e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 92"
## [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 2006 2007 2008 2009 2010 2011
##    2    2    2    1    4    3    2    7    2    1    3    6    3    3    3
## 2012
##    2
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2006 2007 2008 2009 2010 2011
##    0    1    2    1    1    2    2    7    2    0    2    6    2    2    3
## 2012

```



```
##      2
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2006 2007 2008 2009 2010 2011
##    0    1    2    1    1    2    2    7    2    0    2    6    2    2    3
## 2012
##      2
## [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 1.090e+15  1      3.301e+07
## LastAuthorFemale  5.797e+00  1      2.408e+00
## Year              5.183e+15 13      4.022e+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
## -7.46e-01 -2.30e-01  4.44e-16  2.33e-01  4.63e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3685     0.1558   8.78 4.1e-08 ***
## FirstAuthorFemale1 -0.1535     0.1558  -0.99 0.33680
## LastAuthorFemale1  0.4417     0.1980   2.23 0.03794 *
## Year1998          -0.3240     0.2979  -1.09 0.29030
## Year1999          -0.7525     0.1558  -4.83 0.00012 ***
## Year2000          -0.3395     0.1558  -2.18 0.04212 *
## Year2001          -0.5100     0.2302  -2.22 0.03911 *
## Year2002          -0.4470     0.2314  -1.93 0.06841 .
## Year2003          -0.4265     0.1511  -2.82 0.01088 *
## Year2004          -0.8526     0.1652  -5.16 5.5e-05 ***
## Year2007          -0.5944     0.2264  -2.63 0.01667 *
## Year2008          -0.7075     0.2091  -3.38 0.00312 **
## Year2009          -0.7984     0.3278  -2.44 0.02489 *
## Year2010          -0.2795     0.2347  -1.19 0.24837
## Year2011          -0.6224     0.3273  -1.90 0.07250 .
## Year2012          -0.2698     0.0779  -3.46 0.00261 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.456
## Multiple R-squared:  0.309, Adjusted R-squared:  -0.237
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  5 weights are ~= 1. The remaining 30 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.771  0.954  0.972   0.957   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      2.86e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500           50          2          1          1000      200
##      trace.lev      mts      compute.rd
##      0           1000          0
##      psi          subsampling          cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)

```

```
## [1] "Regression 3: First author gender, Year as factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 9.161e+13 1      9.571e+06
## Year              9.161e+13 13      3.443e+00

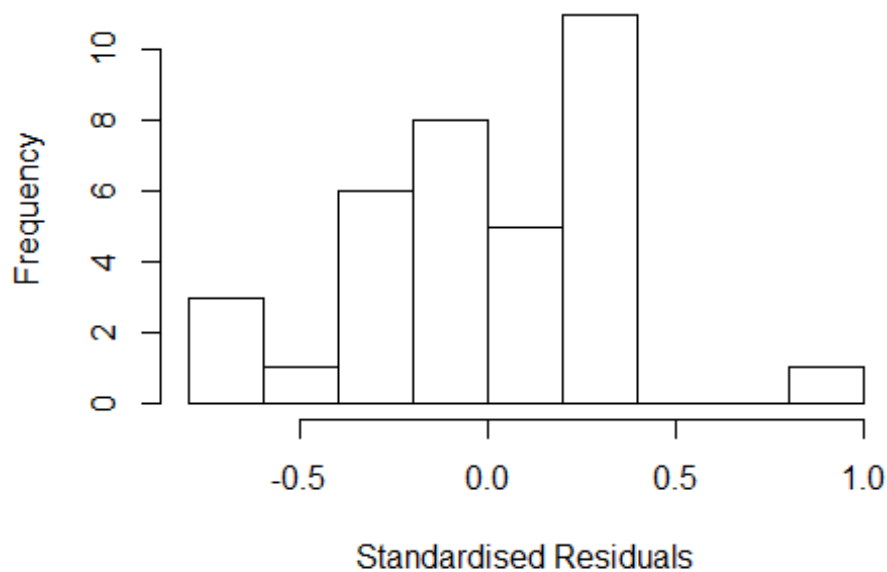
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q    Median        3Q        Max
## -7.07e-01 -2.29e-01  3.29e-15  2.30e-01  8.88e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2903     0.1586   8.14 8.9e-08 ***
## FirstAuthorFemale1 -0.0753     0.1586  -0.47 0.63996
## Year1998          -0.2458     0.2990  -0.82 0.42076
## Year1999          -0.6743     0.1586  -4.25 0.00039 ***
## Year2000          -0.2613     0.1586  -1.65 0.11496
## Year2001          -0.4318     0.2314  -1.87 0.07682 .
## Year2002          -0.3688     0.2326  -1.59 0.12855
## Year2003          -0.3266     0.1207  -2.71 0.01362 *
## Year2004          -0.5927     0.2254  -2.63 0.01607 *
## Year2007          -0.2953     0.2864  -1.03 0.31479
## Year2008          -0.6123     0.2501  -2.45 0.02370 *
## Year2009          -0.4993     0.1730  -2.89 0.00912 **
## Year2010          -0.2013     0.2359  -0.85 0.40354
## Year2011          -0.5831     0.3065  -1.90 0.07157 .
## Year2012          -0.2307     0.0836  -2.76 0.01208 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.48
## Multiple R-squared: 0.197, Adjusted R-squared: -0.365
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~ 1. The remaining 28 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.713  0.958  0.976  0.950  0.981  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          2.86e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##          500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##          0          1000          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
```

Residuals from first author



```
##          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
## -6.81e-01 -2.29e-01  1.11e-15  2.33e-01  5.48e-01
##
## Coefficients:
##              Estimate Std. Error   t value Pr(>|t|)
## (Intercept)    1.22e+00   0.00e+00      Inf    <2e-16 ***
## LastAuthorFemale1 4.05e-01   2.08e-01   1.95e+00   0.0657 .
## Year1998        -1.70e-01   2.57e-01  -6.60e-01   0.5145
## Year1999        -5.99e-01   3.90e-09  -1.53e+08   <2e-16 ***
## Year2000        -1.86e-01   0.00e+00    -Inf    <2e-16 ***
## Year2001        -3.56e-01   1.70e-01  -2.10e+00   0.0484 *
## Year2002        -2.93e-01   1.71e-01  -1.71e+00   0.1020
## Year2003        -3.51e-01   1.02e-01  -3.44e+00   0.0026 **
## Year2004        -7.57e-01   1.08e-01  -7.04e+00   8e-07 ***
## Year2007        -4.22e-01   1.30e-01  -3.26e+00   0.0039 **
## Year2008        -6.02e-01   2.23e-01  -2.70e+00   0.0137 *
## Year2009        -6.26e-01   2.55e-01  -2.46e+00   0.0233 *
## Year2010        -1.26e-01   1.76e-01  -7.20e-01   0.4818
## Year2011        -5.34e-01   3.23e-01  -1.65e+00   0.1136
## Year2012        -1.93e-01   5.33e-02  -3.62e+00   0.0017 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.451
## Multiple R-squared:  0.286, Adjusted R-squared:  -0.214
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 27 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.803  0.945  0.968  0.950  0.977  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      2.86e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample max.it best.r.s k.fast.s k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev mts compute.rd
##      0          1000          0
##      psi          subsampling          cov
```

```

##          "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##          "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 35"
## [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
##    6    5    5    5    6    7    6    8    9    7    10   11   11   13   14
## 2011 2012
##   15    8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    0    3    0    3    2    2    3    1    6    8    8    6    7
## 2011 2012
##    4    2
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    0    3    0    3    0    2    3    1    6    8    7    5    6
## 2011 2012
##    4    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: 2.6207413942089"

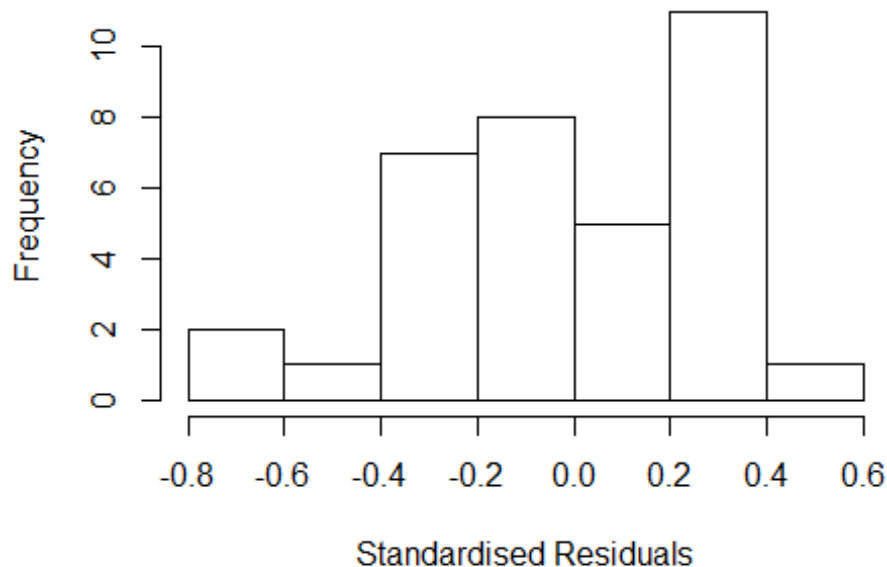
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## 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: 2.44948974278318"

## 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 = 0.6
## 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 1.077e+15  1      3.282e+07
## LastAuthorFemale  2.256e+15  1      4.750e+07
## UniqueAuthors    9.934e+43  4      3.160e+05
## Year              3.193e+43 13      4.712e+01

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
```

```

##           Min           1Q           Median           3Q           Max
## -9.35e-01 -1.83e-01  4.35e-16  1.91e-01  5.39e-01
##
## Coefficients:
##               Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.07e+00   1.93e-09  5.56e+08 < 2e-16 ***
## FirstAuthorFemale1 6.99e-02   2.44e-01  2.90e-01  0.77601
## LastAuthorFemale1 -3.27e-01   3.77e-01 -8.70e-01  0.39243
## UniqueAuthors2    1.26e-01   1.96e-01  6.40e-01  0.52649
## UniqueAuthors3    1.19e-01   3.35e-01  3.60e-01  0.72400
## UniqueAuthors4    4.25e-01   3.64e-01  1.17e+00  0.25160
## UniqueAuthors5   -2.46e-01   4.08e-01 -6.00e-01  0.55141
## Year1997          -3.00e-01   3.19e-01 -9.40e-01  0.35384
## Year1999          -3.53e-01   1.68e-01 -2.10e+00  0.04353 *
## Year2001          -1.11e-01   1.64e-01 -6.80e-01  0.50361
## Year2003          -2.85e-01   4.08e-01 -7.00e-01  0.48997
## Year2004           2.70e-01   1.08e-01  2.50e+00  0.01781 *
## Year2005          -2.65e-01   1.96e-01 -1.35e+00  0.18674
## Year2006          -2.59e-01   3.44e-01 -7.50e-01  0.45681
## Year2007          -3.40e-01   2.79e-01 -1.22e+00  0.23284
## Year2008          -2.74e-01   1.70e-01 -1.61e+00  0.11699
## Year2009          -4.48e-01   1.02e-01 -4.39e+00  0.00012 ***
## Year2010           1.16e-02   2.13e-01  5.00e-02  0.95691
## Year2011          -5.29e-02   3.86e-01 -1.40e-01  0.89176
## Year2012          -3.91e-02   2.50e-01 -1.60e-01  0.87664
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.359
## Multiple R-squared:  0.33, Adjusted R-squared:  -0.0671
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 45 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.477  0.886  0.970  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           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

```

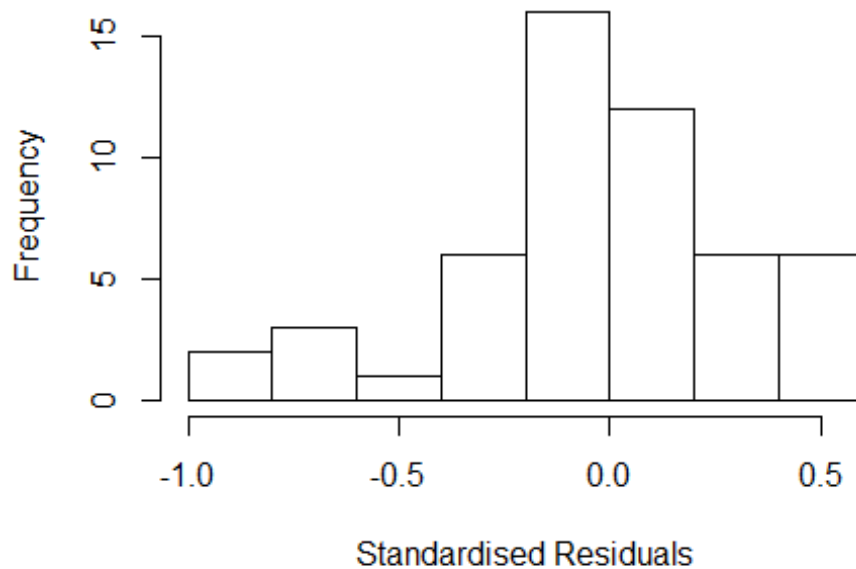


```
##                                "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"

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

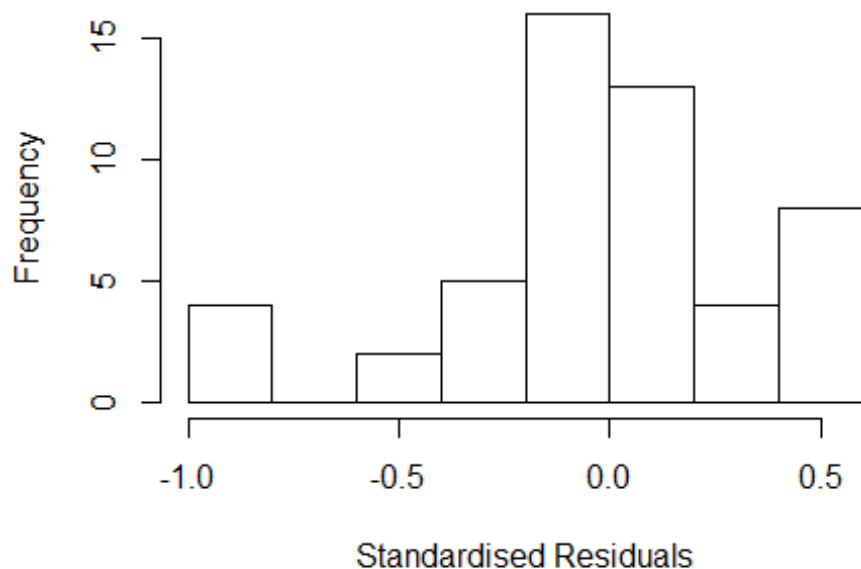
## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

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 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
## -9.03e-01 -1.83e-01  6.66e-16  1.88e-01  6.00e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0750     0.0000      Inf < 2e-16 ***
## FirstAuthorFemale1 -0.0186     0.1508    -0.12  0.90248
## LastAuthorFemale1  -0.2229     0.2577    -0.87  0.39267
## Year1997          -0.3155     0.3127    -1.01  0.31983
## Year1999          -0.2749     0.1119    -2.46  0.01897 *
## Year2001          -0.0696     0.1699    -0.41  0.68432
## Year2003          -0.2927     0.3786    -0.77  0.44442
## Year2004           0.3799     0.0876     4.34  0.00011 ***
## Year2005          -0.1390     0.0000    -Inf < 2e-16 ***
## Year2006          -0.1716     0.2389    -0.72  0.47708
## Year2007          -0.2467     0.2016    -1.22  0.22906
## Year2008          -0.2156     0.1767    -1.22  0.23027
```

```

## Year2009          -0.3860      0.1118    -3.45  0.00144 **
## Year2010           0.1232      0.0935     1.32  0.19578
## Year2011           0.0629      0.3182     0.20  0.84442
## Year2012           0.0648      0.1532     0.42  0.67483
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.392
## Multiple R-squared:  0.286, Adjusted R-squared:  -0.0113
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 46 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.575  0.905  0.976  0.922  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      1.92e-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 1.904e+14  1      1.380e+07
## Year              1.904e+14 13      3.542e+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
## -1.12e+00 -1.67e-01 -4.58e-16  1.90e-01  6.21e-01
##
## Coefficients:

```

```

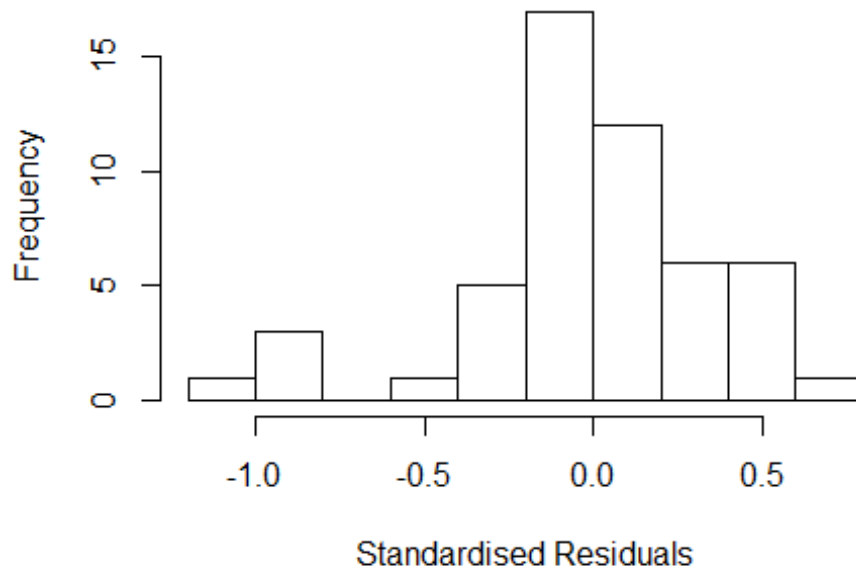
##               Estimate Std. Error   t value Pr(>|t|)
## (Intercept)      1.07e+00   2.05e-08  5.23e+07  <2e-16 ***
## FirstAuthorFemale1 -6.52e-02   2.04e-01 -3.20e-01   0.7514
## Year1997          -4.92e-01   2.04e-01 -2.41e+00   0.0212 *
## Year1999          -3.33e-01   9.98e-02 -3.34e+00   0.0019 **
## Year2001          -6.90e-02   1.71e-01 -4.00e-01   0.6898
## Year2003          -3.81e-01   2.61e-01 -1.46e+00   0.1528
## Year2004           3.95e-01   1.15e-01  3.44e+00   0.0014 **
## Year2005          -1.39e-01   2.10e-08 -6.60e+06  <2e-16 ***
## Year2006          -1.65e-01   2.50e-01 -6.60e-01   0.5127
## Year2007          -2.42e-01   2.10e-01 -1.15e+00   0.2561
## Year2008          -2.37e-01   1.80e-01 -1.32e+00   0.1953
## Year2009          -3.80e-01   1.26e-01 -3.03e+00   0.0045 **
## Year2010           9.25e-02   7.28e-02  1.27e+00   0.2119
## Year2011           1.11e-01   3.70e-01  3.00e-01   0.7654
## Year2012           8.81e-02   1.54e-01  5.70e-01   0.5711
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.374
## Multiple R-squared:  0.291, Adjusted R-squared:  0.0222
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 47 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.349  0.900  0.976  0.914  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           1.92e-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

```

Residuals from first author



```
##          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
## -9.14e-01 -1.86e-01 -5.41e-16  1.89e-01  6.04e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0750     0.0000      Inf < 2e-16 ***
## LastAuthorFemale1 -0.2287     0.2893    -0.79  0.4342
## Year1997          -0.3283     0.2893    -1.14  0.2637
## Year1999          -0.2792     0.1051    -2.66  0.0116 *
## Year2001          -0.0694     0.1703    -0.41  0.6862
## Year2003          -0.2992     0.3854    -0.78  0.4425
## Year2004           0.3739     0.0738     5.07  1.2e-05 ***
## Year2005          -0.1390     0.0000    -Inf < 2e-16 ***
```

```

## Year2006          -0.1690      0.2402    -0.70    0.4859
## Year2007          -0.2450      0.2032    -1.21    0.2356
## Year2008          -0.2202      0.1552    -1.42    0.1641
## Year2009          -0.3891      0.1146    -3.39    0.0017 **
## Year2010           0.1213      0.0941     1.29    0.2051
## Year2011           0.0681      0.3269     0.21    0.8362
## Year2012           0.0555      0.1411     0.39    0.6963
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.385
## Multiple R-squared:  0.288, Adjusted R-squared:  0.019
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 46 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.551  0.903   0.974   0.919   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      1.92e-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 1913"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1998 1999 2000 2003 2004 2005 2008 2010 2012
##    2    1    1    2    1    1    2    2    1
##
## 1998 1999 2000 2003 2004 2005 2008 2010 2012
##    1    0    1    2    1    1    1    2    0
##
## 1998 1999 2000 2003 2004 2005 2008 2010 2012

```

```

##      1      0      1      2      0      1      1      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: 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: 8"
## [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]"
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##      2      3      2      3      1      1      4      3      6      6      8      6      13      13      3
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##      2      2      2      3      1      0      3      2      6      6      7      5      11      12      3
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##      2      2      2      3      1      0      3      2      4      6      7      5      10      11      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: 1.62238960361098"

## 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 = 0.7
## 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: 1.66055153809915"

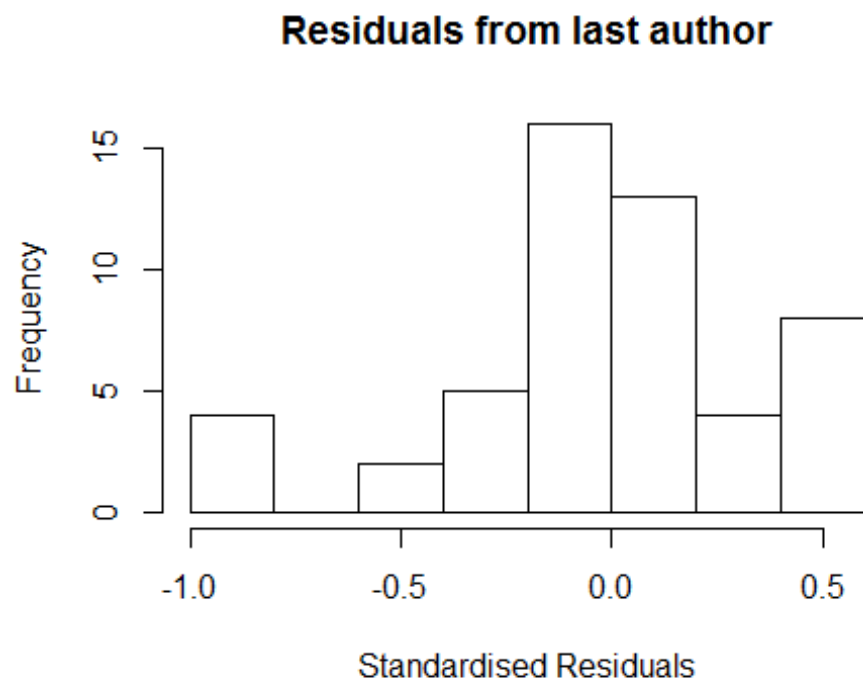
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction

```

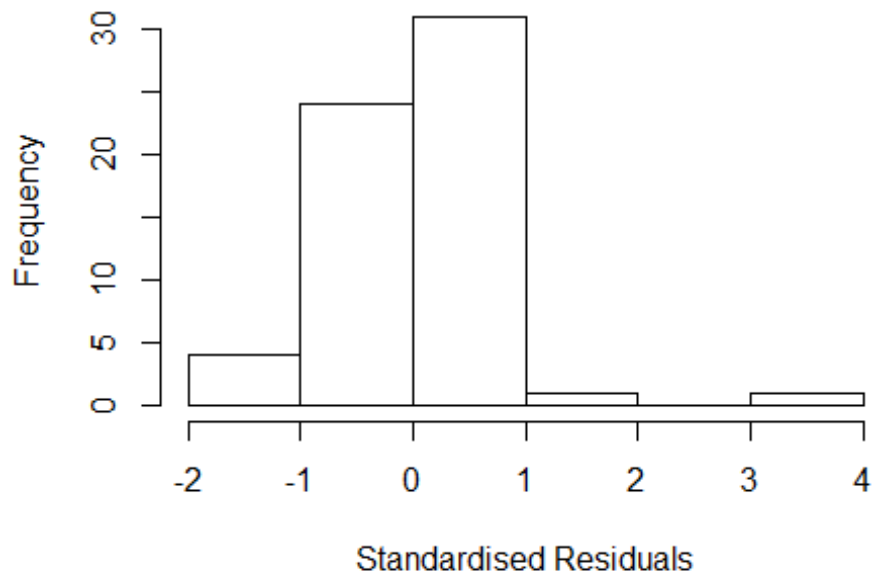
```
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5.5, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"

## Warning in outlierStats(ret, x, control): Detected possible local
breakdown of SM-estimate in coefficient 'Year2000'.
## Use lmrob argument 'setting="KS2014"' to avoid this problem.
```



```
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -4.138e+25  1      NaN
## LastAuthorFemale  1.700e+01  1      4.123
## UniqueAuthors    1.131e+16  3     473.759
## Year              8.414e+15 13      4.097
```


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 3042846612 5.179 2000      2000      1      3.725
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
##      Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##      k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.6449 -0.4221  0.0784  0.4221  3.7250
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.7611    0.1230   14.32 < 2e-16 ***
## FirstAuthorFemale1  0.1482    0.3548    0.42  0.6783
## LastAuthorFemale1  0.0157    0.4428    0.04  0.9719
## UniqueAuthors2    0.5842    0.2045    2.86  0.0066 **
## UniqueAuthors3    1.2583    0.5998    2.10  0.0420 *
## UniqueAuthors4    0.8379    0.3516    2.38  0.0218 *
## Year1999         -1.0598    0.6999   -1.51  0.1375
## Year2000          -0.3071    0.1230   -2.50  0.0166 *
## Year2001          -0.9539    0.4257   -2.24  0.0304 *
## Year2002          -0.2761    0.1230   -2.24  0.0302 *
```

```

## Year2004          -1.6195      0.5014   -3.23   0.0024 **
## Year2005          -2.0714      0.4068   -5.09   7.9e-06 ***
## Year2006          -0.8185      0.2690   -3.04   0.0040 **
## Year2007          -0.8547      0.2796   -3.06   0.0039 **
## Year2008          -0.4023      0.2237   -1.80   0.0793 .
## Year2009          -0.6416      0.4732   -1.36   0.1824
## Year2010          -0.6204      0.2370   -2.62   0.0122 *
## Year2011          -0.8651      0.2699   -3.21   0.0026 **
## Year2012          -0.4978      0.3742   -1.33   0.1906
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.729
## Multiple R-squared:  0.338, Adjusted R-squared:  0.0544
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## observation 6 is an outlier with |weight| = 0 ( < 0.0016);
## 11 weights are ~= 1. The remaining 49 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.590  0.911  0.951   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           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 2: First author gender, Last author gender, Year as
factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 10.33 1           3.214
## LastAuthorFemale  10.96 1           3.311
## Year              23.53 13           1.129

## [1] "List of 1 outliers with residuals above 2.5"
##   ScopusId  NLCS Year OneField Fields residuals
## 7 3042846612 5.179 2000      2000      1      3.586
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =

```

```

AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q    Median        3Q        Max
## -1.2589 -0.5955  0.0504   0.5422   3.5863
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.7911      0.1265   14.16  <2e-16 ***
## FirstAuthorFemale1  0.1925      0.2786    0.69   0.4932
## LastAuthorFemale1 -0.0888      0.3202   -0.28   0.7829
## Year1999          -0.5279      0.6011   -0.88   0.3845
## Year2000          -0.1984      0.9474   -0.21   0.8351
## Year2001          -0.5834      0.7087   -0.82   0.4147
## Year2002          -0.3061      0.1265   -2.42   0.0196 *
## Year2004          -1.0380      0.3683   -2.82   0.0072 **
## Year2005          -1.2024      0.6781   -1.77   0.0830 .
## Year2006          -0.8485      0.2684   -3.16   0.0028 **
## Year2007          -0.6363      0.2563   -2.48   0.0168 *
## Year2008          -0.1467      0.2499   -0.59   0.5602
## Year2009          -0.2077      0.5475   -0.38   0.7063
## Year2010          -0.5322      0.2676   -1.99   0.0529 .
## Year2011          -0.4962      0.1876   -2.65   0.0112 *
## Year2012          -0.1276      0.4848   -0.26   0.7935
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.854
## Multiple R-squared:  0.138, Adjusted R-squared:  -0.149
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 53 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
## 0.0386 0.9060 0.9480 0.9220 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          1.64e-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 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: 61"
## [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]"
##
## 1998 2000 2003 2004 2005 2006 2007 2008 2009 2011 2012
##    1    1    1    3    3    4    3    5    4    4    8
##
## 1998 2000 2003 2004 2005 2006 2007 2008 2009 2011 2012
##    1    1    1    2    3    2    2    4    4    4    6
##
## 1998 2000 2003 2004 2005 2006 2007 2008 2009 2011 2012
##    1    1    1    2    3    2    2    4    3    4    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: 2.1689435423954"

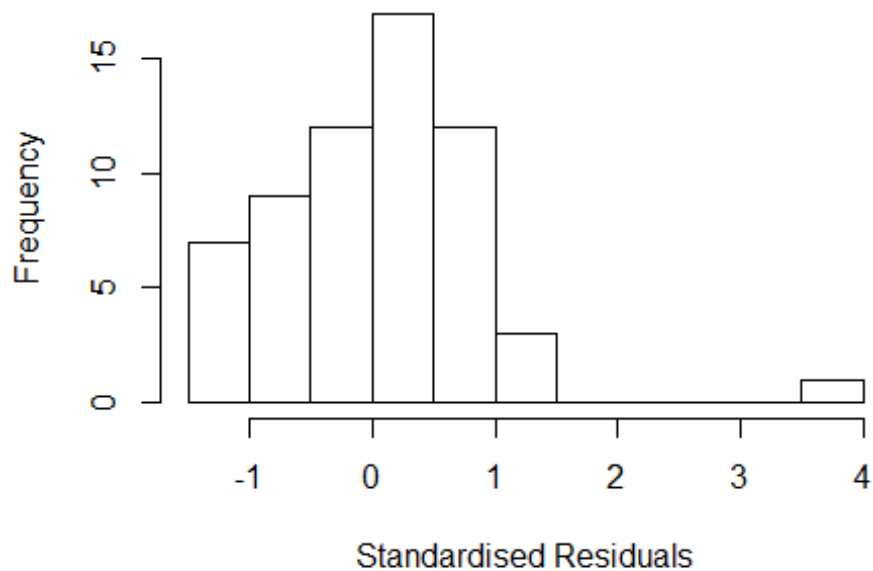
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0, p-value = 0.05
## 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: 2.21336383940064"

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

```

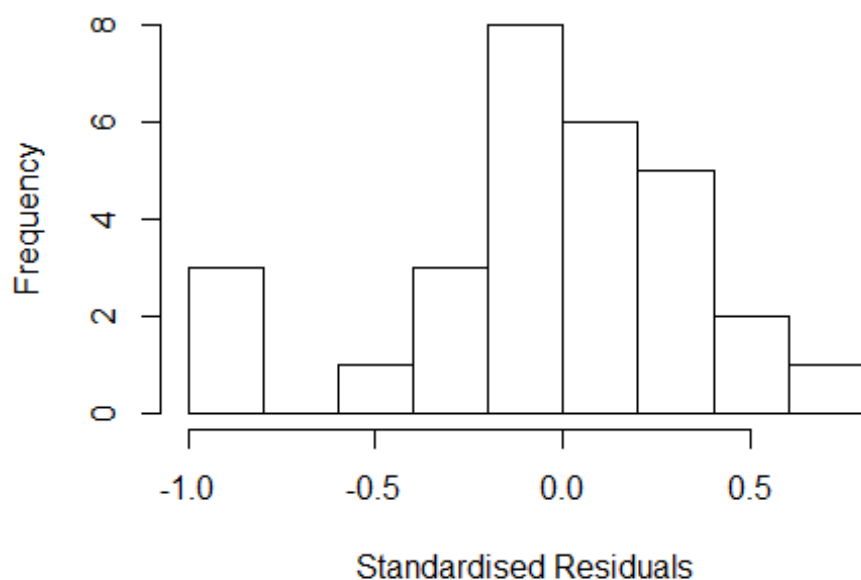
Residuals from first and last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1.5, 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"
## [1] "Regression 2: First author gender, Last author gender, Year as
##      factors"
##
```

| | GVIF | Df | GVIF ^{1/(2*Df)} |
|----------------------|-----------|----|--------------------------|
| ## FirstAuthorFemale | 1.980e+13 | 1 | 4.450e+06 |
| ## LastAuthorFemale | 4.036e+13 | 1 | 6.353e+06 |
| ## Year | 4.188e+26 | 10 | 2.143e+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
## -8.58e-01 -1.97e-01  1.11e-15  2.66e-01  7.07e-01
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.24e+00   1.51e-08  8.18e+07  < 2e-16 ***
## FirstAuthorFemale1  5.13e-02   1.09e-01  4.70e-01  0.64537
## LastAuthorFemale1  5.35e-01   7.40e-02  7.23e+00  2.0e-06 ***
## Year2000         -2.05e-01   7.40e-02 -2.77e+00  0.01358 *
## Year2003          2.07e-02   1.09e-01  1.90e-01  0.85207
## Year2004         -7.31e-01   1.53e-01 -4.78e+00  0.00021 ***
## Year2005          3.38e-01   5.41e-02  6.25e+00  1.2e-05 ***
## Year2006         -4.21e-01   2.23e-01 -1.89e+00  0.07698 .
## Year2007         -7.00e-03   4.16e-01 -2.00e-02  0.98678
## Year2008         -3.78e-01   3.64e-01 -1.04e+00  0.31359
## Year2009         -4.85e-01   4.30e-01 -1.13e+00  0.27565
## Year2011         -4.06e-01   2.56e-01 -1.59e+00  0.13172
```

```

## Year2012          9.23e-03   1.12e-01   8.00e-02   0.93535
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.519
## Multiple R-squared:  0.464, Adjusted R-squared:  0.0624
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 25 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.766  0.921  0.976  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           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 -5.293e+15  1           NaN
## Year              -5.293e+15 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
## -0.852 -0.241  0.000  0.296  0.713
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   1.24e+00   2.33e-08  5.31e+07 < 2e-16 ***
## FirstAuthorFemale1 1.19e-01   2.02e-01  5.90e-01  0.56266

```

```

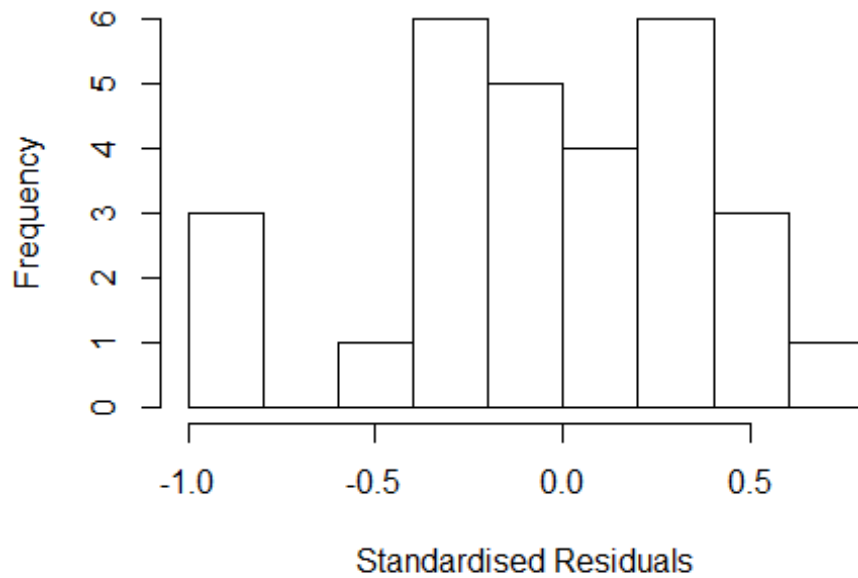
## Year2000          3.30e-01  1.99e-08  1.66e+07  < 2e-16 ***
## Year2003          -4.72e-02  2.02e-01 -2.30e-01  0.81795
## Year2004          -7.65e-01  1.54e-01 -4.96e+00  0.00012 ***
## Year2005           4.92e-01  1.55e-01  3.18e+00  0.00552 **
## Year2006          -4.21e-01  2.19e-01 -1.92e+00  0.07198 .
## Year2007          -7.00e-03  3.97e-01 -2.00e-02  0.98613
## Year2008          -3.84e-01  3.42e-01 -1.12e+00  0.27740
## Year2009          -3.17e-02  4.41e-01 -7.00e-02  0.94362
## Year2011          -4.17e-01  2.49e-01 -1.67e+00  0.11266
## Year2012           7.56e-02  1.17e-01  6.50e-01  0.52546
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.592
## Multiple R-squared:  0.403, Adjusted R-squared:  0.0167
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 3 weights are ~= 1. The remaining 26 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.820  0.941  0.977  0.954  0.986  0.999
## Algorithmic parameters:
##           tuning.chi                bb           tuning.psi           refine.tol
##           1.55e+00                5.00e-01           4.69e+00           1.00e-07
##           rel.tol                solve.tol           eps.outlier           eps.x
##           1.00e-07                1.00e-07           3.45e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01                5.00e-01
## nResample      max.it    best.r.s    k.fast.s      k.max maxit.scale
##           500           50         2         1         1000         200
## trace.lev      mts    compute.rd
##           0           1000         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

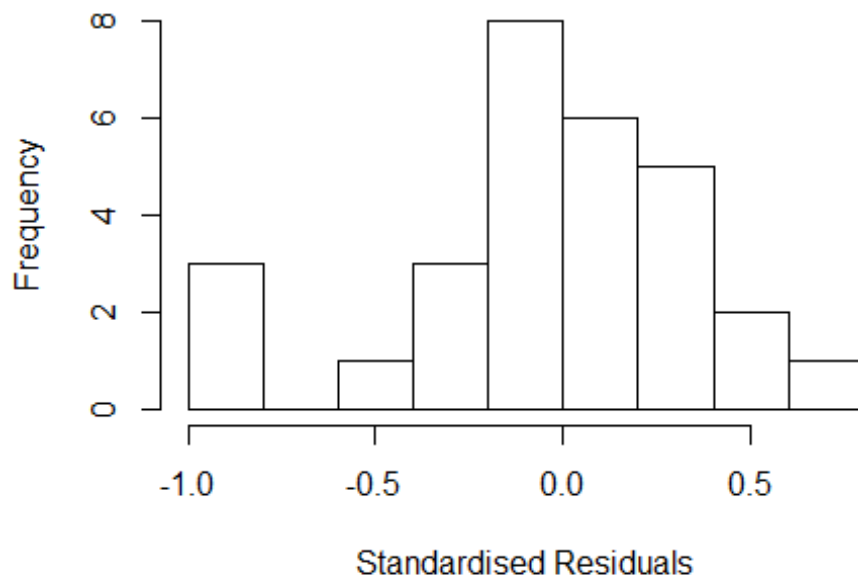
```


Residuals from first author



| | | | |
|---------------------|------|----|---------------------|
| ## | GVIF | Df | $GVIF^{(1/(2*Df))}$ |
| ## LastAuthorFemale | NaN | 1 | NaN |
| ## Year | NaN | 10 | NaN |

Residuals from last author



```

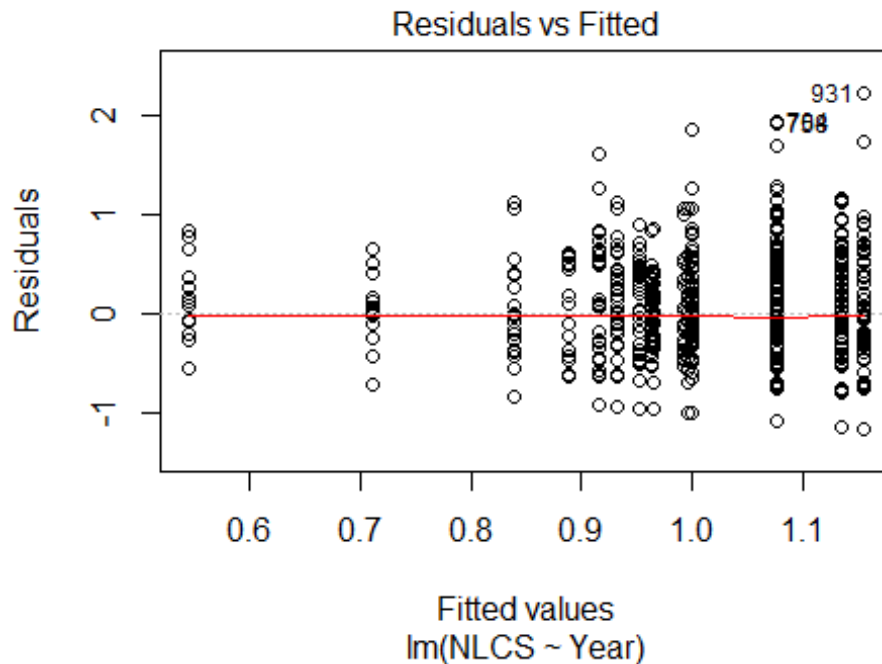
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8623 -0.1995 0.0185 0.2576 0.7027
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.24e+00 2.32e-08 5.32e+07 < 2e-16 ***
## LastAuthorFemale1 5.46e-01 8.34e-02 6.55e+00 5.0e-06 ***
## Year2000 -2.16e-01 8.34e-02 -2.59e+00 0.01890 *
## Year2003 7.20e-02 0.00e+00 Inf < 2e-16 ***
## Year2004 -7.06e-01 1.64e-01 -4.31e+00 0.00048 ***
## Year2005 3.52e-01 3.91e-02 8.99e+00 7.2e-08 ***
## Year2006 -4.21e-01 2.25e-01 -1.87e+00 0.07888 .
## Year2007 -7.00e-03 4.32e-01 -2.00e-02 0.98725
## Year2008 -3.74e-01 3.79e-01 -9.90e-01 0.33799
## Year2009 -4.34e-01 4.36e-01 -9.90e-01 0.33379
## Year2011 -3.98e-01 2.57e-01 -1.55e+00 0.14001
## Year2012 2.45e-02 9.77e-02 2.50e-01 0.80516
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.48
## Multiple R-squared: 0.468, Adjusted R-squared: 0.124
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 25 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.727 0.908 0.974 0.935 0.986 0.999
## Algorithmic parameters:
## tuning.chi bb tuning.psi refine.tol
## 1.55e+00 5.00e-01 4.69e+00 1.00e-07
## rel.tol solve.tol eps.outlier eps.x
## 1.00e-07 1.00e-07 3.45e-03 1.82e-12
## warn.limit.reject warn.limit.meanrw
## 5.00e-01 5.00e-01
## nResample max.it best.r.s k.fast.s k.max maxit.scale
## 500 50 2 1 1000 200
## trace.lev mts compute.rd
## 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 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
##    22   31   22   29   33   27   23   24   46   36   43   46   63   58   81
## 2011 2012
##    85   78
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    14   24   19   24   24   21   19   22   37   34   39   43   54   52   74
## 2011 2012
##    75   71
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    14   23   19   22   23   21   19   21   36   32   36   38   49   47   67
## 2011 2012
##    68   61
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 51, df = 16, p-value = 1e-05

```



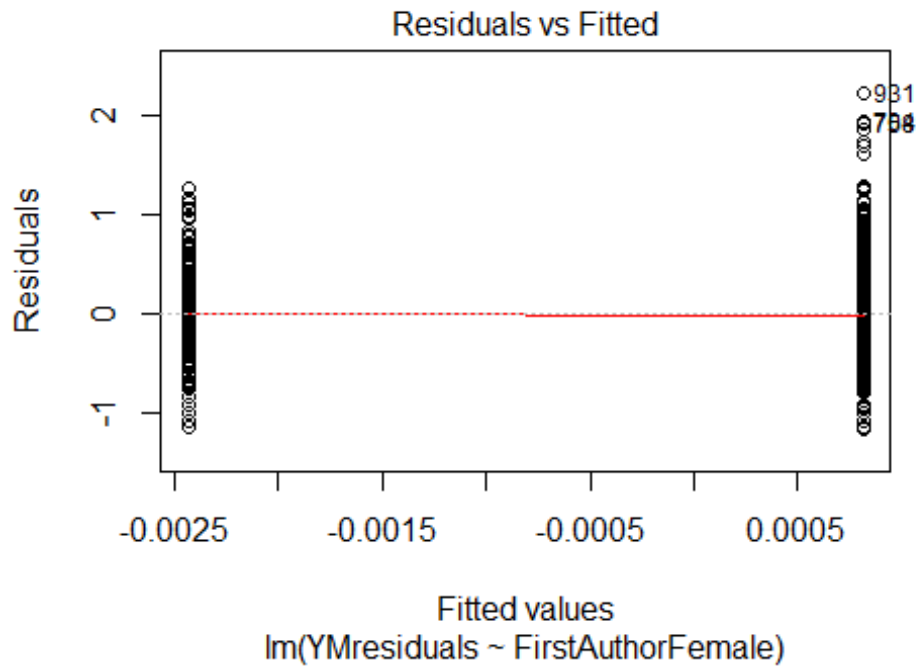
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.85, df = 1, p-value = 0.4

## [1] "Female first author team size 2018 geometric mean: 1.76865991418384"
## [1] "Male first author team size 2018 geometric mean: 1.70495309703477"

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

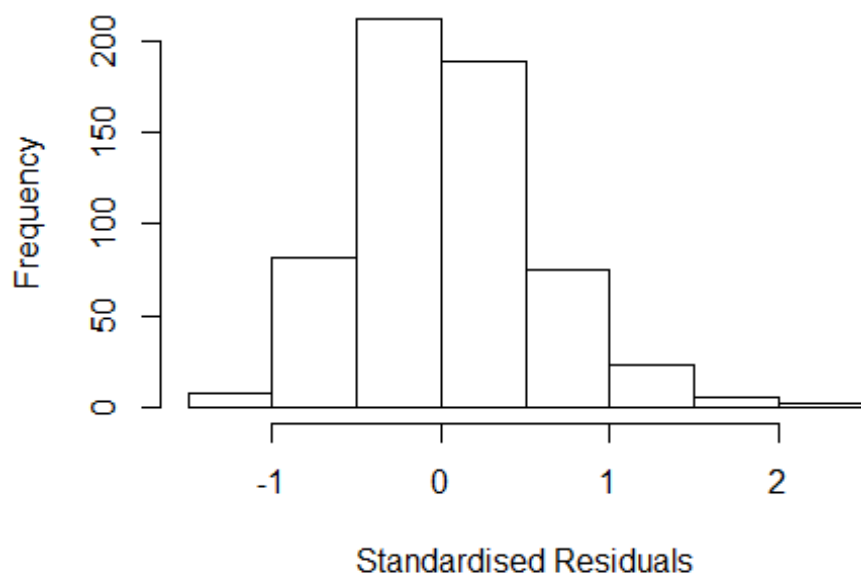
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 300, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.63057279084246"
## [1] "Male last author team size 2018 geometric mean: 1.75260319068767"

## Warning in wilcox.test.default(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 = 0.6
## 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 2.037 1          1.427
## LastAuthorFemale  1.927 1          1.388
## UniqueAuthors    2.057 4          1.094
## Year              2.265 16         1.026
```

Residuals from first and last author and team size



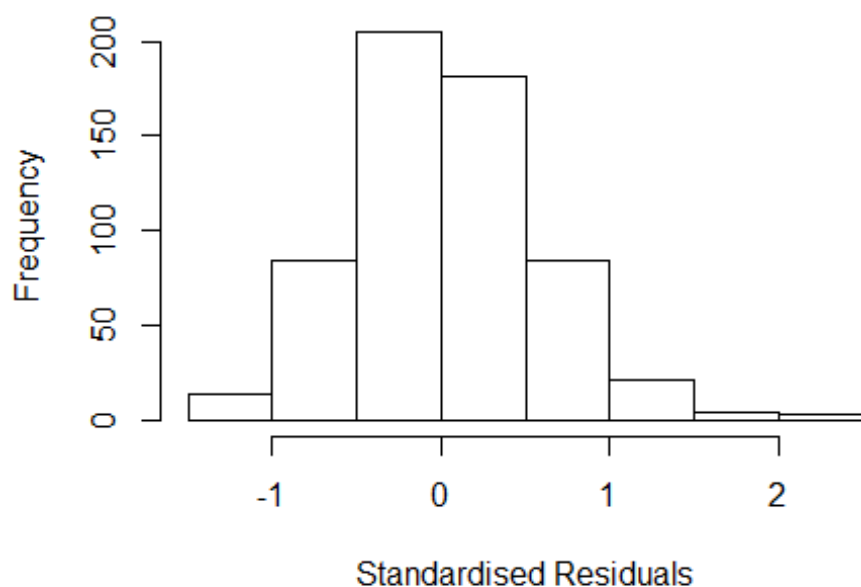
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.23829 -0.34382 -0.00671 0.37707 2.33408
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6914 0.1093 6.33 5e-10 ***
## FirstAuthorFemale1 0.0421 0.0713 0.59 0.5552
## LastAuthorFemale1 -0.0130 0.0687 -0.19 0.8503
## UniqueAuthors2 0.1741 0.0545 3.19 0.0015 **
## UniqueAuthors3 0.2054 0.0769 2.67 0.0078 **
## UniqueAuthors4 0.2026 0.1201 1.69 0.0923 .
## UniqueAuthors5 0.5037 0.2316 2.18 0.0300 *
## Year1997 0.0339 0.1349 0.25 0.8016
## Year1998 -0.2162 0.1518 -1.42 0.1550
## Year1999 0.1715 0.1570 1.09 0.2753
```

```

## Year2000          0.2622      0.1372      1.91      0.0565 .
## Year2001          0.2068      0.1448      1.43      0.1538
## Year2002          0.0991      0.1594      0.62      0.5342
## Year2003          0.1646      0.1469      1.12      0.2631
## Year2004          0.2005      0.1250      1.60      0.1094
## Year2005          0.2107      0.1494      1.41      0.1588
## Year2006          0.0553      0.1482      0.37      0.7094
## Year2007          0.1588      0.1463      1.09      0.2781
## Year2008          0.2622      0.1576      1.66      0.0967 .
## Year2009          0.1514      0.1502      1.01      0.3141
## Year2010          0.2129      0.1399      1.52      0.1285
## Year2011          0.2963      0.1398      2.12      0.0345 *
## Year2012          0.3415      0.1408      2.43      0.0156 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.522
## Multiple R-squared:  0.0837, Adjusted R-squared:  0.0486
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 62 weights are ~= 1. The remaining 534 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0082 0.8660 0.9480 0.8980 0.9840 0.9990
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          1.68e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.908 1          1.381
## LastAuthorFemale 1.788 1          1.337
## Year          1.212 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
## -1.1622 -0.3455 -0.0173 0.3850 2.2371
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7103 0.1003 7.08 4.1e-12 ***
## FirstAuthorFemale1 0.0490 0.0708 0.69 0.4895
## LastAuthorFemale1 -0.0216 0.0681 -0.32 0.7513
## Year1997 0.0642 0.1280 0.50 0.6160
## Year1998 -0.1821 0.1505 -1.21 0.2265
## Year1999 0.2479 0.1415 1.75 0.0803 .
## Year2000 0.2729 0.1286 2.12 0.0343 *
## Year2001 0.2953 0.1306 2.26 0.0241 *
## Year2002 0.1755 0.1545 1.14 0.2563
## Year2003 0.2184 0.1388 1.57 0.1163
## Year2004 0.2370 0.1135 2.09 0.0371 *
## Year2005 0.2566 0.1468 1.75 0.0811 .
```

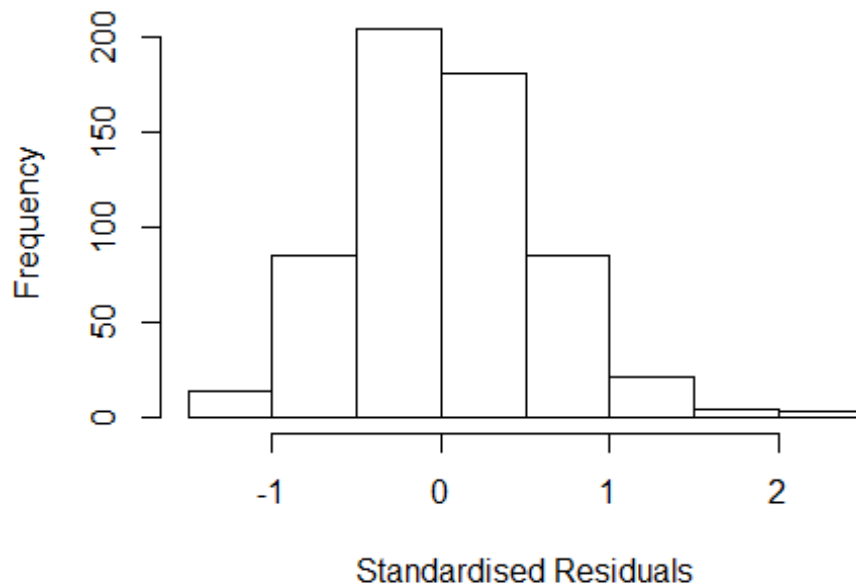


```

## Year2006          0.0980      0.1464      0.67      0.5032
## Year2007          0.2247      0.1355      1.66      0.0978 .
## Year2008          0.3424      0.1497      2.29      0.0225 *
## Year2009          0.2313      0.1384      1.67      0.0952 .
## Year2010          0.2669      0.1314      2.03      0.0427 *
## Year2011          0.4029      0.1293      3.12      0.0019 **
## Year2012          0.4196      0.1296      3.24      0.0013 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.529
## Multiple R-squared:  0.058, Adjusted R-squared:  0.0286
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 41 weights are ~= 1. The remaining 555 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0347 0.8720 0.9490 0.9020 0.9860 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      1.68e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.134 1      1.065
## Year              1.134 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
## -1.146 -0.355 -0.018 0.385 2.238
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7094 0.1008 7.04 5.6e-12 ***
## FirstAuthorFemale1 0.0356 0.0547 0.65 0.5161
## Year1997 0.0618 0.1279 0.48 0.6289
## Year1998 -0.1843 0.1505 -1.23 0.2210
## Year1999 0.2457 0.1413 1.74 0.0827 .
## Year2000 0.2726 0.1288 2.12 0.0347 *
## Year2001 0.2952 0.1310 2.25 0.0246 *
## Year2002 0.1737 0.1553 1.12 0.2638
## Year2003 0.2178 0.1394 1.56 0.1188
## Year2004 0.2376 0.1139 2.09 0.0374 *
## Year2005 0.2540 0.1463 1.74 0.0830 .
## Year2006 0.0972 0.1471 0.66 0.5088
```

```

## Year2007          0.2244      0.1359      1.65      0.0993 .
## Year2008          0.3407      0.1499      2.27      0.0234 *
## Year2009          0.2280      0.1379      1.65      0.0988 .
## Year2010          0.2647      0.1316      2.01      0.0448 *
## Year2011          0.4009      0.1294      3.10      0.0020 **
## Year2012          0.4197      0.1301      3.23      0.0013 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.528
## Multiple R-squared:  0.0578, Adjusted R-squared:  0.0301
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 40 weights are ~= 1. The remaining 556 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0329 0.8720 0.9500 0.9020 0.9860 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      1.68e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.06 1      1.030
## Year      1.06 16      1.002

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.1397 -0.3534 -0.0216  0.3805  2.2273

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.71225    0.09915   7.18 2.1e-12 ***
## LastAuthorFemale1 0.00873    0.05237   0.17 0.86758
## Year1997        0.06185    0.12676   0.49 0.62580
## Year1998       -0.18009    0.14874  -1.21 0.22648
## Year1999        0.24897    0.14142   1.76 0.07884 .
## Year2000        0.27557    0.12731   2.16 0.03083 *
## Year2001        0.29912    0.12909   2.32 0.02084 *
## Year2002        0.17138    0.15389   1.11 0.26588
## Year2003        0.22645    0.13870   1.63 0.10307
## Year2004        0.24419    0.11140   2.19 0.02878 *
## Year2005        0.25906    0.14514   1.78 0.07480 .
## Year2006        0.10301    0.14594   0.71 0.48055
## Year2007        0.22763    0.13415   1.70 0.09027 .
## Year2008        0.34546    0.14756   2.34 0.01957 *
## Year2009        0.23152    0.13751   1.68 0.09280 .
## Year2010        0.27262    0.13065   2.09 0.03735 *
## Year2011        0.40471    0.12875   3.14 0.00176 **
## Year2012        0.42747    0.12745   3.35 0.00085 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.535
## Multiple R-squared:  0.0566, Adjusted R-squared:  0.0288
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 39 weights are ~= 1. The remaining 557 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0441 0.8770 0.9510 0.9040 0.9860 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.68e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 596"
## [1] ""

```

```

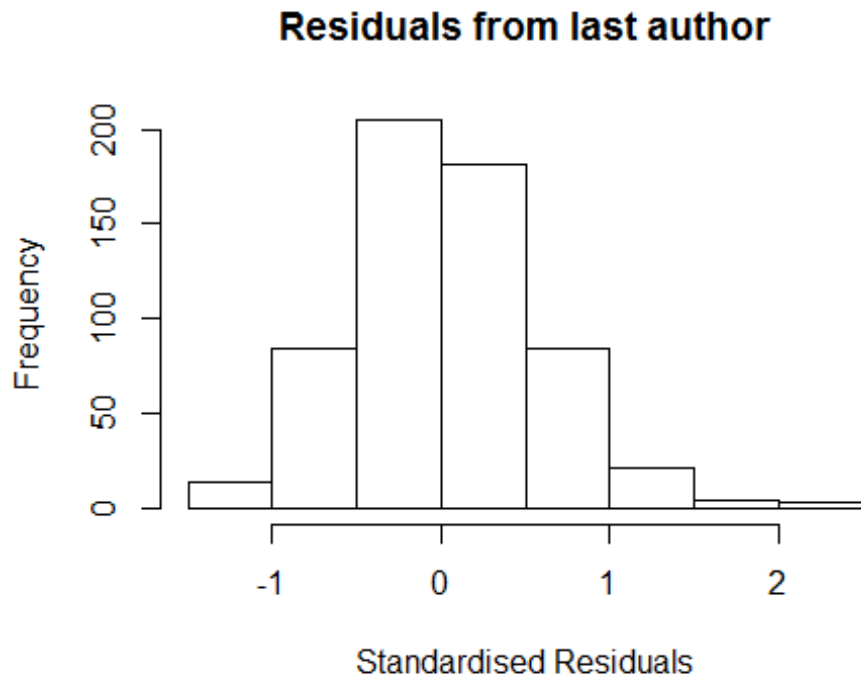
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2003"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    3    4    3    8    7    5    2   14   11   22   12   16   11   15
## 2011 2012
##   12   17
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    3    4    3    6    6    5    1   13   11   21   11   16    9   14
## 2011 2012
##   12   17
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    3    4    3    5    6    5    1   13   10   19   10   16    9   14
## 2011 2012
##   11   13
## [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: 1.782602457966"

## Warning in wilcox.test.default(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 = 0.7
## 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: 1.81712059283214"

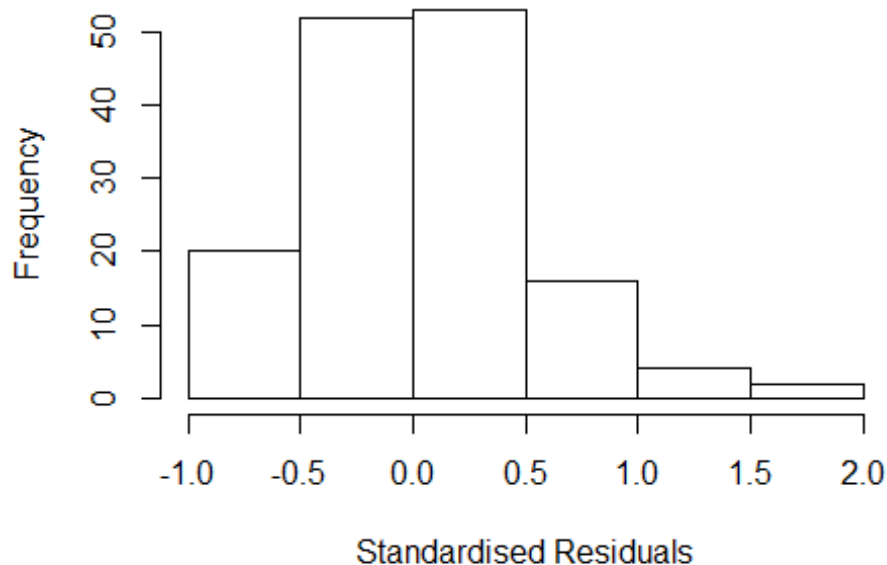
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 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"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 7.061 1      2.657
## LastAuthorFemale  4.764 1      2.183
## UniqueAuthors    3.096 2      1.327
## Year             11.292 16      1.079
```

Residuals from first and last author and team size



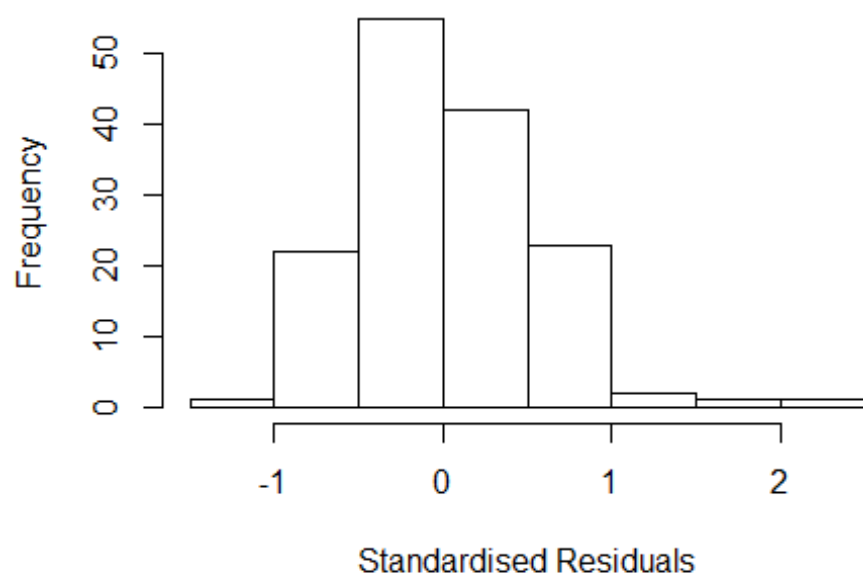
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9752 -0.3201 0.0153 0.3258 1.8418
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7364 0.1696 4.34 2.9e-05 ***
## FirstAuthorFemale1 0.1047 0.1727 0.61 0.5456
## LastAuthorFemale1 0.0106 0.1489 0.07 0.9434
## UniqueAuthors2 0.2971 0.1119 2.66 0.0089 **
## UniqueAuthors3 0.2441 0.2079 1.17 0.2426
## Year1997 0.0498 0.2205 0.23 0.8216
## Year1998 -0.4682 0.2125 -2.20 0.0294 *
## Year1999 0.1743 0.3242 0.54 0.5917
## Year2000 0.4572 0.2020 2.26 0.0253 *
## Year2001 0.1004 0.3834 0.26 0.7938
```

```

## Year2002          0.1638      0.4325      0.38      0.7055
## Year2003         -0.2164      0.1696     -1.28      0.2043
## Year2004          0.0489      0.1905      0.26      0.7980
## Year2005          0.0318      0.2168      0.15      0.8837
## Year2006         -0.0688      0.1912     -0.36      0.7195
## Year2007         -0.0220      0.2379     -0.09      0.9265
## Year2008          0.2273      0.2512      0.90      0.3673
## Year2009          0.3743      0.2923      1.28      0.2028
## Year2010          0.1338      0.3830      0.35      0.7274
## Year2011         -0.0021      0.2342     -0.01      0.9929
## Year2012          0.0523      0.2429      0.22      0.8300
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.467
## Multiple R-squared:  0.192, Adjusted R-squared:  0.0633
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 135 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0841 0.8590 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      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 5.430 1      2.330
## LastAuthorFemale  3.520 1      1.876
## Year              3.779 16      1.042

```


Residuals from first and last author



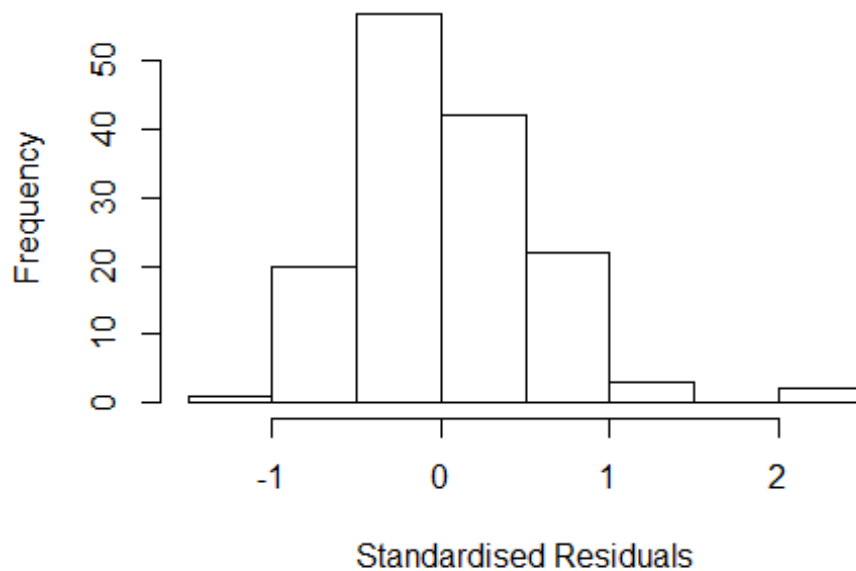
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1564 -0.3178 -0.0129 0.3575 2.0183
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7319 0.1697 4.31 3.2e-05 ***
## FirstAuthorFemale1 0.1420 0.1602 0.89 0.377
## LastAuthorFemale1 -0.0296 0.1413 -0.21 0.834
## Year1997 0.1450 0.2861 0.51 0.613
## Year1998 -0.4625 0.2121 -2.18 0.031 *
## Year1999 0.2546 0.2660 0.96 0.340
## Year2000 0.4614 0.2018 2.29 0.024 *
## Year2001 0.1574 0.3400 0.46 0.644
## Year2002 0.3992 0.3958 1.01 0.315
## Year2003 -0.2119 0.1697 -1.25 0.214
## Year2004 0.1207 0.1842 0.66 0.514
## Year2005 0.1624 0.2377 0.68 0.496
```

```

## Year2006          0.0203      0.2000      0.10      0.919
## Year2007          0.1125      0.2395      0.47      0.639
## Year2008          0.4245      0.2508      1.69      0.093 .
## Year2009          0.4799      0.3104      1.55      0.125
## Year2010          0.2588      0.3366      0.77      0.443
## Year2011          0.2113      0.2349      0.90      0.370
## Year2012          0.2532      0.2281      1.11      0.269
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.507
## Multiple R-squared:  0.137, Adjusted R-squared:  0.0153
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 132 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0767 0.8840 0.9510 0.9020 0.9840 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      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 2.859 1      1.691
## Year              2.859 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
## -1.1524 -0.3150 -0.0137 0.3649 2.0271
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7327 0.1697 4.32 3.1e-05 ***
## FirstAuthorFemale1 0.1244 0.1177 1.06 0.293
## Year1997 0.1393 0.2876 0.48 0.629
## Year1998 -0.4666 0.2124 -2.20 0.030 *
## Year1999 0.2500 0.2637 0.95 0.345
## Year2000 0.4587 0.2027 2.26 0.025 *
## Year2001 0.1614 0.3359 0.48 0.632
## Year2002 0.3987 0.3995 1.00 0.320
## Year2003 -0.2127 0.1697 -1.25 0.212
## Year2004 0.1210 0.1839 0.66 0.512
## Year2005 0.1602 0.2382 0.67 0.503
## Year2006 0.0132 0.2022 0.07 0.948
```

```

## Year2007          0.1054      0.2374      0.44      0.658
## Year2008          0.4196      0.2502      1.68      0.096 .
## Year2009          0.4748      0.3121      1.52      0.131
## Year2010          0.2492      0.3315      0.75      0.454
## Year2011          0.2074      0.2334      0.89      0.376
## Year2012          0.2504      0.2276      1.10      0.273
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.498
## Multiple R-squared:  0.137, Adjusted R-squared:  0.0235
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 134 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0603 0.8820 0.9480 0.9000 0.9830 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      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 2.027 1      1.424
## Year      2.027 16      1.022

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.1594 -0.3263 -0.0155  0.3749  1.9483

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.7316    0.1697    4.31 3.2e-05 ***
## LastAuthorFemale1 0.0588    0.1063    0.55  0.581
## Year1997        0.1648    0.2775    0.59  0.554
## Year1998       -0.4481    0.2091   -2.14  0.034 *
## Year1999        0.2710    0.2785    0.97  0.332
## Year2000        0.4711    0.1979    2.38  0.019 *
## Year2001        0.1897    0.3209    0.59  0.556
## Year2002        0.3994    0.3943    1.01  0.313
## Year2003       -0.2116    0.1697   -1.25  0.215
## Year2004        0.1320    0.1824    0.72  0.471
## Year2005        0.2130    0.2257    0.94  0.347
## Year2006        0.0218    0.1998    0.11  0.913
## Year2007        0.1087    0.2396    0.45  0.651
## Year2008        0.4279    0.2480    1.73  0.087 .
## Year2009        0.5118    0.3232    1.58  0.116
## Year2010        0.3291    0.3264    1.01  0.315
## Year2011        0.2015    0.2329    0.87  0.389
## Year2012        0.2610    0.2314    1.13  0.261
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.51
## Multiple R-squared:  0.131, Adjusted R-squared:  0.0165
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 133 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.113  0.885   0.949   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      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] ""

```

```

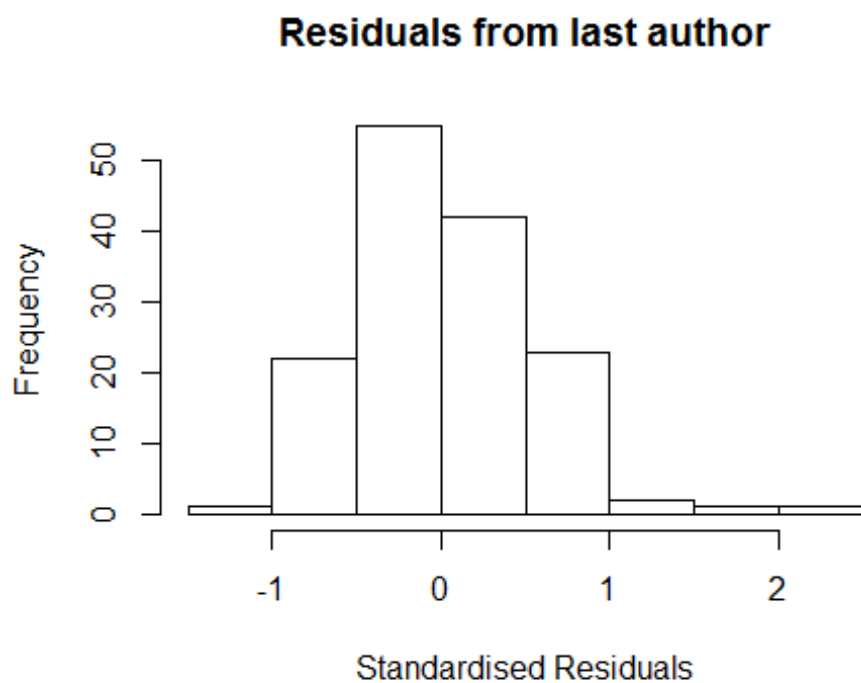
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2100"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1997 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    2    2    1    3    2    2    2    9   18   30   27   24   24
##
## 1997 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    0    0    0    1    2    1    1    1    8    8   19    9   16   14
##
## 1997 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    0    0    0    0    2    0    1    1    8    8   16    9   15   12
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.49146187923103"
## [1] "Male first author team size 2018 geometric mean: 3.03250120335048"

## Warning in wilcox.test.default(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 = 0.4
## 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: 2.73053814226843"

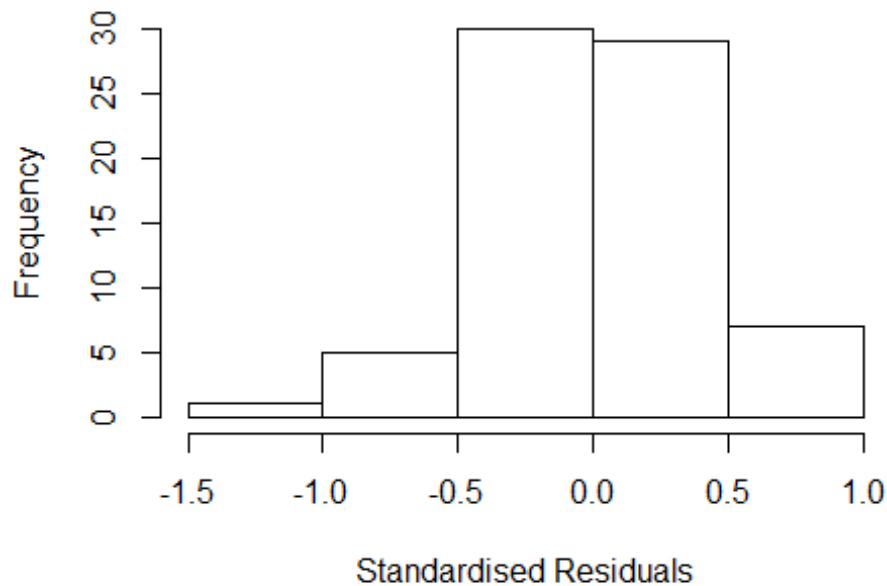
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 39, p-value = 0.7
## 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 1.739e+14 1      1.319e+07
## LastAuthorFemale  3.233e+14 1      1.798e+07
## UniqueAuthors    2.433e+15 4      8.380e+01
## Year              1.098e+16 8      1.006e+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
## -1.41784 -0.25065 0.00948 0.26240 0.89996
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2221 0.2764 4.42 4.5e-05 ***
## FirstAuthorFemale1 0.0502 0.1349 0.37 0.71
## LastAuthorFemale1 -0.0434 0.1839 -0.24 0.81
## UniqueAuthors2 0.2709 0.1636 1.66 0.10
## UniqueAuthors3 0.2649 0.1674 1.58 0.12
## UniqueAuthors4 0.2150 0.2297 0.94 0.35
## UniqueAuthors5 0.2538 0.1616 1.57 0.12
## Year2005 -1.2221 0.2764 -4.42 4.5e-05 ***
## Year2006 0.5542 0.3432 1.61 0.11
## Year2007 -0.0242 0.3091 -0.08 0.94
```

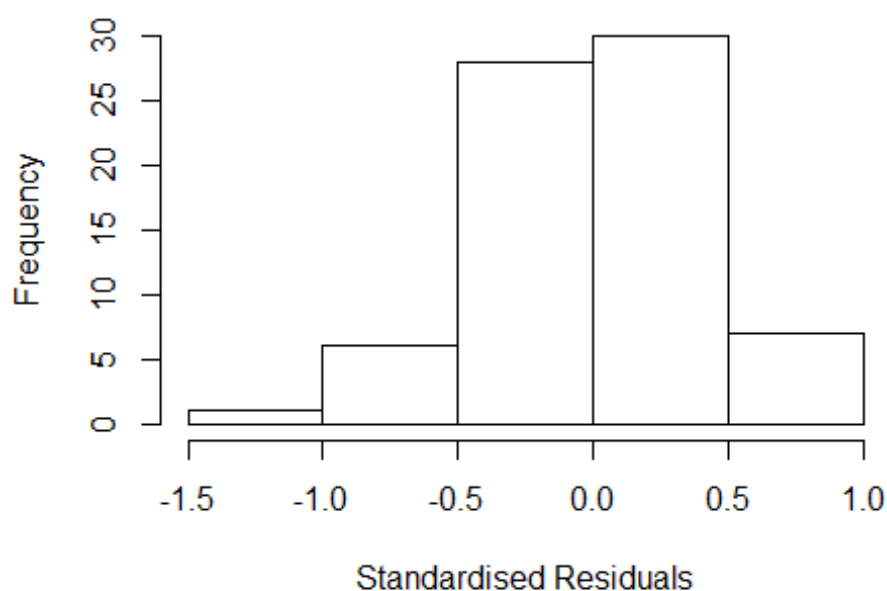


```

## Year2008          0.0245      0.2911      0.08      0.93
## Year2009          0.0809      0.2837      0.29      0.78
## Year2010         -0.0751      0.3408     -0.22      0.83
## Year2011         -0.1970      0.3046     -0.65      0.52
## Year2012         -0.1068      0.3100     -0.34      0.73
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.438
## Multiple R-squared:  0.256, Adjusted R-squared:  0.0727
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 64 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.272  0.903   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      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 1.860e+14 1      1.364e+07
## LastAuthorFemale  3.073e+14 1      1.753e+07
## Year              4.278e+14 8      8.212e+00

```

Residuals from first and last author



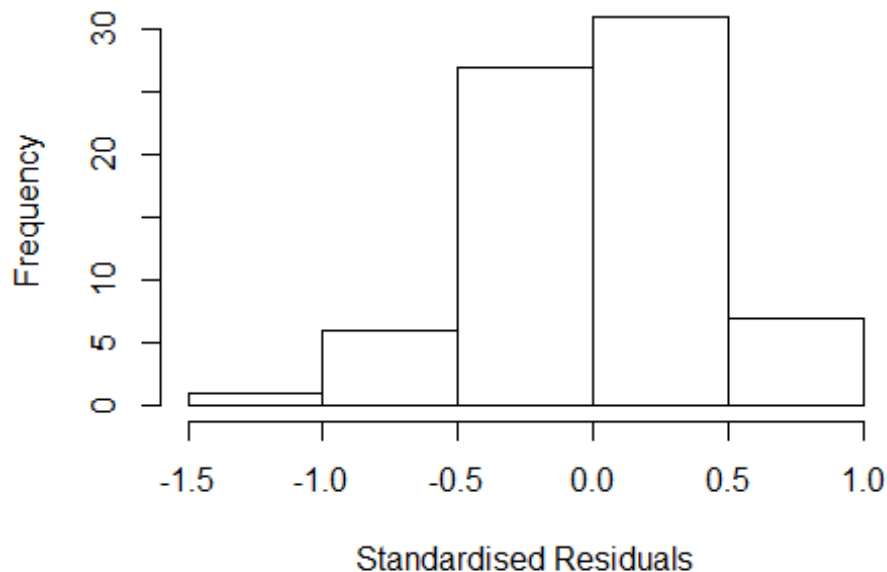
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3439 -0.2728 0.0509 0.2211 0.9331
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3575 0.4256 3.19 0.0023 **
## FirstAuthorFemale1 0.0684 0.1292 0.53 0.5985
## LastAuthorFemale1 -0.0865 0.1660 -0.52 0.6044
## Year2005 -1.3575 0.4256 -3.19 0.0023 **
## Year2006 0.7086 0.4495 1.58 0.1201
## Year2007 0.0690 0.4578 0.15 0.8806
## Year2008 0.1372 0.4344 0.32 0.7532
## Year2009 0.1040 0.4413 0.24 0.8145
## Year2010 -0.0136 0.4532 -0.03 0.9762
## Year2011 -0.1300 0.4539 -0.29 0.7755
## Year2012 0.0127 0.4484 0.03 0.9776
## ---
```

```

## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.431
## Multiple R-squared:  0.213, Adjusted R-squared:  0.0836
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 66 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.311  0.896  0.959  0.914  0.989  0.998
## Algorithmic parameters:
##           tuning.chi                bb           tuning.psi           refine.tol
##           1.55e+00                5.00e-01           4.69e+00           1.00e-07
##           rel.tol                solve.tol           eps.outlier           eps.x
##           1.00e-07                1.00e-07           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 1.257e+14 1      1.121e+07
## Year              1.257e+14 8      7.607e+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
## -1.3279 -0.2770 0.0247 0.2382 0.9784
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3575 0.4157 3.27 0.0018 **
## FirstAuthorFemale1 0.0457 0.1143 0.40 0.6908
## Year2005 -1.3575 0.4157 -3.27 0.0018 **
## Year2006 0.6448 0.4326 1.49 0.1411
## Year2007 0.0516 0.4511 0.11 0.9092
## Year2008 0.1399 0.4245 0.33 0.7428
## Year2009 0.0985 0.4322 0.23 0.8204
## Year2010 -0.0296 0.4465 -0.07 0.9474
## Year2011 -0.1526 0.4380 -0.35 0.7287
## Year2012 0.0150 0.4392 0.03 0.9729
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Robust residual standard error: 0.449
## Multiple R-squared: 0.206, Adjusted R-squared: 0.0909
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 64 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.363  0.905  0.951  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      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 4.289e+14 1      2.071e+07
## Year          4.289e+14 8      8.213e+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
## -1.3469 -0.2759  0.0461  0.2285  0.9887
##
## Coefficients:
##      Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3575     0.4210   3.22  0.002 **
## LastAuthorFemale1 -0.0564     0.1372  -0.41  0.683
## Year2005          -1.3575     0.4210  -3.22  0.002 **
## Year2006           0.7469     0.4445   1.68  0.098 .
## Year2007           0.0732     0.4533   0.16  0.872
```

```

## Year2008          0.1456      0.4289      0.34      0.735
## Year2009          0.1232      0.4344      0.28      0.778
## Year2010         -0.0106      0.4480     -0.02      0.981
## Year2011         -0.1172      0.4520     -0.26      0.796
## Year2012          0.0424      0.4370      0.10      0.923
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.439
## Multiple R-squared:  0.209, Adjusted R-squared:  0.0939
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.327  0.902  0.953  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      1.39e-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 2101"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 2011
##    2
##
## 2011
##    2
##
## 2011
##    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: 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 = 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: 3.55689330449006"
## [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
## 3 6 3 1 2 3 4 2 5 9 6 13 6 8 13
## 2011 2012
## 12 10
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1 1 0 0 2 0 1 2 2 4 3 11 4 5 4
## 2011 2012
## 5 4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 0 1 0 0 2 0 1 1 2 3 3 11 4 4 3
## 2011 2012
## 5 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: 2.45126927006581"

```

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

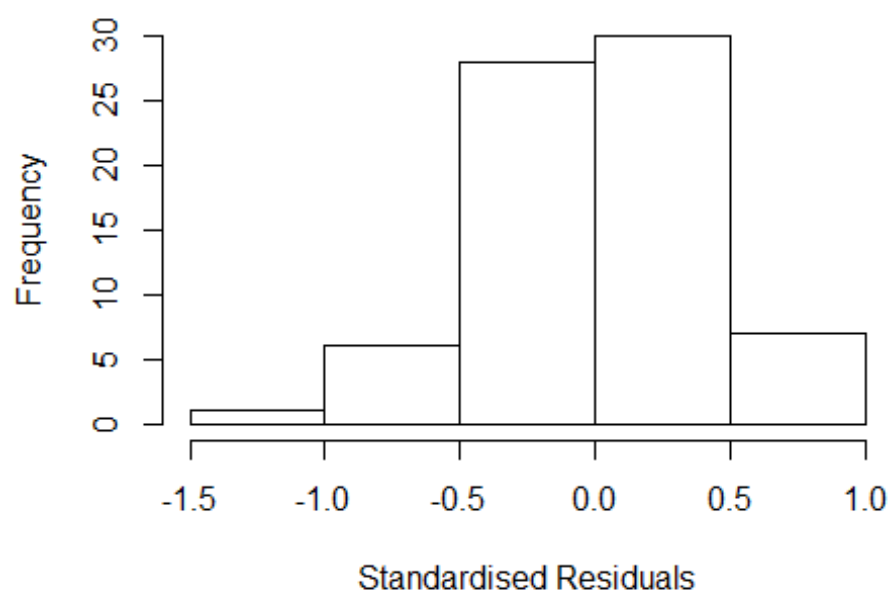
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 11, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6"
## [1] "Male last author team size 2018 geometric mean: 2.45126927006581"

## Warning in wilcox.test.default(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"

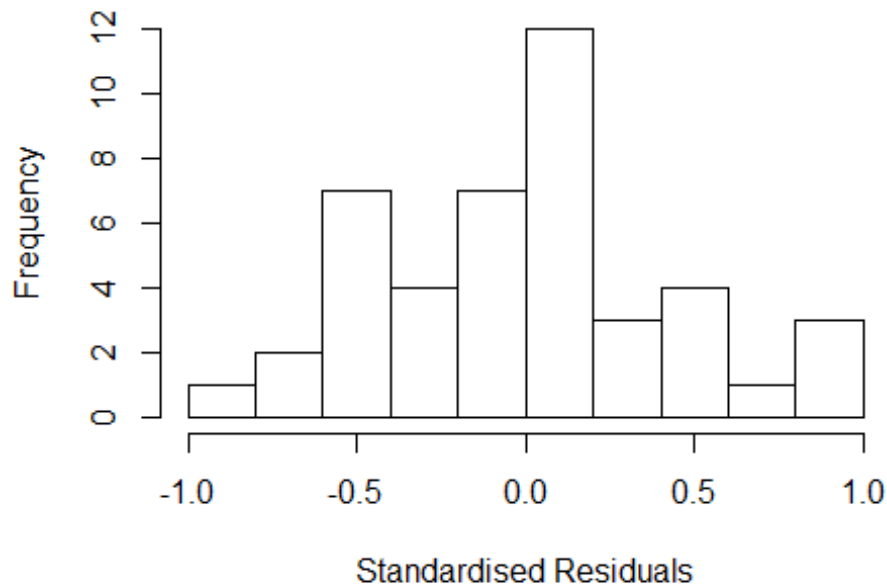
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead
```


Residuals from last author



| | | | | |
|----|-------------------|-----------|----|---------------------|
| ## | | GVIF | Df | $GVIF^{(1/(2*Df))}$ |
| ## | FirstAuthorFemale | 6.592e+13 | 1 | 8.119e+06 |
| ## | LastAuthorFemale | 7.550e+02 | 1 | 2.748e+01 |
| ## | UniqueAuthors | 3.950e+30 | 4 | 6.677e+03 |
| ## | Year | 1.404e+34 | 12 | 2.647e+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
## -0.9605 -0.2993 0.0113 0.2005 0.9311
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.8220 0.0000 Inf < 2e-16 ***
## FirstAuthorFemale1 0.1157 0.1889 0.61 0.54576
## LastAuthorFemale1 -0.0484 0.2311 -0.21 0.83589
## UniqueAuthors2 1.2046 0.3652 3.30 0.00292 **
## UniqueAuthors3 0.9744 0.3774 2.58 0.01609 *
## UniqueAuthors4 1.5551 0.3866 4.02 0.00047 ***
## UniqueAuthors5 0.1953 0.4455 0.44 0.66493
## Year2000 -1.5287 0.4034 -3.79 0.00085 ***
## Year2002 -0.0760 0.4592 -0.17 0.86996
## Year2003 -1.5026 0.3652 -4.11 0.00037 ***
```

```

## Year2004          -0.9526      0.1640    -5.81  4.7e-06 ***
## Year2005          -0.1117      0.3976    -0.28  0.78108
## Year2006          -0.7552      0.3726    -2.03  0.05345 .
## Year2007          -1.5221      0.2896    -5.26  1.9e-05 ***
## Year2008          -1.2449      0.2896    -4.30  0.00023 ***
## Year2009          -1.8867      0.4622    -4.08  0.00040 ***
## Year2010          -0.5724      0.4134    -1.38  0.17842
## Year2011          -1.4883      0.4170    -3.57  0.00148 **
## Year2012          -0.3436      0.5169    -0.66  0.51232
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.569
## Multiple R-squared:  0.68,   Adjusted R-squared:  0.45
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~ 1. The remaining 34 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.757  0.920  0.953   0.936  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 2: First author gender, Last author gender, Year as
factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 4.267e+13  1      6.532e+06
## LastAuthorFemale  1.323e+01  1      3.638e+00
## Year              4.097e+14 12      4.063e+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,

```

```

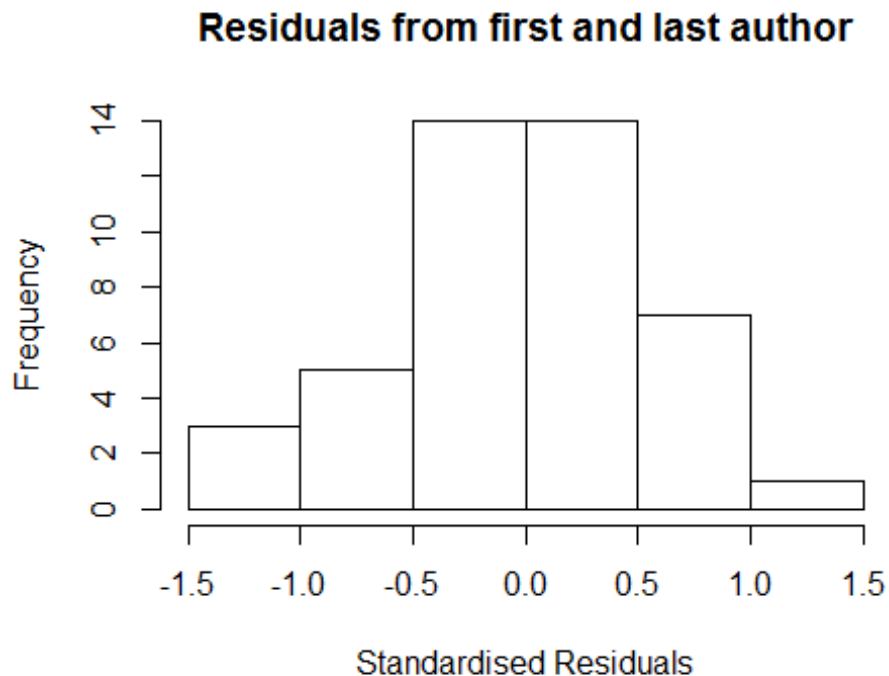
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q    Median        3Q        Max
## -1.1229 -0.2924  0.0418  0.2437  1.3567
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   1.82e+00   3.71e-08  4.91e+07 < 2e-16 ***
## FirstAuthorFemale1 2.43e-01   2.23e-01   1.09e+00  0.28420
## LastAuthorFemale1 2.51e-01   3.35e-01   7.50e-01  0.45922
## Year2000       -1.01e+00   1.49e-01  -6.77e+00   2e-07 ***
## Year2002       -7.97e-03   2.23e-01  -4.00e-02  0.97169
## Year2003       -2.98e-01   3.57e-08  -8.35e+06 < 2e-16 ***
## Year2004       -1.17e+00   2.94e-01  -3.96e+00  0.00044 ***
## Year2005        9.01e-01   2.44e-01   3.69e+00  0.00093 ***
## Year2006        2.59e-03   4.42e-01   1.00e-02  0.99536
## Year2007       -6.99e-01   2.40e-01  -2.91e+00  0.00679 **
## Year2008       -4.69e-01   7.19e-01  -6.50e-01  0.51955
## Year2009       -5.67e-01   2.71e-01  -2.09e+00  0.04561 *
## Year2010        4.31e-01   1.43e-01   3.02e+00  0.00529 **
## Year2011       -3.92e-01   2.48e-01  -1.58e+00  0.12419
## Year2012        6.04e-01   3.21e-01   1.88e+00  0.07016 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.638
## Multiple R-squared:  0.52, Adjusted R-squared:  0.289
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 40 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.630  0.920  0.984    0.932  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      2.27e-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 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 12          NaN

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1707 -0.3242  0.0423  0.2606  1.2403
##
## Coefficients:
```

```

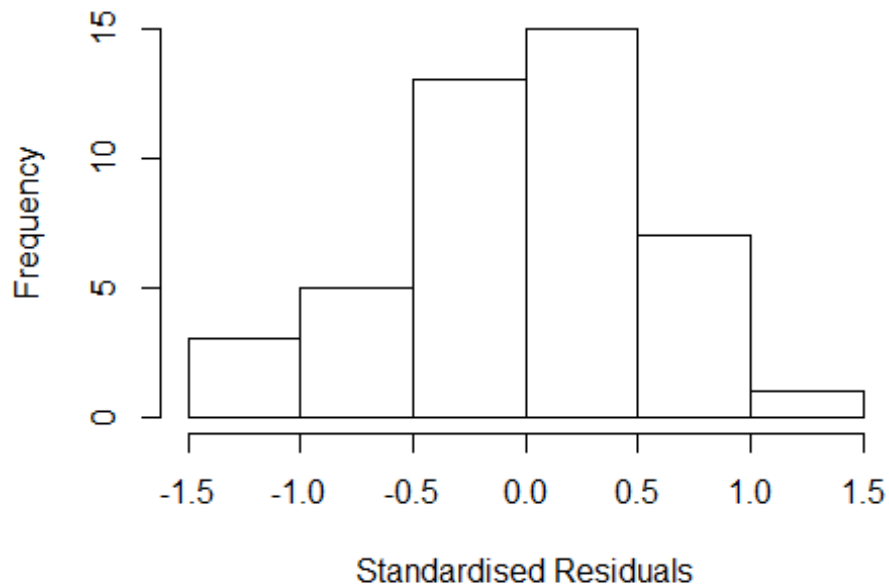
##               Estimate Std. Error   t value Pr(>|t|)
## (Intercept)    1.82e+00   2.42e-08  7.53e+07 < 2e-16 ***
## FirstAuthorFemale1 2.30e-01   2.09e-01   1.10e+00  0.28040
## Year2000       -1.00e+00   1.47e-01  -6.80e+00  1.5e-07 ***
## Year2002         5.04e-03   2.09e-01   2.00e-02  0.98095
## Year2003       -2.98e-01   0.00e+00    -Inf < 2e-16 ***
## Year2004       -1.03e+00   1.28e-01  -8.08e+00  5.2e-09 ***
## Year2005         9.05e-01   2.44e-01   3.71e+00  0.00083 ***
## Year2006         3.98e-03   4.32e-01   1.00e-02  0.99272
## Year2007       -6.66e-01   2.18e-01  -3.05e+00  0.00474 **
## Year2008       -3.52e-01   5.92e-01  -6.00e-01  0.55600
## Year2009       -5.64e-01   2.68e-01  -2.10e+00  0.04401 *
## Year2010         4.35e-01   1.40e-01   3.12e+00  0.00401 **
## Year2011       -3.89e-01   2.44e-01  -1.59e+00  0.12177
## Year2012         6.00e-01   3.14e-01   1.91e+00  0.06530 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.669
## Multiple R-squared:  0.508, Adjusted R-squared:  0.294
## 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.
##   0.711  0.925  0.981  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           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

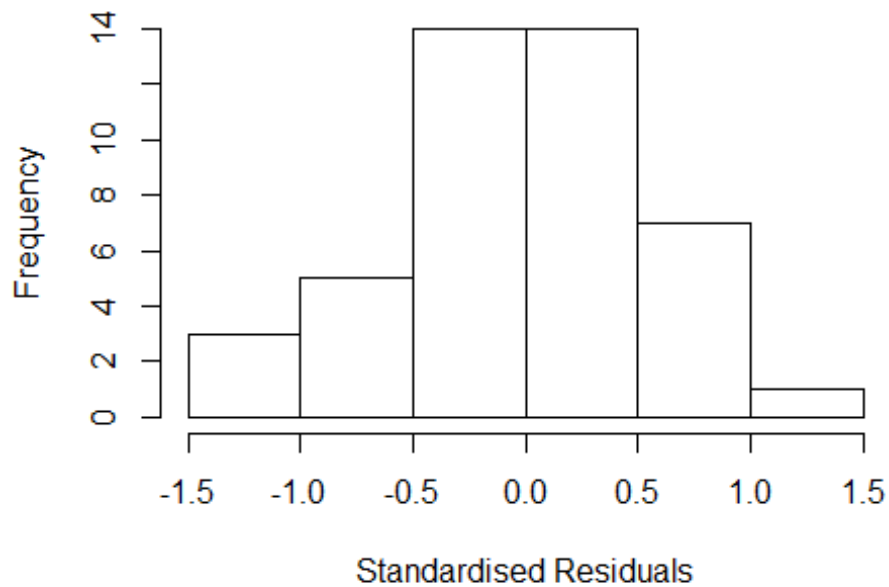
```

Residuals from first author



```
##          GVIF Df GVIF^(1/(2*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
## -1.2179 -0.3708 0.0273 0.1990 1.2313
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.82e+00 1.25e-08 1.46e+08 < 2e-16 ***
## LastAuthorFemale1 2.11e-01 2.75e-01 7.70e-01 0.4497
## Year2000 -8.86e-01 1.89e-01 -4.70e+00 5.4e-05 ***
## Year2002 2.35e-01 0.00e+00 Inf < 2e-16 ***
## Year2003 -2.98e-01 2.94e-08 -1.01e+07 < 2e-16 ***
## Year2004 -1.02e+00 1.53e-01 -6.70e+00 2.0e-07 ***
## Year2005 9.83e-01 2.71e-01 3.63e+00 0.0010 **
## Year2006 1.32e-01 4.20e-01 3.20e-01 0.7549
## Year2007 -6.04e-01 2.26e-01 -2.67e+00 0.0121 *
## Year2008 -3.43e-01 5.97e-01 -5.70e-01 0.5698
## Year2009 -5.62e-01 2.68e-01 -2.10e+00 0.0444 *
## Year2010 5.13e-01 1.32e-01 3.90e+00 0.0005 ***
## Year2011 -3.38e-01 2.46e-01 -1.37e+00 0.1805
## Year2012 6.32e-01 2.69e-01 2.35e+00 0.0256 *
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.704
## Multiple R-squared: 0.49, Adjusted R-squared: 0.269
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 35 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.741 0.901 0.965 0.934 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 2.27e-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 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
##    4    2    3    5    2    4    7    2    3    4    3    6    4    3    8
## 2011 2012
##    4    4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    0    1    2    2    1    2    2    2    2    0    4    3    2    0
## 2011 2012
##    1    0
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    0    1    1    2    1    1    1    2    0    0    4    3    1    0
## 2011 2012
##    1    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: 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: 6"
## [1] "Male last 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] "Regression 1: First author gender, last author gender, team size,

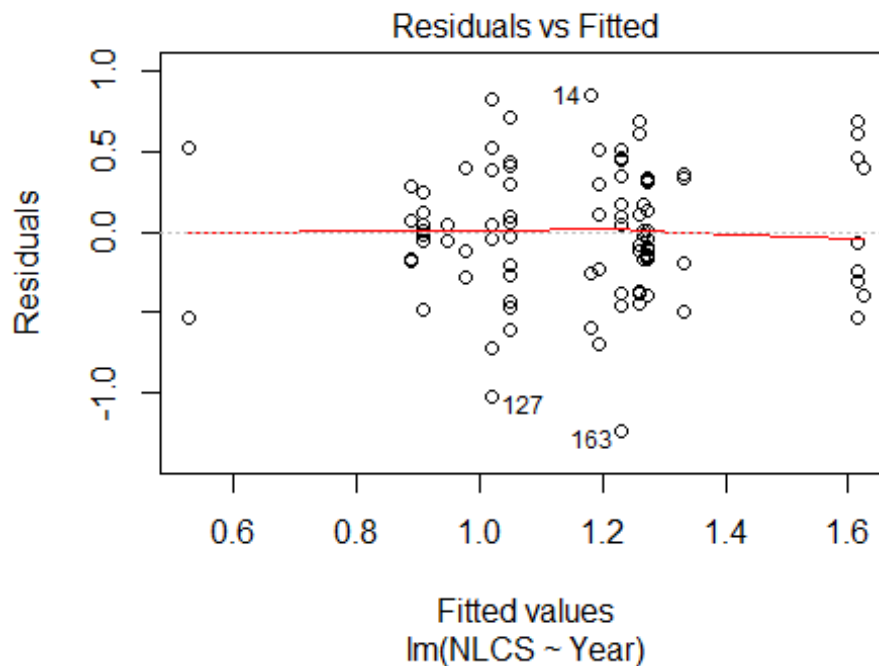
```

```

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: 19"
## [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 2002 2004 2005 2006 2007 2008 2009 2011 2012
##    2    4    3    1    2    1    1    3    3    2    1    4    1    2
##
## 1996 1997 1998 1999 2000 2002 2004 2005 2006 2007 2008 2009 2011 2012
##    2    0    0    0    1    0    1    1    1    1    0    1    0    0
##
## 1996 1997 1998 1999 2000 2002 2004 2005 2006 2007 2008 2009 2011 2012
##    1    0    0    0    1    0    1    0    1    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: 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"
## [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: 5"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2105"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    3    5    3    3    5    4    9    17    9    9    20    17    16    23
## 2011 2012
##   21   23
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    2    3    2    2    3    2    5    8    4    2    12    7    8    10
## 2011 2012
##   12    9

```

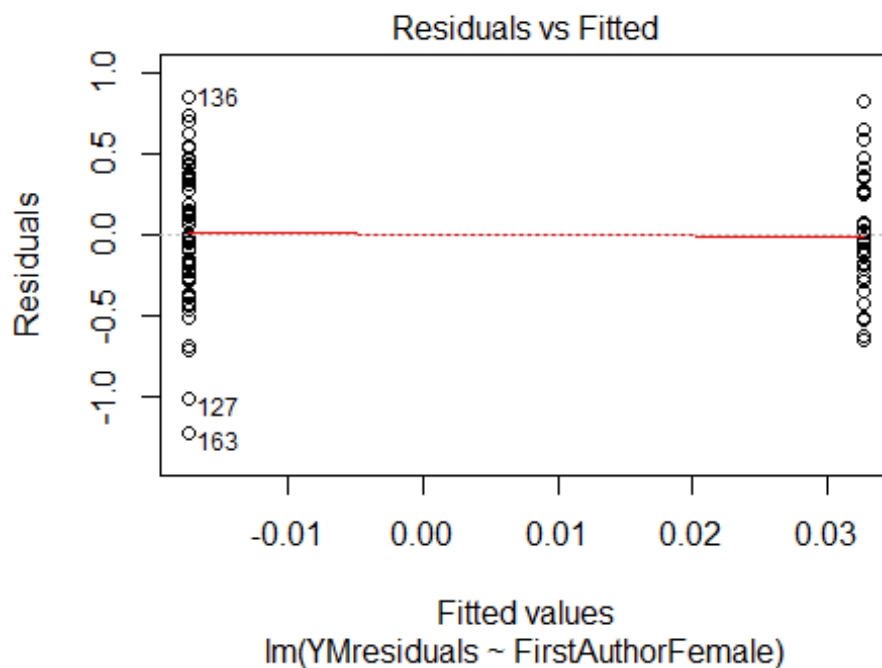
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    2    3    1    2    3    1    5    5    2    2   12    7    7    9
## 2011 2012
##   11    7
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 25, df = 16, p-value = 0.08
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.85, df = 1, p-value = 0.4
## [1] "Female first author team size 2018 geometric mean: 3.0261714988458"
## [1] "Male first author team size 2018 geometric mean: 2.93625788475197"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
##
## Wilcoxon rank sum test with continuity correction
```

```
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 86, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.77976314968462"
## [1] "Male last author team size 2018 geometric mean: 2.79962845899"

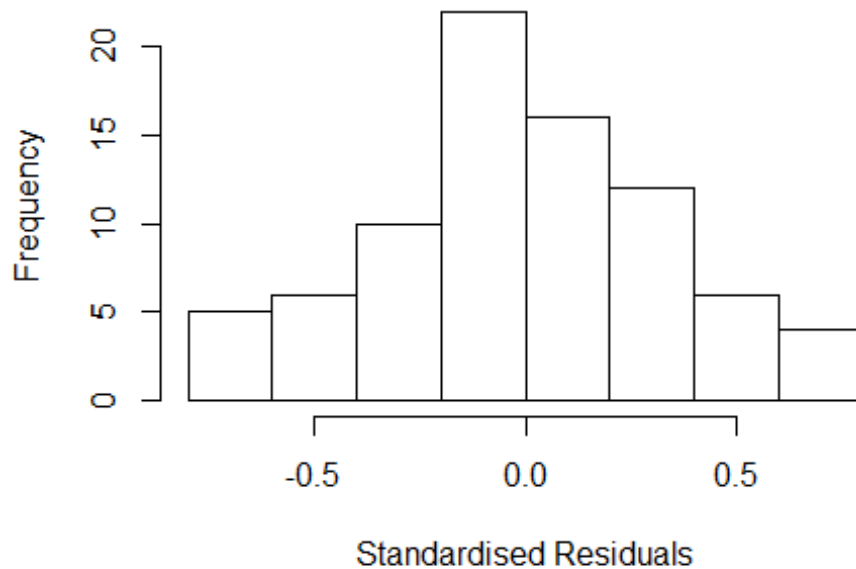
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## 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] "Regression 1: First author gender, last author gender, team size,
## Year as factors"
##
```

| | GVIF | Df | GVIF ^{1/(2*Df)} |
|-------------------|------------|----|--------------------------|
| FirstAuthorFemale | -1.575e+14 | 1 | NaN |
| LastAuthorFemale | 2.511e+00 | 1 | 1.585 |
| UniqueAuthors | -1.699e+16 | 4 | NaN |
| Year | -6.188e+16 | 16 | NaN |

Residuals from first and last author and team size



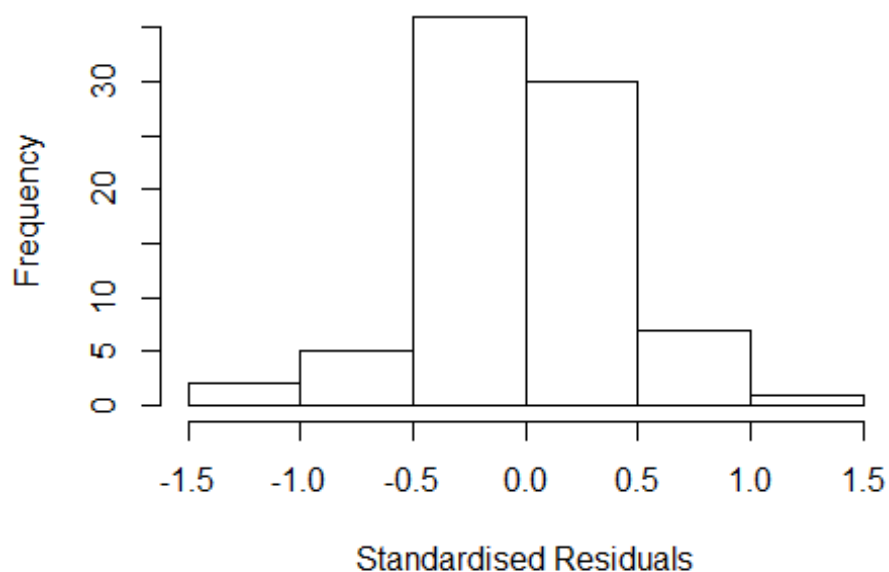
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.79567 -0.21062 -0.00613 0.27348 0.79567
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9374 0.3276 2.86 0.0059 **
## FirstAuthorFemale1 0.0011 0.0960 0.01 0.9909
## LastAuthorFemale1 0.2784 0.1277 2.18 0.0333 *
## UniqueAuthors2 0.3477 0.1540 2.26 0.0277 *
## UniqueAuthors3 0.5393 0.1750 3.08 0.0032 **
## UniqueAuthors4 0.5446 0.2097 2.60 0.0119 *
## UniqueAuthors5 -0.0613 0.3190 -0.19 0.8483
## Year1997 -0.6811 1.8831 -0.36 0.7189
## Year1998 -0.2436 0.5513 -0.44 0.6602
## Year1999 0.5399 0.3645 1.48 0.1440
```

```

## Year2000          -0.4759      0.3520    -1.35    0.1816
## Year2001          -0.2584      0.3314    -0.78    0.4388
## Year2002          -0.3777      0.3448    -1.10    0.2778
## Year2003          -0.1543      0.3619    -0.43    0.6715
## Year2004          -0.3546      0.3672    -0.97    0.3383
## Year2005          -0.5802      0.3509    -1.65    0.1037
## Year2006          -0.1167      0.3464    -0.34    0.7375
## Year2007          -0.3047      0.3324    -0.92    0.3630
## Year2008          -0.1766      0.3772    -0.47    0.6414
## Year2009          -0.0288      0.3374    -0.09    0.9324
## Year2010          -0.0864      0.3702    -0.23    0.8163
## Year2011          -0.1372      0.3404    -0.40    0.6884
## Year2012           0.4533      0.3506     1.29    0.2011
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.441
## Multiple R-squared:  0.412, Adjusted R-squared:  0.188
## Convergence in 39 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 72 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.725  0.910  0.961  0.934  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      1.23e-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 2: First author gender, Last author gender, Year as
factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 8.591e+13 1      9.269e+06
## LastAuthorFemale  3.169e+00 1      1.780e+00
## Year              2.846e+14 16      2.829e+00

```

Residuals from first and last author



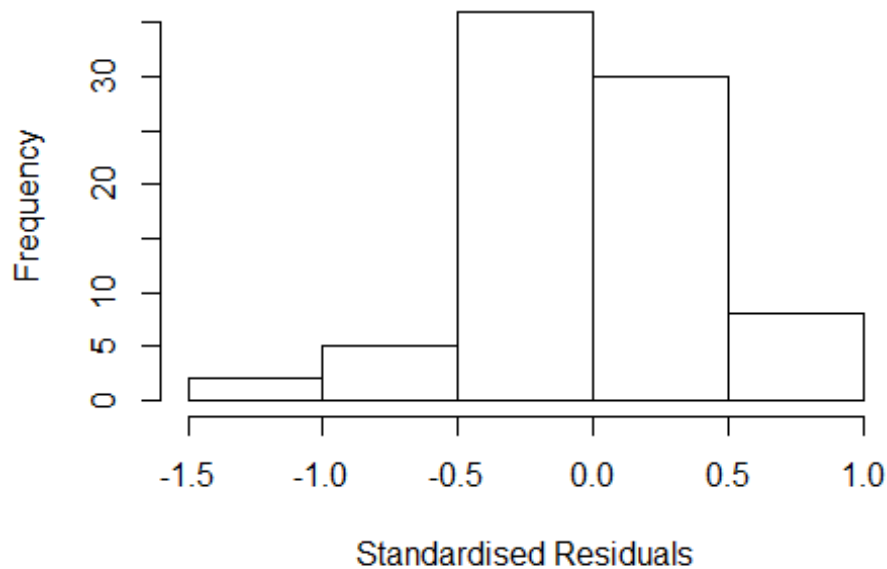
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.315 -0.278 -0.011 0.307 1.096
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1929 0.4493 2.66 0.010 *
## FirstAuthorFemale1 -0.0574 0.1200 -0.48 0.634
## LastAuthorFemale1 0.1736 0.1300 1.34 0.187
## Year1997 -0.6669 0.7088 -0.94 0.350
## Year1998 -0.2026 0.6521 -0.31 0.757
## Year1999 0.8875 0.4612 1.92 0.059 .
## Year2000 -0.2728 0.4616 -0.59 0.557
## Year2001 -0.2245 0.4841 -0.46 0.644
## Year2002 -0.0939 0.4493 -0.21 0.835
## Year2003 0.0109 0.5093 0.02 0.983
## Year2004 -0.3789 0.4521 -0.84 0.405
## Year2005 -0.3105 0.4622 -0.67 0.504
```

```

## Year2006          0.1325      0.4613      0.29      0.775
## Year2007         -0.1802      0.4574     -0.39      0.695
## Year2008         -0.0537      0.5749     -0.09      0.926
## Year2009          0.0720      0.4732      0.15      0.880
## Year2010          0.1224      0.4901      0.25      0.804
## Year2011          0.0666      0.4543      0.15      0.884
## Year2012          0.5492      0.4891      1.12      0.266
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.412
## Multiple R-squared:  0.299, Adjusted R-squared:  0.0959
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 69 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.286  0.887  0.941  0.902  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      1.23e-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"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -8.032e+13  1      NaN
## Year              -8.032e+13 16      NaN

```


Residuals from first author



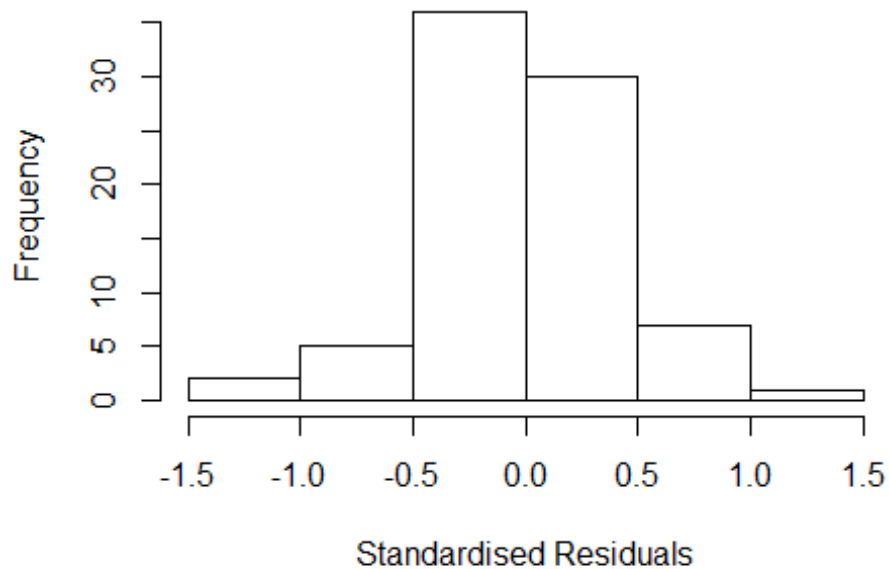
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.30473 -0.26880 -0.00925 0.31972 0.94386
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.25278 0.34735 3.61 0.00061 ***
## FirstAuthorFemale1 -0.00357 0.10400 -0.03 0.97273
## Year1997 -0.72678 0.61516 -1.18 0.24186
## Year1998 -0.16407 0.58387 -0.28 0.77963
## Year1999 0.77378 0.34921 2.22 0.03033 *
## Year2000 -0.29972 0.35104 -0.85 0.39646
## Year2001 -0.28295 0.39089 -0.72 0.47183
## Year2002 -0.15378 0.34735 -0.44 0.65947
## Year2003 -0.03792 0.40416 -0.09 0.92554
## Year2004 -0.38711 0.35927 -1.08 0.28536
## Year2005 -0.31050 0.38528 -0.81 0.42333
## Year2006 0.01878 0.34930 0.05 0.95728
```

```

## Year2007          -0.20598    0.36582   -0.56  0.57539
## Year2008          -0.16394    0.45879   -0.36  0.72205
## Year2009           0.04783    0.38341    0.12  0.90112
## Year2010           0.05195    0.39830    0.13  0.89665
## Year2011          -0.00653    0.35014   -0.02  0.98517
## Year2012           0.47280    0.39139    1.21  0.23156
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.45
## Multiple R-squared:  0.272, Adjusted R-squared:  0.0755
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 70 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.380  0.897  0.942  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      1.23e-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 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 5.584 1          2.363
## Year              5.584 16          1.055

```

Residuals from last author



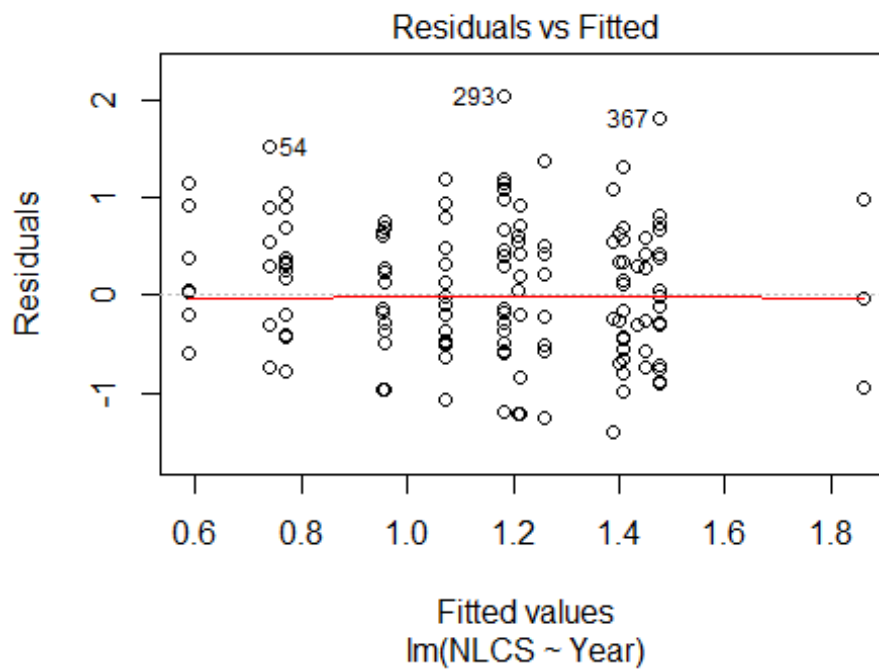
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2982 -0.2788 -0.0054 0.3044 1.0670
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.17608 0.47034 2.50 0.015 *
## LastAuthorFemale1 0.14984 0.11731 1.28 0.206
## Year1997 -0.65008 0.71328 -0.91 0.366
## Year1998 -0.21408 0.67368 -0.32 0.752
## Year1999 0.84692 0.47034 1.80 0.077 .
## Year2000 -0.30150 0.47437 -0.64 0.527
## Year2001 -0.20730 0.50306 -0.41 0.682
## Year2002 -0.07708 0.47034 -0.16 0.870
## Year2003 0.00549 0.52462 0.01 0.992
## Year2004 -0.37418 0.47487 -0.79 0.434
## Year2005 -0.31050 0.47921 -0.65 0.519
## Year2006 0.09192 0.47040 0.20 0.846
```

```

## Year2007          -0.17152      0.47886    -0.36      0.721
## Year2008          -0.07350      0.56969    -0.13      0.898
## Year2009           0.07667      0.49402     0.16      0.877
## Year2010           0.12210      0.50785     0.24      0.811
## Year2011           0.05377      0.47332     0.11      0.910
## Year2012           0.54628      0.50540     1.08      0.284
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.42
## Multiple R-squared:  0.293, Adjusted R-squared:  0.103
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 70 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.319  0.893   0.938   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      1.23e-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: 81"
## [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
##   16  14  18  15  17  10   8   9   7  13   7  10  23  27  39
## 2011 2012
##   28  30
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    8    5    9   11    9    4    2    4    3    7    4    7   15   15   23
## 2011 2012

```

```
## 16 17
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 7 5 9 7 9 4 2 3 2 5 4 6 14 12 22
## 2011 2012
## 13 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 = 8.5, df = 16, p-value = 0.9
```



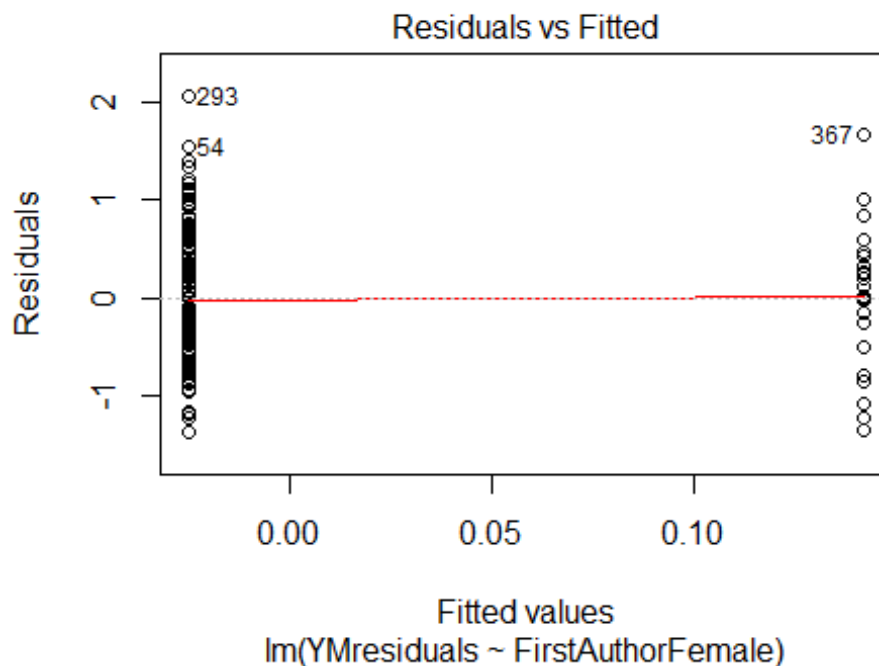
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.073, df = 1, p-value = 0.8

## [1] "Female first author team size 2018 geometric mean: 3.16227766016838"
## [1] "Male first author team size 2018 geometric mean: 1.70673683684508"

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

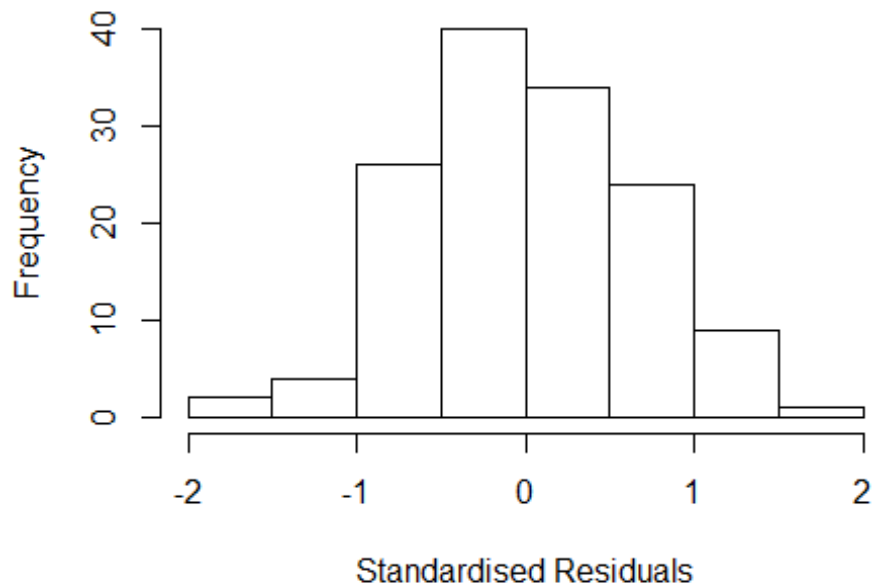
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 12, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5"
## [1] "Male last author team size 2018 geometric mean: 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 = 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 1.534 1      1.239
## LastAuthorFemale  2.068 1      1.438
## UniqueAuthors    4.147 4      1.195
## Year              8.008 16      1.067
```

Residuals from first and last author and team size



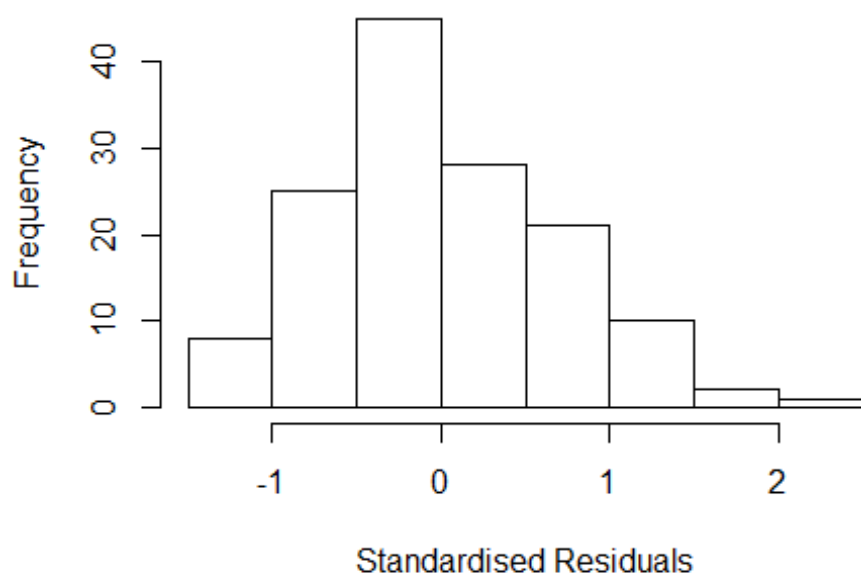
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5908 -0.4656 -0.0365  0.4633  1.5546
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2424    0.3392   3.66 0.00038 ***
## FirstAuthorFemale1  0.1866    0.1671   1.12 0.26651
## LastAuthorFemale1 -0.1421    0.1979  -0.72 0.47429
## UniqueAuthors2     0.0694    0.1533   0.45 0.65165
## UniqueAuthors3    -0.0206    0.1700  -0.12 0.90355
## UniqueAuthors4     0.2960    0.2384   1.24 0.21692
## UniqueAuthors5     0.9637    0.3718   2.59 0.01075 *
## Year1997          -0.2614    0.4407  -0.59 0.55432
## Year1998          -0.6154    0.4969  -1.24 0.21803
## Year1999          -0.6010    0.4074  -1.48 0.14287
```

```

## Year2000          -0.3548      0.3774    -0.94  0.34904
## Year2001          0.1519      0.4375     0.35  0.72897
## Year2002          0.1579      0.4072     0.39  0.69892
## Year2003          0.2863      0.3977     0.72  0.47300
## Year2004          0.9636      0.4358     2.21  0.02899 *
## Year2005          0.1369      0.4885     0.28  0.77970
## Year2006          0.2180      0.6609     0.33  0.74209
## Year2007          0.0657      0.3778     0.17  0.86228
## Year2008         -0.3461      0.3625    -0.95  0.34162
## Year2009         -0.4994      0.3774    -1.32  0.18838
## Year2010         -0.3745      0.3587    -1.04  0.29865
## Year2011         -0.2147      0.3829    -0.56  0.57611
## Year2012         -0.0039      0.3613    -0.01  0.99140
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.778
## Multiple R-squared:  0.249, Adjusted R-squared:  0.108
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 125 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.655  0.908  0.955  0.932  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      7.14e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.745 1      1.321
## LastAuthorFemale  2.967 1      1.723
## Year              4.596 16      1.049

```


Residuals from first and last author



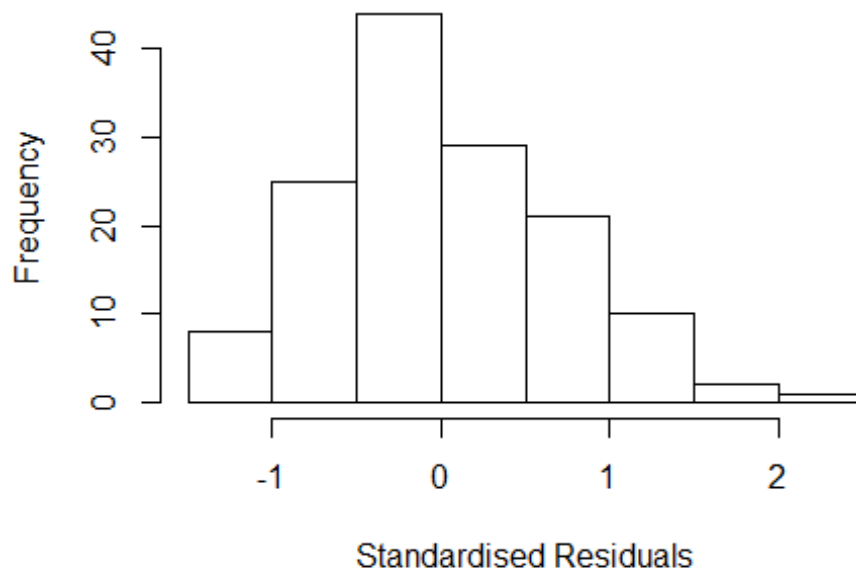
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4618 -0.4648 -0.0511 0.4755 2.1402
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.28393 0.33751 3.80 0.00022 ***
## FirstAuthorFemale1 0.29461 0.17616 1.67 0.09704 .
## LastAuthorFemale1 -0.02387 0.25713 -0.09 0.92619
## Year1997 -0.31208 0.44971 -0.69 0.48903
## Year1998 -0.59784 0.45359 -1.32 0.18999
## Year1999 -0.61358 0.42456 -1.45 0.15098
## Year2000 -0.42042 0.39315 -1.07 0.28703
## Year2001 0.11993 0.43426 0.28 0.78288
## Year2002 0.15107 0.40479 0.37 0.70964
## Year2003 0.24405 0.40683 0.60 0.54970
## Year2004 0.90277 0.43844 2.06 0.04164 *
## Year2005 0.10734 0.46792 0.23 0.81896
```

```

## Year2006          0.17783      0.67884      0.26  0.79380
## Year2007          0.07661      0.39303      0.19  0.84579
## Year2008         -0.32234      0.37936     -0.85  0.39717
## Year2009         -0.49752      0.39253     -1.27  0.20742
## Year2010         -0.20915      0.39149     -0.53  0.59415
## Year2011         -0.00836      0.38763     -0.02  0.98284
## Year2012          0.11364      0.38890      0.29  0.77064
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.777
## Multiple R-squared:  0.166, Adjusted R-squared:  0.0421
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 127 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.429  0.902  0.957  0.926  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      7.14e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##      trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.579 1      1.257
## Year              1.579 16      1.014

```

Residuals from first author



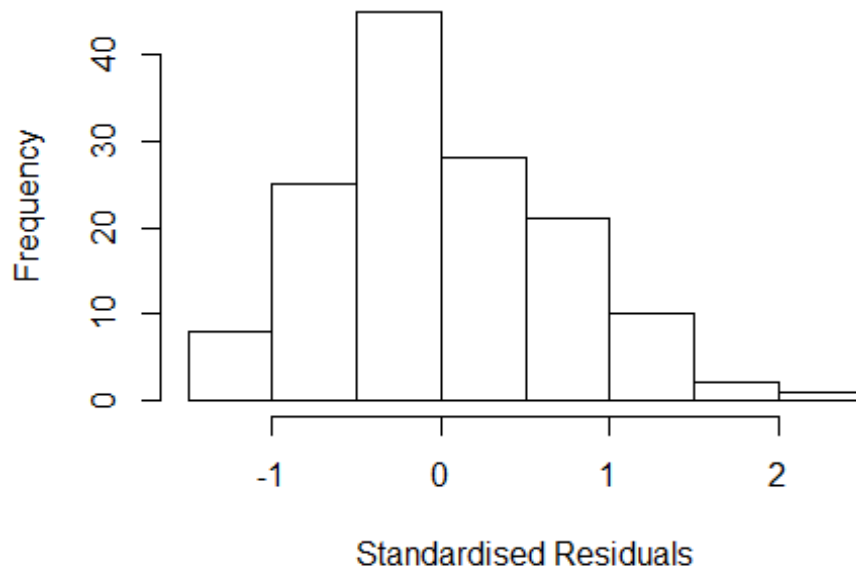
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4535 -0.4637 -0.0514 0.4816 2.1413
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2839 0.3378 3.80 0.00023 ***
## FirstAuthorFemale1 0.2916 0.1887 1.55 0.12484
## Year1997 -0.3176 0.4410 -0.72 0.47285
## Year1998 -0.6012 0.4539 -1.32 0.18780
## Year1999 -0.6137 0.4249 -1.44 0.15115
## Year2000 -0.4224 0.3903 -1.08 0.28138
## Year2001 0.1200 0.4345 0.28 0.78295
## Year2002 0.1511 0.4050 0.37 0.70977
## Year2003 0.2293 0.3660 0.63 0.53210
## Year2004 0.9043 0.4407 2.05 0.04232 *
## Year2005 0.1017 0.4571 0.22 0.82431
## Year2006 0.1696 0.6565 0.26 0.79659
```

```

## Year2007          0.0726      0.3895      0.19  0.85252
## Year2008         -0.3228      0.3801     -0.85  0.39731
## Year2009         -0.4970      0.3930     -1.26  0.20841
## Year2010         -0.2102      0.3917     -0.54  0.59247
## Year2011         -0.0095      0.3867     -0.02  0.98045
## Year2012          0.1124      0.3903      0.29  0.77382
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.775
## Multiple R-squared:  0.166, Adjusted R-squared:  0.0502
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 127 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.425  0.902   0.957   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      7.14e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 2.883 1          1.698
## Year            2.883 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
## -1.441 -0.483 -0.054 0.511 2.114
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2840 0.3369 3.81 0.00022 ***
## LastAuthorFemale1 0.0286 0.2936 0.10 0.92261
## Year1997 -0.3246 0.4517 -0.72 0.47376
## Year1998 -0.6031 0.4540 -1.33 0.18648
## Year1999 -0.6132 0.4238 -1.45 0.15049
## Year2000 -0.3262 0.3874 -0.84 0.40147
## Year2001 0.1198 0.4336 0.28 0.78271
## Year2002 0.1510 0.4042 0.37 0.70932
## Year2003 0.3091 0.4166 0.74 0.45953
## Year2004 1.0500 0.5151 2.04 0.04365 *
## Year2005 0.0944 0.4688 0.20 0.84079
## Year2006 0.1571 0.6866 0.23 0.81943
```

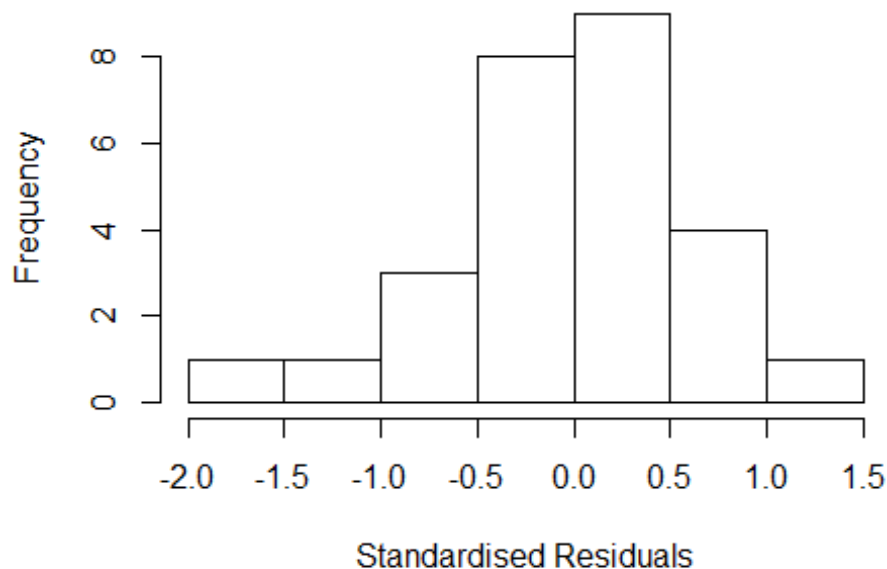
```

## Year2007          0.0675      0.3916      0.17  0.86338
## Year2008          -0.2875      0.3749     -0.77  0.44460
## Year2009          -0.4395      0.3893     -1.13  0.26116
## Year2010          -0.1827      0.3945     -0.46  0.64413
## Year2011           0.0850      0.3804      0.22  0.82358
## Year2012           0.1502      0.3872      0.39  0.69872
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.784
## Multiple R-squared:  0.15,   Adjusted R-squared:  0.032
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 131 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.447  0.899  0.958  0.928  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      7.14e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 140"
## [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]"
##
## 1997 1998 1999 2000 2001 2002 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    4    2    2    1    2    3    3    1    5    5    6    1    7    7    8
##
## 1997 1998 1999 2000 2001 2002 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    0    0    2    0    2    0    2    0    3    3    4    1    5    7    5
##
## 1997 1998 1999 2000 2001 2002 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    0    0    2    0    2    0    1    0    2    3    3    0    4    6    4

```

```
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.71080601082953"
## [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 100.68 1          10.034
## LastAuthorFemale  41.64 1           6.453
## Year              362.82 8           1.445
```

Residuals from first and last author



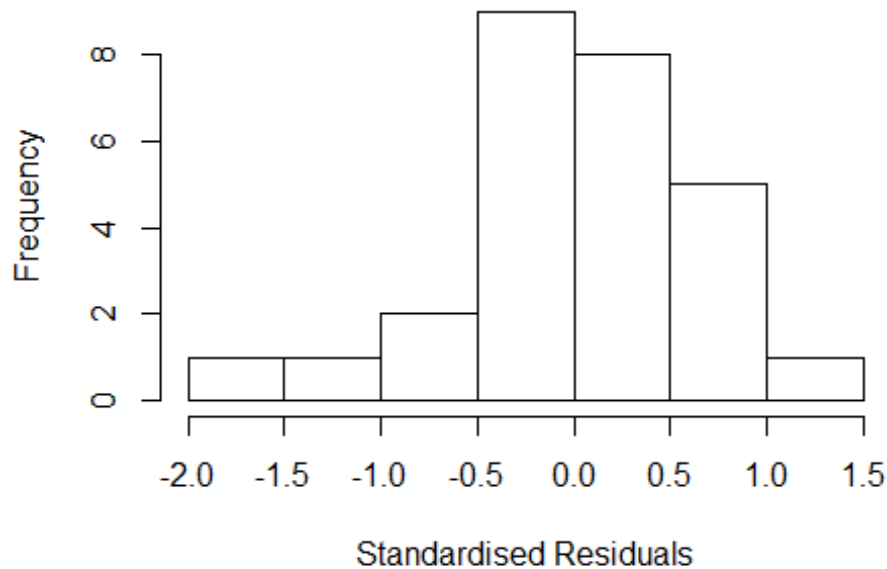
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5974 -0.3452  0.0099  0.4568  1.0708
##
## Coefficients:
```

```

##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.90365    0.58997   1.53   0.15
## FirstAuthorFemale1 -0.09634    0.30291  -0.32   0.75
## LastAuthorFemale1 -0.26462    0.38559  -0.69   0.50
## Year2001          0.27835    0.71970   0.39   0.70
## Year2003           0.08935    0.58997   0.15   0.88
## Year2006          -0.07365    0.59349  -0.12   0.90
## Year2007           0.00341    0.63603   0.01   1.00
## Year2008           0.19579    0.45968   0.43   0.68
## Year2010           0.60353    0.64044   0.94   0.36
## Year2011           0.69372    0.60481   1.15   0.27
## Year2012           0.44806    0.69385   0.65   0.53
##
## Robust residual standard error: 0.657
## Multiple R-squared:  0.225, Adjusted R-squared:  -0.259
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 21 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.534  0.939   0.950   0.914   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      3.70e-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 6.61 1          2.571
## Year              6.61 8          1.125

```


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
## -1.5318 -0.3709 0.0423 0.4812 1.0451
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.829771 0.658605 1.26 0.22
## FirstAuthorFemale1 -0.154771 0.279210 -0.55 0.59
## Year2001 0.352229 0.753125 0.47 0.65
## Year2003 0.163229 0.658605 0.25 0.81
## Year2006 0.000229 0.661760 0.00 1.00
## Year2007 0.076920 0.704052 0.11 0.91
## Year2008 0.220799 0.599117 0.37 0.72
## Year2010 0.703150 0.687736 1.02 0.32
## Year2011 0.702028 0.676682 1.04 0.31
## Year2012 0.455436 0.723820 0.63 0.54
##
## Robust residual standard error: 0.702
```

```

## Multiple R-squared:  0.194, Adjusted R-squared:  -0.232
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 2 weights are ~= 1. The remaining 25 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.614  0.924  0.959   0.936   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           3.70e-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"
##           GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 24.89 1         4.989
## Year              24.89 8         1.223

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ 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 -2.91e-01 -3.05e-16  4.25e-01  1.13e+00
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.8405     0.4790   1.75   0.097 .
## LastAuthorFemale1 -0.3310     0.4451  -0.74   0.467
## Year2001          0.3415     0.6391   0.53   0.600
## Year2003          0.1525     0.4790   0.32   0.754
## Year2006         -0.0105     0.4833  -0.02   0.983
## Year2007          0.0675     0.5303   0.13   0.900
## Year2008          0.2168     0.4314   0.50   0.622

```

```

## Year2010          0.6121      0.6503      0.94      0.360
## Year2011          0.8042      0.5851      1.37      0.187
## Year2012          0.5472      0.6085      0.90      0.381
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.571
## Multiple R-squared:  0.257, Adjusted R-squared:  -0.136
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 22 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.387  0.910  0.938  0.894  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      3.70e-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: 27"
## [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 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    4    2    4    5    2    1    3    2    2    6    9    2    9    6    4
## 2012
##    6
##
## 1996 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    2    1    2    3    0    0    2    1    1    2    8    1    5    3    2
## 2012
##    4
##
## 1996 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011

```

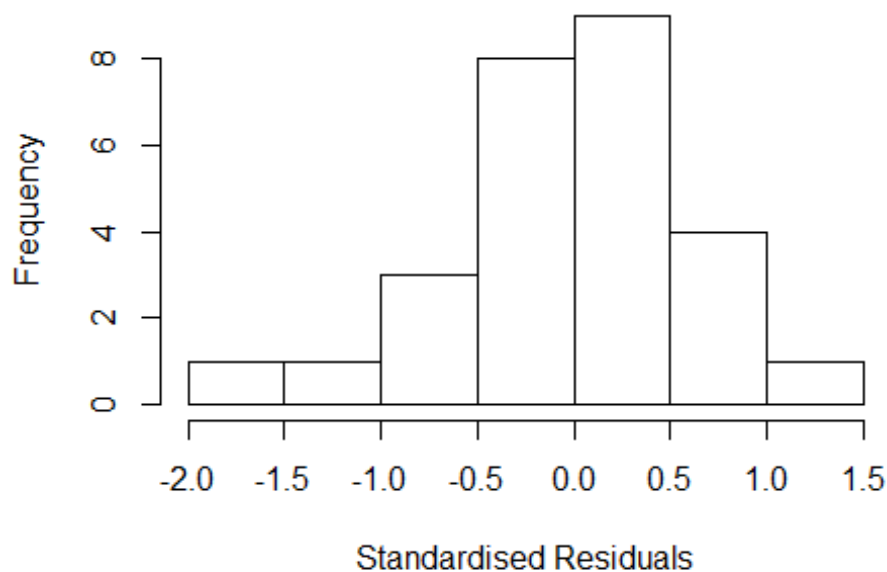
```
##      2      0      2      3      0      0      2      0      0      2      8      1      5      3      2
## 2012
##      4
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3"
## [1] "Male first author team size 2018 geometric mean: 2.35215804504935"

## Warning in wilcox.test.default(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 = 0.8
## 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: 2.44948974278318"
## [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
```

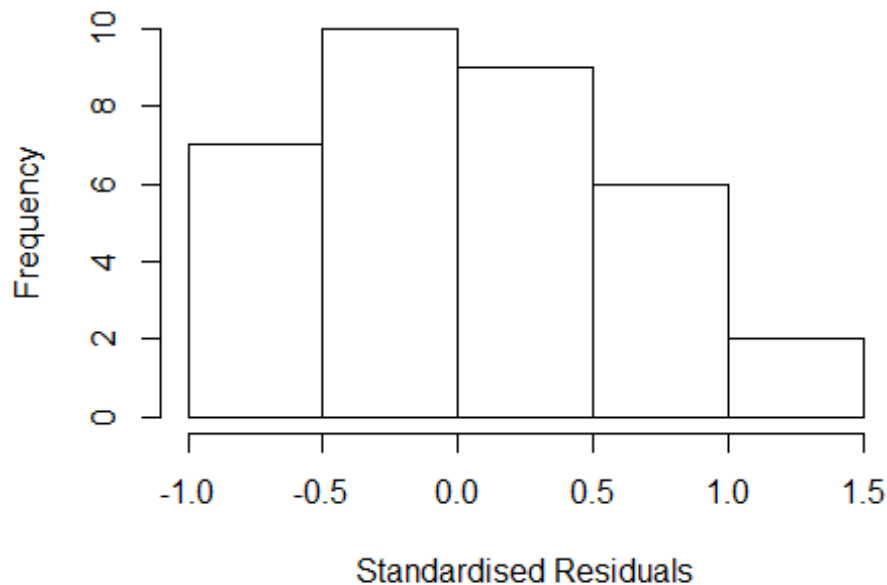
Residuals from last author



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

```
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.737e+15  1      4.168e+07
## LastAuthorFemale  6.533e+02  1      2.556e+01
## Year              8.308e+17 10      7.870e+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
## -0.9102 -0.3108  0.0145  0.2179  1.2264
```

```
##
```

```
## Coefficients:
```

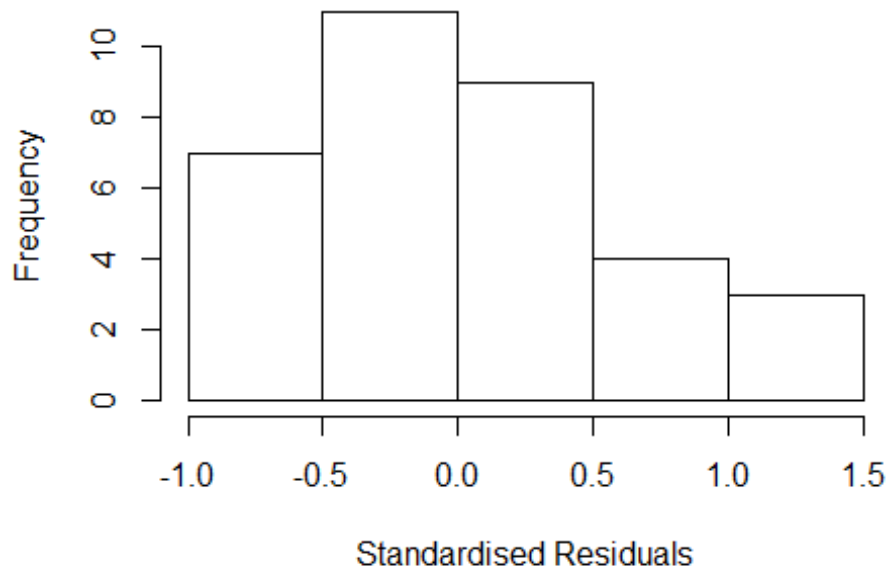
```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.8985     0.0656   13.70  6.1e-12 ***
## FirstAuthorFemale1 -0.2951     0.3325    -0.89   0.3850
## LastAuthorFemale1  0.4274     0.3988     1.07   0.2961
## Year1998          -1.0308     0.4340    -2.37   0.0272 *
## Year2000           0.8390     0.3077     2.73   0.0126 *
```

```

## Year2003          0.5185      0.0687      7.55  2.1e-07 ***
## Year2006          0.2070      0.6306      0.33  0.7460
## Year2007          0.0117      0.3087      0.04  0.9701
## Year2008         -0.2519      0.4042     -0.62  0.5399
## Year2009         -0.0616      0.4204     -0.15  0.8849
## Year2010         -0.3430      0.2932     -1.17  0.2552
## Year2011          0.2510      0.0761      3.30  0.0034 **
## Year2012          0.2021      0.4871      0.41  0.6825
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.79
## Multiple R-squared:  0.3,   Adjusted R-squared:  -0.1
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 24 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.792  0.904   0.946   0.937   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      2.94e-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"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.714e+15  1      4.140e+07
## Year              1.714e+15 10      5.777e+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
## -9.86e-01 -2.32e-01 -2.22e-16 8.63e-02 1.27e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8985 0.0656 13.70 3.0e-12 ***
## FirstAuthorFemale1 -0.1891 0.3216 -0.59 0.5625
## Year1998 -0.7094 0.3282 -2.16 0.0418 *
## Year2000 0.8397 0.3107 2.70 0.0130 *
## Year2003 0.5185 0.0687 7.54 1.5e-07 ***
## Year2006 0.2070 0.6512 0.32 0.7536
## Year2007 0.0880 0.3476 0.25 0.8025
## Year2008 0.1755 0.0656 2.68 0.0138 *
## Year2009 -0.0653 0.4335 -0.15 0.8817
## Year2010 -0.1967 0.1995 -0.99 0.3349
## Year2011 0.2510 0.0761 3.30 0.0033 **
## Year2012 0.1579 0.4900 0.32 0.7504
```

```

## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.736
## Multiple R-squared:  0.283, Adjusted R-squared:  -0.0756
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 20 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.747  0.876  0.914  0.911  0.972  0.999
## Algorithmic parameters:
##           tuning.chi           bb           tuning.psi           refine.tol
##           1.55e+00           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 1270  1           35.64
## Year              1270 10           1.43
##
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##           Min           1Q       Median           3Q          Max
## -8.44e-01 -3.04e-01 -2.22e-16  1.91e-01  1.33e+00
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.8985     0.0656   13.70 3.0e-12 ***
## LastAuthorFemale1  0.3529     0.4433    0.80  0.4345
## Year1998          -1.2514     0.4481   -2.79  0.0106 *

```



```

## Year2000          0.8394      0.3096      2.71      0.0127 *
## Year2003          0.5185      0.0687      7.55      1.5e-07 ***
## Year2006          0.2070      0.6437      0.32      0.7508
## Year2007         -0.0546      0.2648     -0.21      0.8386
## Year2008         -0.1774      0.4481     -0.40      0.6960
## Year2009         -0.0639      0.4162     -0.15      0.8793
## Year2010         -0.3184      0.2940     -1.08      0.2905
## Year2011          0.2510      0.0761      3.30      0.0033 **
## Year2012          0.0994      0.4193      0.24      0.8147
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.754
## Multiple R-squared:  0.291, Adjusted R-squared:  -0.0628
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~ 1. The remaining 24 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.737  0.891  0.943   0.930   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          2.94e-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 2203"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 2000 2004 2006 2007 2008 2009 2010 2011 2012
##    1    1    1    3    2    4    1    5    3
##
## 2000 2004 2006 2007 2008 2009 2010 2011 2012
##    1    1    1    2    0    0    1    4    2

```

```

##
## 2000 2004 2006 2007 2008 2009 2010 2011 2012
##    1    1    1    2    0    0    1    4    2
## [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: 4.89897948556636"
##
## Wilcoxon rank sum test
##
## 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: 1.41421356237309"
## [1] "Male last author team size 2018 geometric mean: 4.16016764610381"
##
## Wilcoxon rank sum test
##
## 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"
## [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 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
##    7    4    8    7    7    8    9   13   14   18   26   40   35   61   69
## 2011 2012
##   58   81
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    3    4    0    2    1    2    2    6   11    9   22   21   37   45
## 2011 2012
##   40   51
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    3    3    0    2    0    2    2    4   10    8   20   16   29   34

```

```
## 2011 2012
## 32 42
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.36521391433382"
## [1] "Male first author team size 2018 geometric mean: 3.03936708283482"

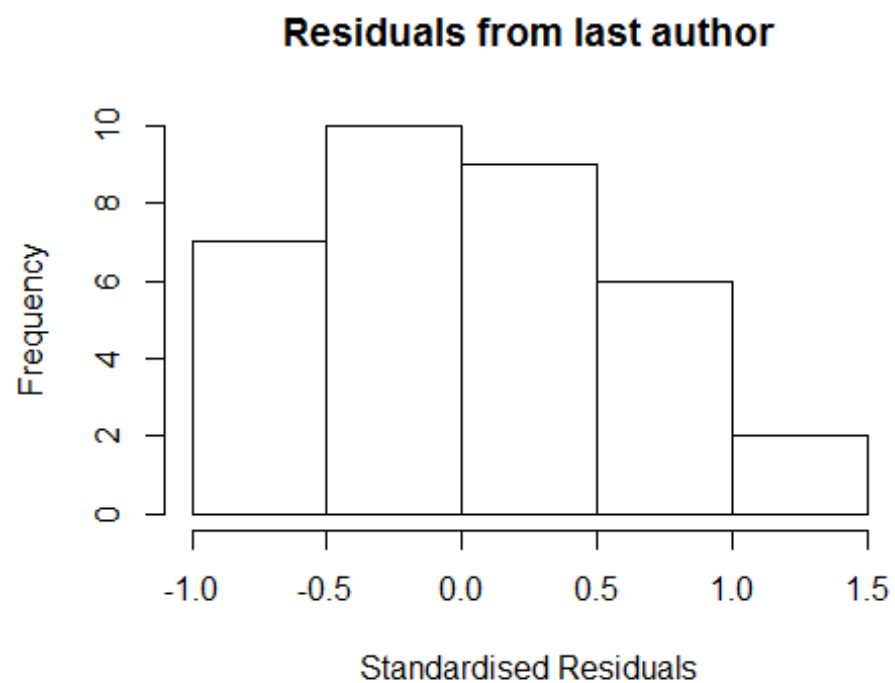
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 130, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.96329033279254"
## [1] "Male last author team size 2018 geometric mean: 2.98727580893905"

## Warning in wilcox.test.default(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 = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"

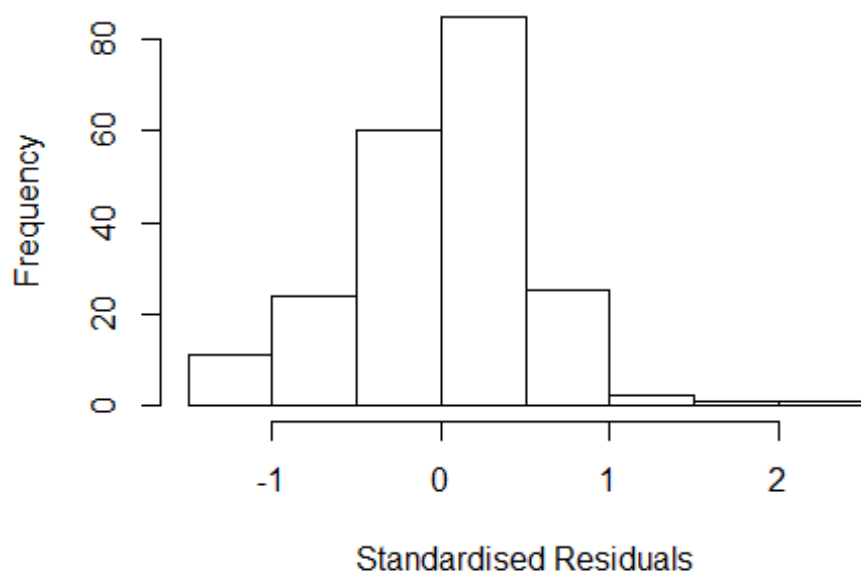
## Warning in lmrob.fit(x, y, control, init = init, mf = mf): M-step did NOT
## converge. Returning unconverged SM-estimate
```



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

| ## | | GVIF | Df | $GVIF^{1/(2 \cdot Df)}$ |
|----|-------------------|-------|----|-------------------------|
| ## | FirstAuthorFemale | 1.383 | 1 | 1.176 |
| ## | LastAuthorFemale | 1.296 | 1 | 1.138 |
| ## | Year | 1.645 | 14 | 1.018 |

Residuals from first and last author



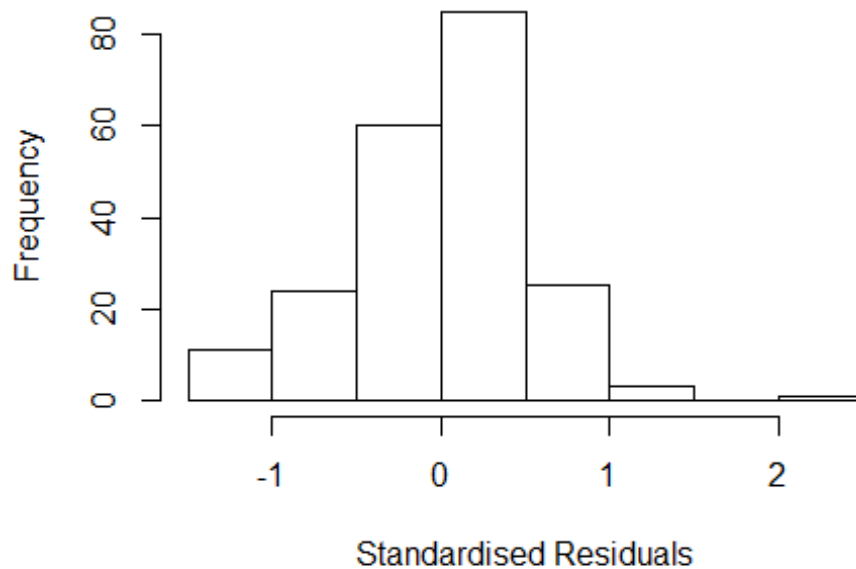
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2526 -0.2965 0.0516 0.3092 2.0945
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0835 0.4659 2.33 0.021 *
## FirstAuthorFemale1 -0.0084 0.0907 -0.09 0.926
## LastAuthorFemale1 0.0383 0.1123 0.34 0.734
## Year1997 0.1177 0.7636 0.15 0.878
## Year1998 -0.6100 0.4684 -1.30 0.194
## Year2000 -0.2213 0.4857 -0.46 0.649
## Year2002 0.3099 0.5090 0.61 0.543
## Year2003 0.6897 0.4808 1.43 0.153
## Year2004 0.1266 0.6277 0.20 0.840
## Year2005 0.1344 0.4850 0.28 0.782
## Year2006 0.1748 0.4859 0.36 0.719
## Year2007 0.0282 0.5006 0.06 0.955
```

```

## Year2008          0.2766      0.4789      0.58      0.564
## Year2009          0.1333      0.4721      0.28      0.778
## Year2010          0.1592      0.4733      0.34      0.737
## Year2011          0.1691      0.4733      0.36      0.721
## Year2012          0.0989      0.4749      0.21      0.835
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.477
## Multiple R-squared:  0.0684, Adjusted R-squared:  -0.00928
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 190 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0146 0.8670 0.9550 0.8940 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      4.78e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.304 1          1.142
## Year              1.304 14          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
## -1.2586 -0.2989 0.0455 0.3069 2.0901
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.08350 0.47013 2.30 0.022 *
## FirstAuthorFemale1 -0.00463 0.08852 -0.05 0.958
## Year1997 0.12759 0.77937 0.16 0.870
## Year1998 -0.61125 0.47246 -1.29 0.197
## Year2000 -0.22318 0.48926 -0.46 0.649
## Year2002 0.32900 0.51659 0.64 0.525
## Year2003 0.68782 0.48506 1.42 0.158
## Year2004 0.12196 0.63329 0.19 0.847
## Year2005 0.13304 0.48918 0.27 0.786
## Year2006 0.17300 0.48994 0.35 0.724
## Year2007 0.02927 0.50513 0.06 0.954
## Year2008 0.28286 0.48286 0.59 0.559
```

```

## Year2009          0.13562    0.47617    0.28    0.776
## Year2010          0.16242    0.47751    0.34    0.734
## Year2011          0.17507    0.47696    0.37    0.714
## Year2012          0.10340    0.47874    0.22    0.829
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.47
## Multiple R-squared:  0.0686, Adjusted R-squared:  -0.00374
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 189 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0098 0.8640 0.9550 0.8910 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      4.78e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.224 1          1.106
## Year            1.224 14          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
## -1.2509 -0.3021  0.0449  0.3056  2.0953
##
## Coefficients:

```



```

##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0835    0.4695   2.31  0.022 *
## LastAuthorFemale1 0.0362    0.1094   0.33  0.741
## Year1997          0.1262    0.7788   0.16  0.871
## Year1998          -0.6128    0.4709  -1.30  0.195
## Year2000          -0.2255    0.4862  -0.46  0.643
## Year2002           0.3109    0.5126   0.61  0.545
## Year2003           0.6855    0.4828   1.42  0.157
## Year2004           0.1196    0.6292   0.19  0.849
## Year2005           0.1316    0.4873   0.27  0.787
## Year2006           0.1713    0.4879   0.35  0.726
## Year2007           0.0287    0.5045   0.06  0.955
## Year2008           0.2739    0.4805   0.57  0.569
## Year2009           0.1316    0.4755   0.28  0.782
## Year2010           0.1598    0.4769   0.34  0.738
## Year2011           0.1674    0.4765   0.35  0.726
## Year2012           0.0982    0.4782   0.21  0.838
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.471
## Multiple R-squared:  0.0689, Adjusted R-squared:  -0.00348
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 189 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0096 0.8630 0.9550 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           4.78e-04           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01           5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500           50           2           1           1000           200
## trace.lev      mts      compute.rd
##           0           1000           0
##           psi           subsampling           cov
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 209"
## [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
##    4    5    2    6    7   14    4    6    7   15   13   18   18   22   26
## 2011 2012
##   22   29
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    4    1    2    4    6    2    5    1    5    6    2   11   15   12
## 2011 2012
##    9   14
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    4    1    1    4    4    1    4    1    3    3    2    9   15   11
## 2011 2012
##    8   11
## [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: 2.37800092589143"

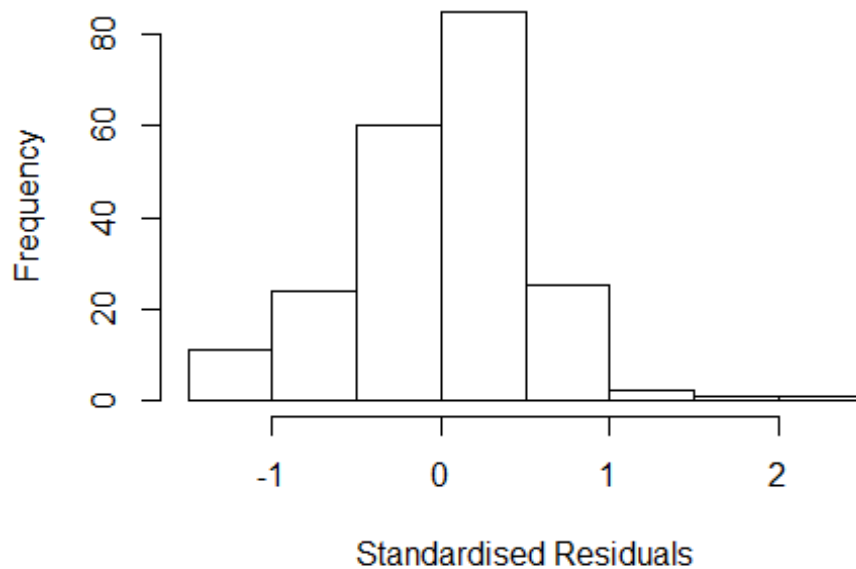
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 24, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.05976714390712"
## [1] "Male last author team size 2018 geometric mean: 2.34877748309538"

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

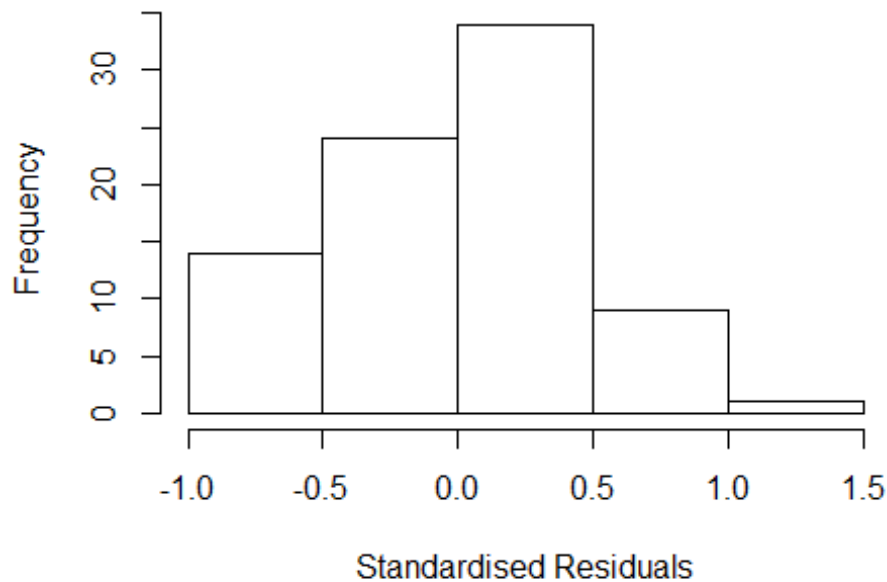
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 23, 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 6.549e+00 1          2.559
## LastAuthorFemale  6.486e+00 1          2.547
## UniqueAuthors    -2.877e+30 4           NaN
## Year              -1.507e+31 15          NaN
```

Residuals from first and last author and team size



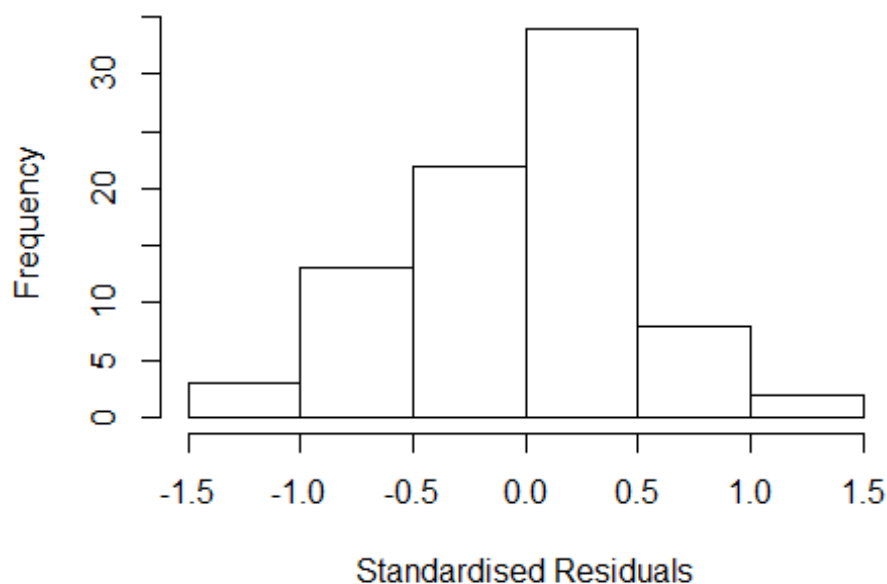
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9944 -0.3087 0.0518 0.2871 1.2976
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.81036 0.24030 3.37 0.00131 **
## FirstAuthorFemale1 -0.18093 0.17007 -1.06 0.29163
## LastAuthorFemale1 0.57089 0.19587 2.91 0.00500 **
## UniqueAuthors2 0.30671 0.18523 1.66 0.10298
## UniqueAuthors3 0.47118 0.17251 2.73 0.00827 **
## UniqueAuthors4 0.60884 0.22877 2.66 0.00997 **
## UniqueAuthors5 0.68405 0.36358 1.88 0.06476 .
## Year1998 -1.11706 0.28290 -3.95 0.00021 ***
## Year1999 0.12864 0.24030 0.54 0.59440
## Year2000 0.04944 0.34170 0.14 0.88544
```

```

## Year2001          0.52059      0.40489      1.29  0.20347
## Year2002         -0.75906      0.28290     -2.68  0.00941 **
## Year2003          0.03740      0.43161      0.09  0.93124
## Year2004         -0.63406      0.28290     -2.24  0.02872 *
## Year2005          0.56776      0.28036      2.03  0.04731 *
## Year2006         -0.33723      0.31452     -1.07  0.28792
## Year2007         -0.25545      0.55349     -0.46  0.64609
## Year2008          0.00686      0.30130      0.02  0.98191
## Year2009         -0.03770      0.29417     -0.13  0.89845
## Year2010          0.32322      0.30263      1.07  0.28979
## Year2011          0.14641      0.27778      0.53  0.60007
## Year2012          0.21770      0.29835      0.73  0.46842
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.512
## Multiple R-squared:  0.377, Adjusted R-squared:  0.159
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 71 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.501  0.872   0.965   0.915   0.985   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.22e-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 2.708 1      1.646
## LastAuthorFemale  3.767 1      1.941
## Year              3.193 15      1.039

```

Residuals from first and last author



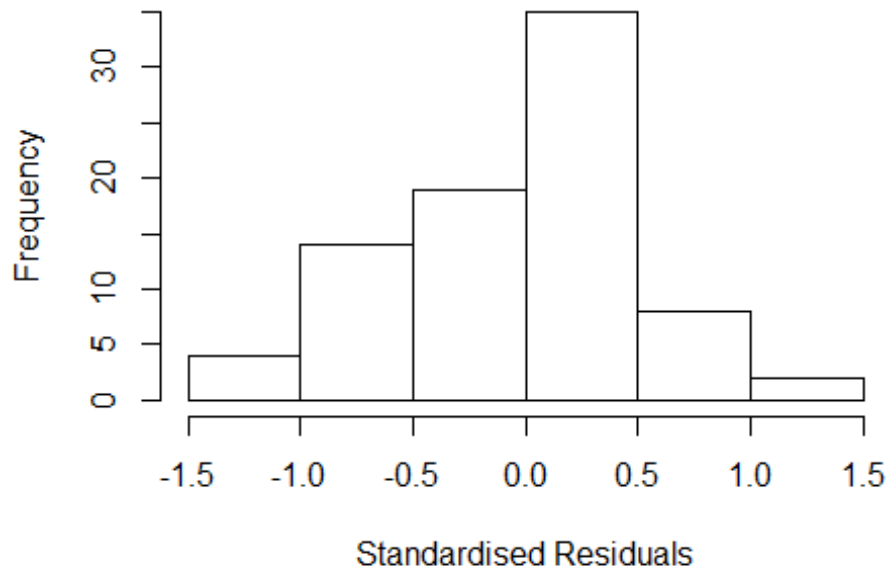
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1707 -0.2281 0.0508 0.2604 1.2599
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0384 0.2204 4.71 1.4e-05 ***
## FirstAuthorFemale1 -0.2870 0.1880 -1.53 0.132
## LastAuthorFemale1 0.5208 0.2200 2.37 0.021 *
## Year1998 -1.0384 0.2204 -4.71 1.4e-05 ***
## Year1999 -0.0994 0.2204 -0.45 0.653
## Year2000 0.1156 0.4830 0.24 0.812
## Year2001 0.8223 0.3590 2.29 0.025 *
## Year2002 -0.6804 0.2204 -3.09 0.003 **
## Year2003 0.1545 0.3864 0.40 0.691
## Year2004 -0.5554 0.2204 -2.52 0.014 *
## Year2005 0.5958 0.2303 2.59 0.012 *
## Year2006 -0.2978 0.2364 -1.26 0.212
```

```

## Year2007          -0.2479      1.5392    -0.16     0.873
## Year2008           0.1784      0.3011     0.59     0.556
## Year2009           0.0787      0.2777     0.28     0.778
## Year2010           0.4969      0.2473     2.01     0.049 *
## Year2011           0.2516      0.2738     0.92     0.362
## Year2012           0.3409      0.2758     1.24     0.221
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.457
## Multiple R-squared:  0.327, Adjusted R-squared:  0.149
## Convergence in 30 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 72 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.427  0.838  0.939   0.883  0.985   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          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 1.268 1          1.126
## Year              1.268 15          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
## -1.1154 -0.3422 0.0316 0.3224 1.1395
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1109 0.2316 4.80 9.8e-06 ***
## FirstAuthorFemale1 -0.0580 0.1631 -0.36 0.7234
## Year1998 -1.1109 0.2316 -4.80 9.8e-06 ***
## Year1999 -0.1719 0.2316 -0.74 0.4606
## Year2000 -0.0695 0.4096 -0.17 0.8658
## Year2001 0.6944 0.3784 1.84 0.0711 .
## Year2002 -0.7529 0.2316 -3.25 0.0018 **
## Year2003 0.2869 0.3472 0.83 0.4117
## Year2004 -0.6279 0.2316 -2.71 0.0086 **
## Year2005 0.5233 0.2411 2.17 0.0336 *
## Year2006 -0.3701 0.2467 -1.50 0.1385
## Year2007 -0.3204 0.9827 -0.33 0.7454
```

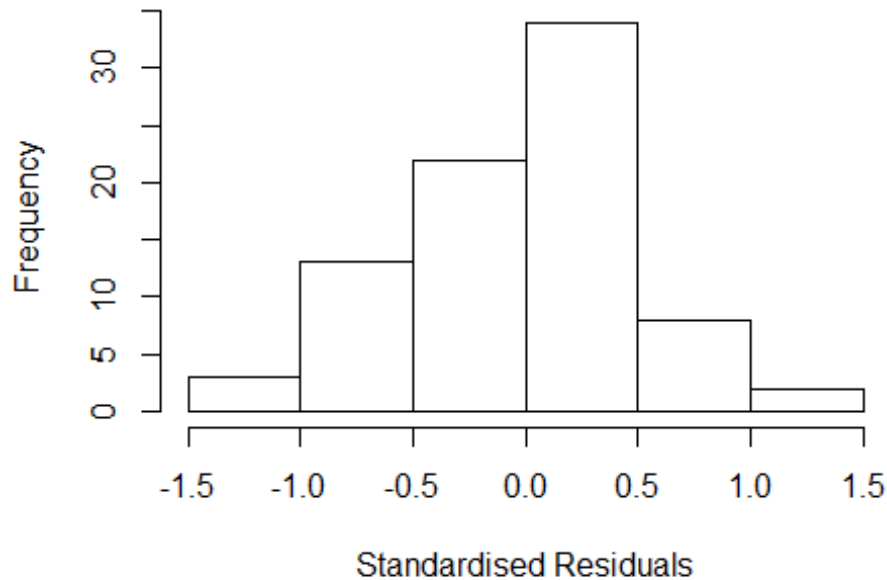


```

## Year2008          0.0518      0.2874      0.18      0.8576
## Year2009          0.1525      0.2924      0.52      0.6037
## Year2010          0.4229      0.2636      1.60      0.1135
## Year2011          0.2110      0.3091      0.68      0.4973
## Year2012          0.2176      0.2907      0.75      0.4569
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.55
## Multiple R-squared:  0.248, Adjusted R-squared:  0.0632
## 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.
##  0.647  0.857   0.951   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      1.22e-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 1.471  1          1.213
## Year            1.471 15          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
## -1.1411 -0.3565  0.0414  0.3033  1.2454
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9966    0.2141   4.66 1.6e-05 ***
## LastAuthorFemale1 0.3619    0.1879   1.93  0.059 .
## Year1998      -0.9966    0.2141  -4.66 1.6e-05 ***
## Year1999      -0.0576    0.2141  -0.27  0.789
## Year2000       0.0462    0.4055   0.11  0.910
## Year2001       0.8346    0.3679   2.27  0.027 *
## Year2002      -0.6386    0.2141  -2.98  0.004 **
## Year2003       0.2487    0.3544   0.70  0.485
## Year2004      -0.5136    0.2141  -2.40  0.019 *
## Year2005       0.6376    0.2243   2.84  0.006 **
## Year2006      -0.2559    0.2303  -1.11  0.271
## Year2007      -0.2061    1.1819  -0.17  0.862
```

```

## Year2008          0.1108      0.3060      0.36      0.719
## Year2009          0.1351      0.2624      0.51      0.608
## Year2010          0.4889      0.2446      2.00      0.050 *
## Year2011          0.3031      0.2781      1.09      0.280
## Year2012          0.3200      0.2895      1.11      0.273
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.498
## Multiple R-squared:  0.29,   Adjusted R-squared:  0.116
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 70 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.511  0.860  0.943  0.896  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.22e-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 2206"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    2    3    2    2    2    1    2    5    6    5    4    5    3    2
## 2011 2012
##    2    1
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    0    0    0    0    2    0    0    3    3    5    2    2    3    1
## 2011 2012
##    1    0

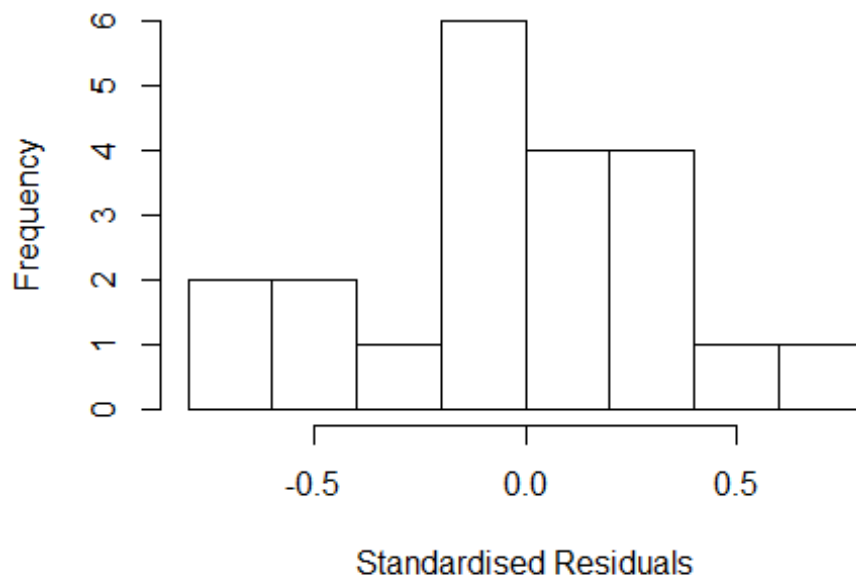
```

```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    0    0    0    0    2    0    0    3    3    4    2    2    3    1
## 2011 2012
##    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: 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 34.59  1          5.881
## Year              34.59  8          1.248

## [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
```

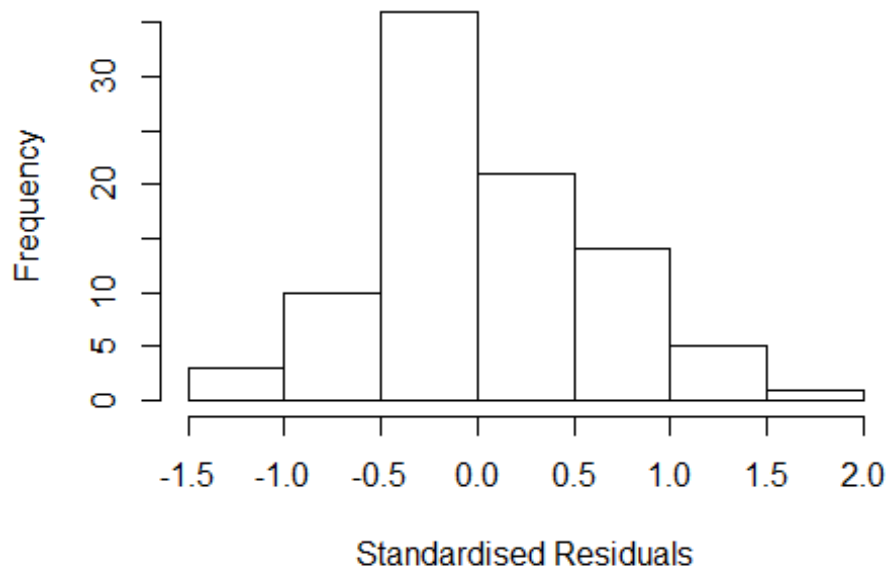
Residuals from first author



```
## [1] "Sample size for the above analysis: 21"
## [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
##    5    3   10   11    3    6   12   10    8    9   12   20   12   18    7
## 2011 2012
##   14   20
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    9    5    2    2    7    7    5    3    7   13    6   10    4
## 2011 2012
##   11   16
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    4    4    2    0    6    6    4    2    6   11    5    7    3
## 2011 2012
##   11   16
## [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: 2.70192007704123"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale  7.020  1          2.649
## LastAuthorFemale   4.924  1          2.219
## UniqueAuthors     136.388  4          1.849
## Year               258.885 15          1.203
```

Residuals from first and last author and team size



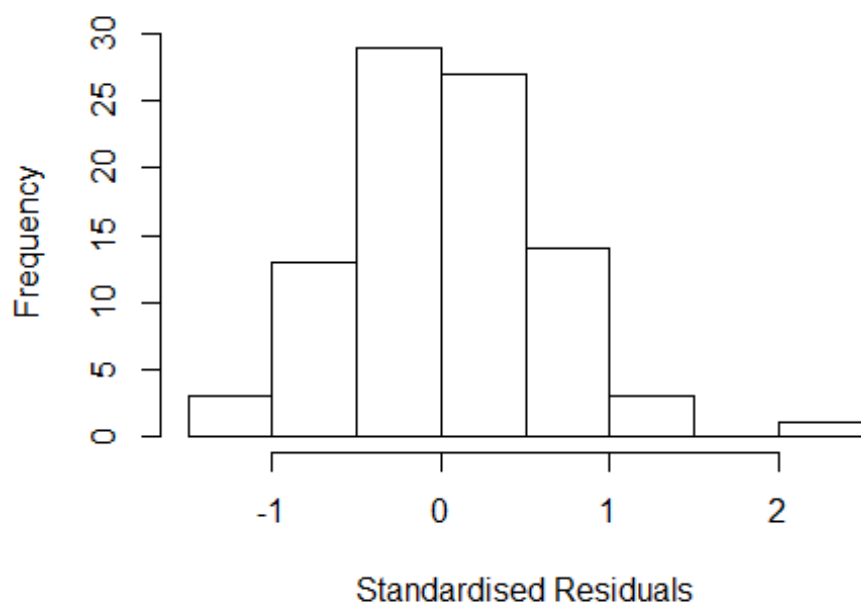
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2114 -0.4034 -0.0721 0.4555 1.5110
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2179 0.2407 5.06 3.4e-06 ***
## FirstAuthorFemale1 -0.4406 0.3753 -1.17 0.245
## LastAuthorFemale1 -0.0456 0.4119 -0.11 0.912
## UniqueAuthors2 0.4276 0.2627 1.63 0.108
## UniqueAuthors3 0.2073 0.2092 0.99 0.325
## UniqueAuthors4 0.2874 0.2091 1.37 0.174
## UniqueAuthors5 0.3538 0.2947 1.20 0.234
## Year1997 0.4226 0.3801 1.11 0.270
## Year1998 -0.8144 0.4761 -1.71 0.092 .
## Year1999 0.1484 0.3844 0.39 0.701
```

```

## Year2000          0.2853      0.3364      0.85      0.399
## Year2002         -0.5240      0.3185     -1.65      0.105
## Year2003         -0.4366      0.3356     -1.30      0.198
## Year2004         -0.3066      0.4784     -0.64      0.524
## Year2005         -0.2160      0.7921     -0.27      0.786
## Year2006         -0.1934      0.3020     -0.64      0.524
## Year2007         -0.3535      0.3132     -1.13      0.263
## Year2008         -0.7186      0.3995     -1.80      0.077 .
## Year2009         -0.1267      0.4163     -0.30      0.762
## Year2010         -0.0755      0.3823     -0.20      0.844
## Year2011         -0.5467      0.3535     -1.55      0.127
## Year2012         -0.4341      0.2916     -1.49      0.141
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.636
## Multiple R-squared:  0.232, Adjusted R-squared:  -0.00516
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 82 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.552  0.901  0.954  0.924  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          1.11e-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 13.928 1      3.732
## LastAuthorFemale  2.269 1      1.506
## Year              15.335 15      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
## -1.06788 -0.39646 0.00107 0.42714 2.00649
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3202 0.1673 7.89 2.5e-11 ***
## FirstAuthorFemale1 -0.2914 0.3210 -0.91 0.367
## LastAuthorFemale1 -0.1796 0.3781 -0.48 0.636
## Year1997 0.7478 0.1673 4.47 2.9e-05 ***
## Year1998 -0.9847 0.4214 -2.34 0.022 *
## Year1999 0.2559 0.2672 0.96 0.342
## Year2000 0.3968 0.1681 2.36 0.021 *
## Year2002 -0.4365 0.2902 -1.50 0.137
## Year2003 -0.3835 0.2405 -1.59 0.115
## Year2004 -0.0308 0.3861 -0.08 0.937
## Year2005 -0.1455 0.9880 -0.15 0.883
## Year2006 -0.2227 0.2433 -0.92 0.363
```

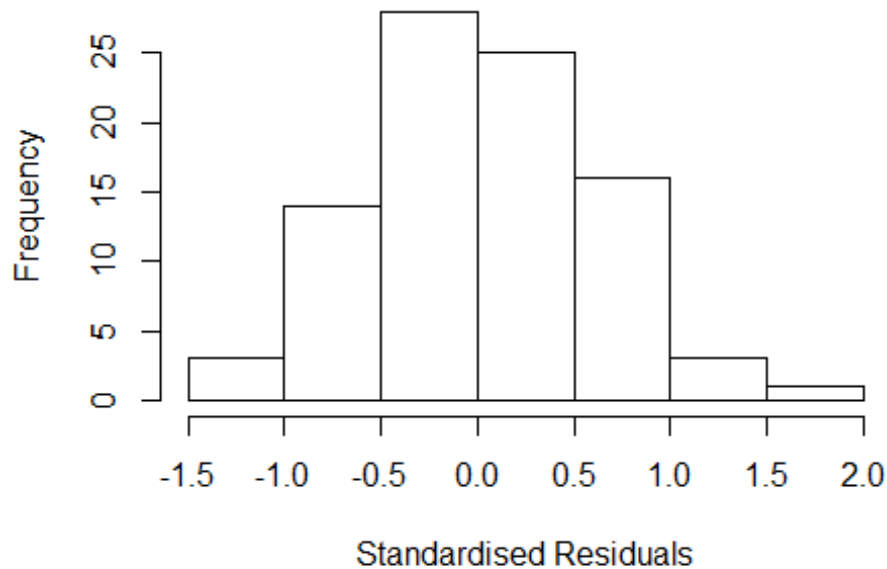


```

## Year2007          -0.2713      0.2480    -1.09     0.278
## Year2008          -0.5133      0.3504    -1.46     0.147
## Year2009           0.0537      0.3604     0.15     0.882
## Year2010           0.1048      0.2511     0.42     0.678
## Year2011          -0.3883      0.2510    -1.55     0.126
## Year2012          -0.2606      0.2028    -1.28     0.203
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.65
## Multiple R-squared:  0.206, Adjusted R-squared:  0.0184
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 76 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.321  0.884   0.951   0.917   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      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 26.72 1          5.169
## Year              26.72 15          1.116

```

Residuals from first author

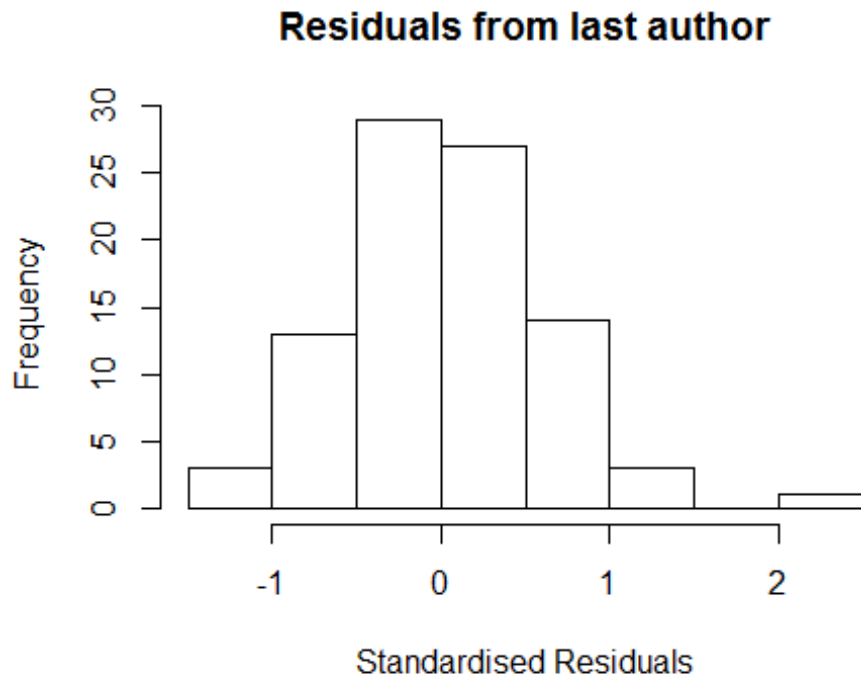


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.06793 -0.39303 0.00013 0.42068 1.99710
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3489 0.1386 9.73 8.0e-15 ***
## FirstAuthorFemale1 -0.3489 0.2719 -1.28 0.204
## Year1997 0.7191 0.1386 5.19 1.8e-06 ***
## Year1998 -1.0040 0.4379 -2.29 0.025 *
## Year1999 0.2272 0.2502 0.91 0.367
## Year2000 0.3681 0.1395 2.64 0.010 *
## Year2002 -0.4631 0.2820 -1.64 0.105
## Year2003 -0.4312 0.2227 -1.94 0.057 .
## Year2004 -0.1056 0.3128 -0.34 0.737
## Year2005 -0.1455 0.9073 -0.16 0.873
## Year2006 -0.2517 0.2231 -1.13 0.263
## Year2007 -0.2872 0.2464 -1.17 0.248
```

```

## Year2008          -0.5414      0.3385    -1.60      0.114
## Year2009           0.0250      0.3465      0.07      0.943
## Year2010           0.0762      0.2335      0.33      0.745
## Year2011          -0.4137      0.2360     -1.75      0.084 .
## Year2012          -0.2950      0.1707     -1.73      0.088 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.657
## Multiple R-squared:  0.203, Adjusted R-squared:  0.0277
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 76 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.336  0.890  0.951  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      1.11e-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"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.152  1          1.073
## Year            1.152 15          1.005

```



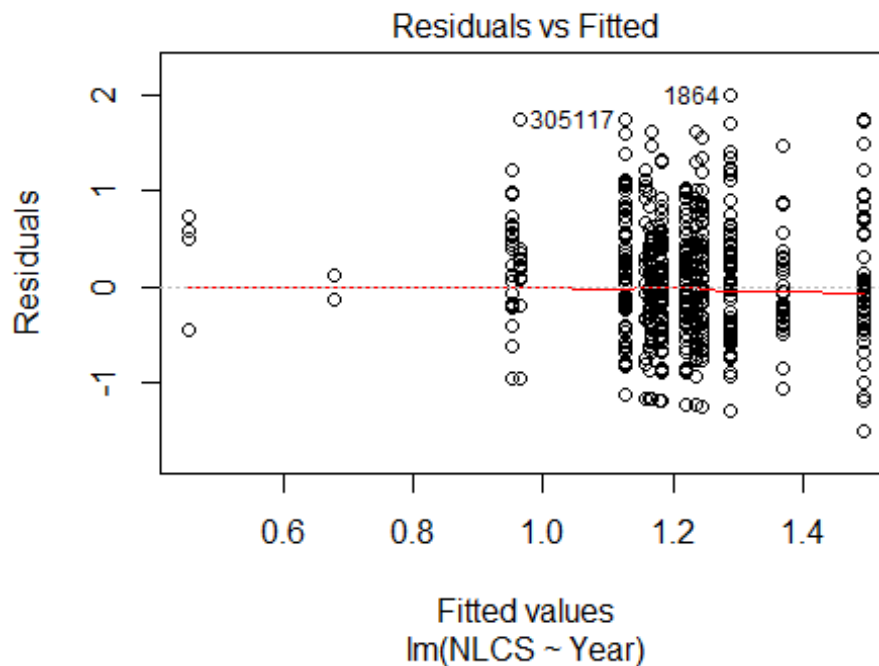
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.068 -0.394 0.015 0.403 2.018
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1745 0.1533 7.66 6.1e-11 ***
## LastAuthorFemale1 -0.3872 0.3414 -1.13 0.26042
## Year1997 0.8935 0.1533 5.83 1.4e-07 ***
## Year1998 -0.8505 0.4132 -2.06 0.04310 *
## Year1999 0.4015 0.2591 1.55 0.12553
## Year2000 0.5425 0.1541 3.52 0.00075 ***
## Year2002 -0.2935 0.2847 -1.03 0.30587
## Year2003 -0.2514 0.2415 -1.04 0.30138
## Year2004 0.1637 0.3800 0.43 0.66798
## Year2005 -0.1455 1.5567 -0.09 0.92579
## Year2006 -0.0767 0.2331 -0.33 0.74308
## Year2007 -0.1876 0.2777 -0.68 0.50144
```

```

## Year2008          -0.3683      0.3465    -1.06   0.29132
## Year2009           0.1993      0.3555      0.56   0.57669
## Year2010           0.2504      0.2426      1.03   0.30547
## Year2011          -0.2651      0.2285     -1.16   0.24971
## Year2012          -0.1412      0.2105     -0.67   0.50456
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.642
## Multiple R-squared:  0.199, Adjusted R-squared:  0.0231
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 81 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.302  0.894  0.956  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      1.11e-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 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
##   33   46   57    2   50   52   59   43   63   87  106  132   82  125  121
## 2011 2012
##  128  106
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7   20   29    2   19   16   28   23   29   50   46   69   41   63   52
## 2011 2012
##   70   76

```

```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6   18   25    2   15   14   23   21   25   41   40   49   36   54   43
## 2011 2012
##   64   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 = 31, df = 16, p-value = 0.01
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.18, df = 1, p-value = 0.7

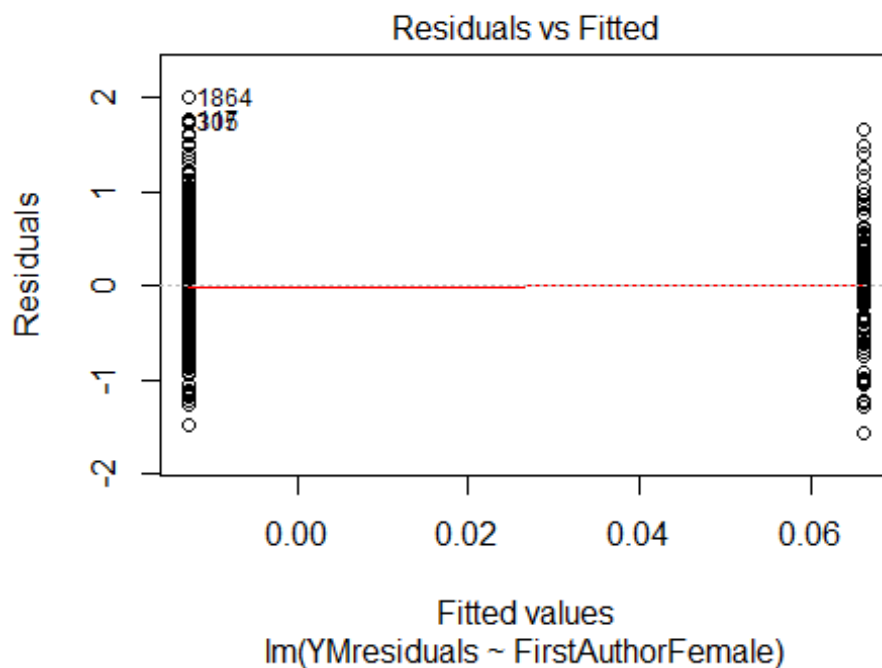
## [1] "Female first author team size 2018 geometric mean: 6"
## [1] "Male first author team size 2018 geometric mean: 3.47742961675988"

## Warning in wilcox.test.default(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 = 0.3
## 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: 3.56667922655813"

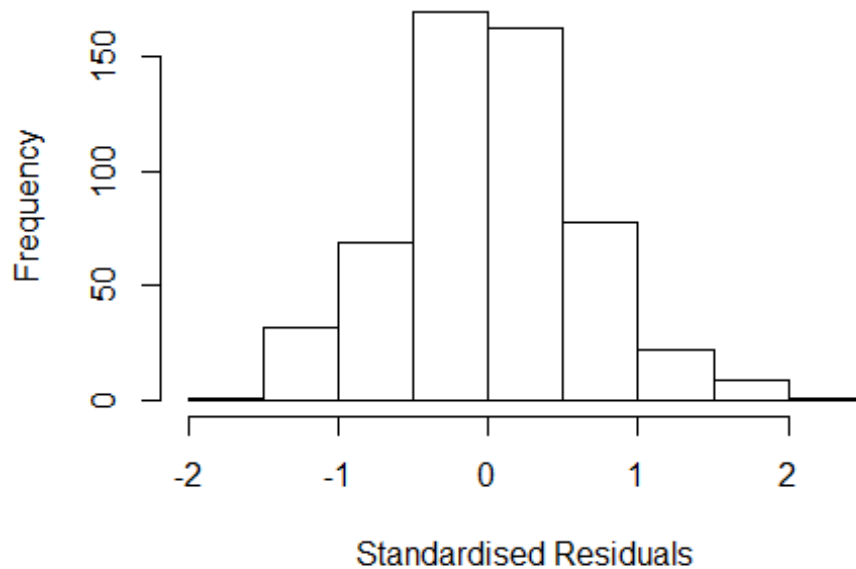
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 9, p-value = 0.9
## 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 | 1.287 | 1 | 1.135 |
| LastAuthorFemale | 1.204 | 1 | 1.097 |
| UniqueAuthors | 2.297 | 4 | 1.110 |
| Year | 2.819 | 16 | 1.033 |

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
## -1.60393 -0.38544 0.00165 0.39897 2.10093
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.3049 0.1487 2.05 0.04077 *
## FirstAuthorFemale1 0.0324 0.0747 0.43 0.66481
## LastAuthorFemale1 -0.0241 0.0990 -0.24 0.80773
## UniqueAuthors2 0.3613 0.1122 3.22 0.00136 **
## UniqueAuthors3 0.5559 0.1228 4.53 7.4e-06 ***
## UniqueAuthors4 0.3946 0.1158 3.41 0.00071 ***
## UniqueAuthors5 0.4101 0.1194 3.43 0.00064 ***
## Year1997 0.6614 0.2840 2.33 0.02024 *
## Year1998 0.5035 0.2733 1.84 0.06595 .
## Year1999 0.1743 0.2783 0.63 0.53141
```

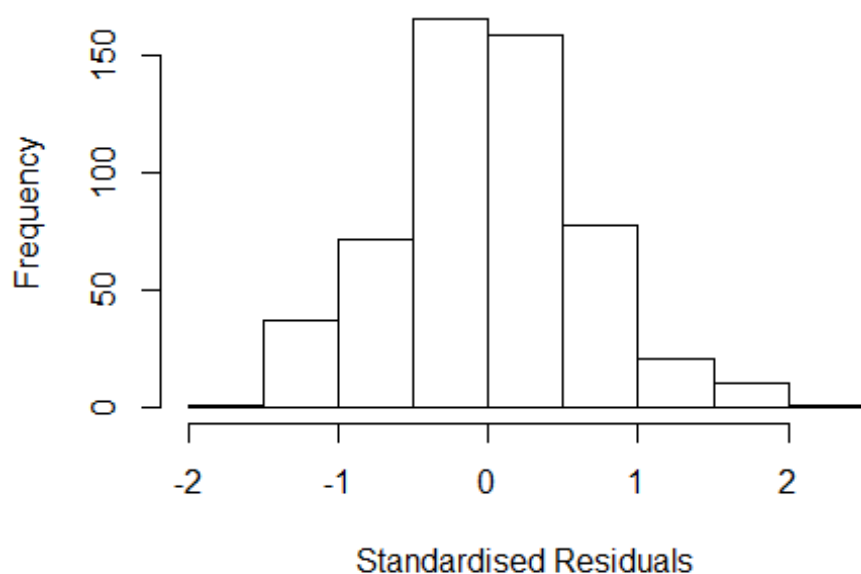


```

## Year2000          0.5487      0.2151      2.55  0.01103 *
## Year2001          0.3147      0.2327      1.35  0.17683
## Year2002          0.4150      0.2128      1.95  0.05166 .
## Year2003          0.5248      0.1822      2.88  0.00413 **
## Year2004          0.6722      0.1780      3.78  0.00018 ***
## Year2005          0.7107      0.1907      3.73  0.00021 ***
## Year2006          0.4831      0.1633      2.96  0.00324 **
## Year2007          0.6629      0.1720      3.85  0.00013 ***
## Year2008          0.4391      0.1844      2.38  0.01758 *
## Year2009          0.5178      0.1784      2.90  0.00385 **
## Year2010          0.4894      0.1716      2.85  0.00453 **
## Year2011          0.5182      0.1613      3.21  0.00139 **
## Year2012          0.4766      0.1787      2.67  0.00789 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.59
## Multiple R-squared:  0.0958, Adjusted R-squared:  0.0577
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 48 weights are ~= 1. The remaining 497 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.178  0.838  0.953  0.896  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          1.83e-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 1.215  1          1.102
## LastAuthorFemale  1.184  1          1.088
## Year              1.435 16          1.011

```

Residuals from first and last author



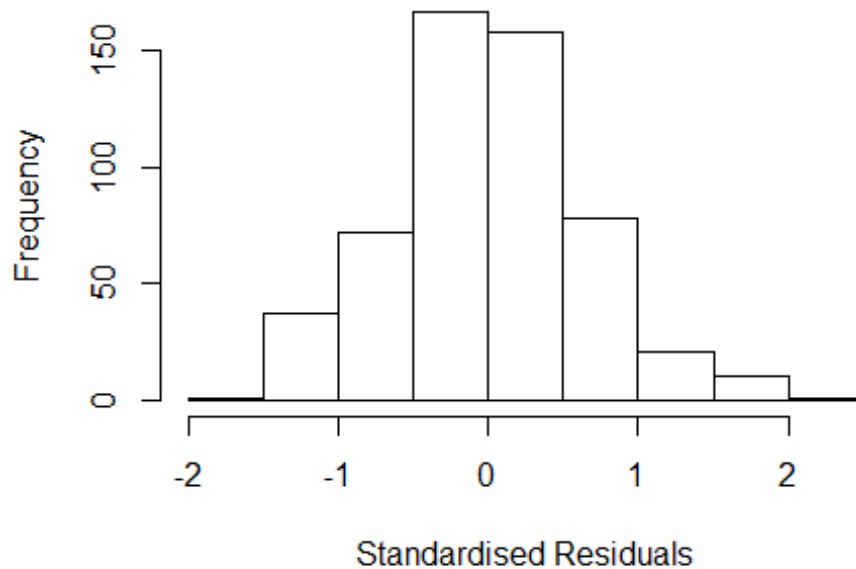
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.53446 -0.40007 -0.00393 0.38305 2.09993
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5243 0.2529 2.07 0.0387 *
## FirstAuthorFemale1 0.0912 0.0751 1.21 0.2254
## LastAuthorFemale1 0.0275 0.1039 0.27 0.7911
## Year1997 0.7242 0.3624 2.00 0.0462 *
## Year1998 0.5806 0.3440 1.69 0.0920 .
## Year1999 0.1522 0.2678 0.57 0.5700
## Year2000 0.6892 0.2914 2.37 0.0184 *
## Year2001 0.4309 0.2999 1.44 0.1514
## Year2002 0.4335 0.3070 1.41 0.1585
## Year2003 0.6266 0.2819 2.22 0.0267 *
## Year2004 0.8155 0.2737 2.98 0.0030 **
## Year2005 0.9190 0.2806 3.27 0.0011 **
```

```

## Year2006          0.6130      0.2657      2.31      0.0214 *
## Year2007          0.7741      0.2680      2.89      0.0040 **
## Year2008          0.6007      0.2807      2.14      0.0328 *
## Year2009          0.6980      0.2710      2.58      0.0103 *
## Year2010          0.6895      0.2680      2.57      0.0104 *
## Year2011          0.7178      0.2613      2.75      0.0062 **
## Year2012          0.6528      0.2722      2.40      0.0168 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.603
## Multiple R-squared:  0.0462, Adjusted R-squared:  0.0136
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 52 weights are ~= 1. The remaining 493 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.200  0.840  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      1.83e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##      trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.208 1      1.099
## Year              1.208 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
## -1.53798 -0.40233 -0.00486 0.38272 2.09767
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5243 0.2525 2.08 0.03836 *
## FirstAuthorFemale1 0.0917 0.0751 1.22 0.22254
## Year1997 0.7238 0.3634 1.99 0.04691 *
## Year1998 0.5815 0.3414 1.70 0.08912 .
## Year1999 0.1522 0.2674 0.57 0.56956
## Year2000 0.6906 0.2923 2.36 0.01850 *
## Year2001 0.4336 0.2989 1.45 0.14751
## Year2002 0.4349 0.3071 1.42 0.15730
## Year2003 0.6309 0.2813 2.24 0.02533 *
## Year2004 0.8229 0.2707 3.04 0.00249 **
## Year2005 0.9219 0.2785 3.31 0.00099 ***
## Year2006 0.6145 0.2649 2.32 0.02074 *
```

```

## Year2007          0.7755      0.2672      2.90  0.00385 **
## Year2008          0.6013      0.2801      2.15  0.03226 *
## Year2009          0.7020      0.2696      2.60  0.00949 **
## Year2010          0.6915      0.2675      2.58  0.01001 *
## Year2011          0.7196      0.2607      2.76  0.00598 **
## Year2012          0.6550      0.2694      2.43  0.01539 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.606
## Multiple R-squared:  0.046, Adjusted R-squared:  0.0153
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 51 weights are ~= 1. The remaining 494 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.207  0.844  0.953  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      1.83e-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 1.172 1      1.083
## Year      1.172 16      1.005
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ 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 -0.4205  0.0162  0.3902  2.0895

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.5243    0.2528   2.07 0.03857 *
## LastAuthorFemale1 0.0318    0.1037   0.31 0.75940
## Year1997          0.7509    0.3601   2.09 0.03752 *
## Year1998          0.5900    0.3454   1.71 0.08817 .
## Year1999          0.1522    0.2677   0.57 0.56988
## Year2000          0.7032    0.2923   2.41 0.01649 *
## Year2001          0.4504    0.2990   1.51 0.13259
## Year2002          0.4444    0.3073   1.45 0.14877
## Year2003          0.6312    0.2814   2.24 0.02531 *
## Year2004          0.8143    0.2730   2.98 0.00299 **
## Year2005          0.9292    0.2787   3.33 0.00092 ***
## Year2006          0.6425    0.2638   2.44 0.01519 *
## Year2007          0.8003    0.2667   3.00 0.00282 **
## Year2008          0.6150    0.2809   2.19 0.02899 *
## Year2009          0.7109    0.2700   2.63 0.00871 **
## Year2010          0.7053    0.2679   2.63 0.00872 **
## Year2011          0.7312    0.2606   2.81 0.00521 **
## Year2012          0.6632    0.2698   2.46 0.01427 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.605
## Multiple R-squared:  0.0436, Adjusted R-squared:  0.0128
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 45 weights are ~= 1. The remaining 500 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.208  0.845   0.952   0.896   0.987   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.83e-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: 545"
## [1] ""

```

```

## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2209"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1   14   22   23   21   37   37   22   39   28   29   36   46   45   44
## 2011 2012
##   30   40
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    1    5   10   12   12   10    9   12   10    9   21   20   28   28
## 2011 2012
##   16   27
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    1    5    9   12   10    9    9    7    7    7   15   10   20   22
## 2011 2012
##   11   21
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.46500296612651"
## [1] "Male first author team size 2018 geometric mean: 2.93726886214801"

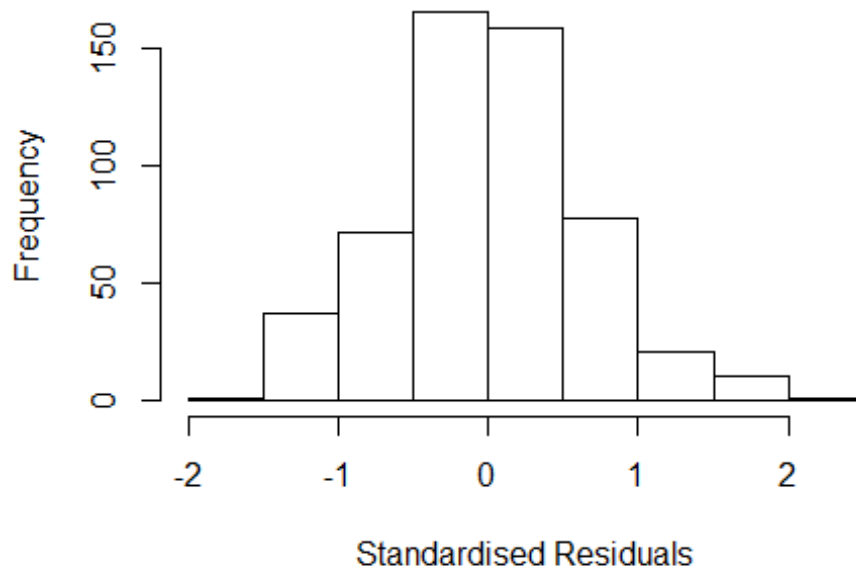
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 80, p-value = 0.5
## 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: 3.16925931910958"

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

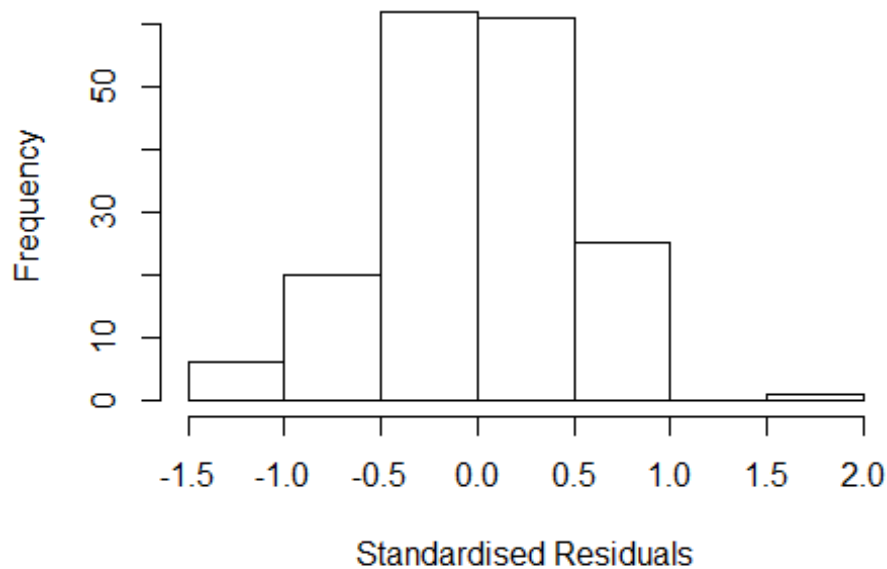
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 31, 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 2.874 1          1.695
## LastAuthorFemale  3.042 1          1.744
## UniqueAuthors    20.464 4          1.458
## Year              47.597 15         1.137
```


Residuals from first and last author and team size



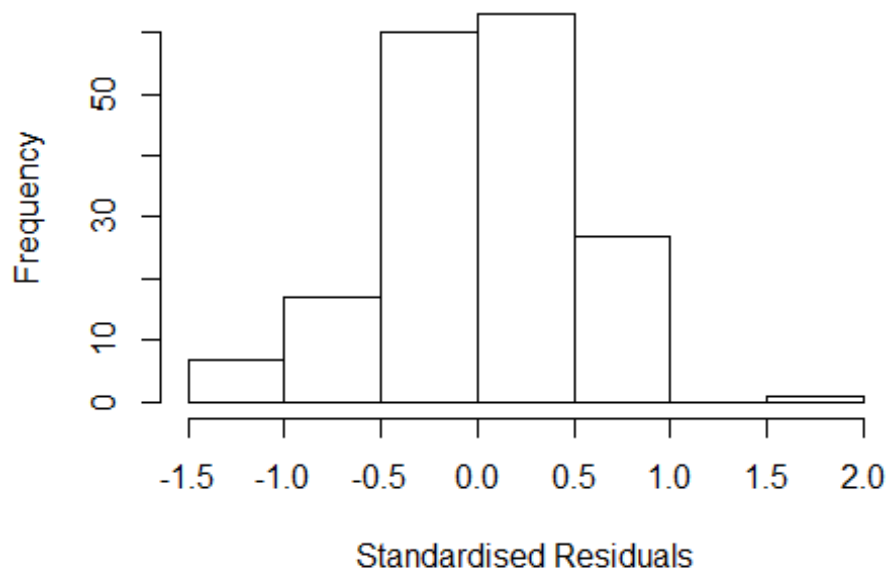
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.33e+00 -2.70e-01 -1.01e-14 3.15e-01 1.92e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.0892 0.1888 11.07 < 2e-16 ***
## FirstAuthorFemale1 0.0154 0.0755 0.20 0.8387
## LastAuthorFemale1 -0.1247 0.1035 -1.20 0.2302
## UniqueAuthors2 0.1464 0.1802 0.81 0.4177
## UniqueAuthors3 0.3978 0.1927 2.07 0.0406 *
## UniqueAuthors4 0.2838 0.1879 1.51 0.1330
## UniqueAuthors5 0.2691 0.1949 1.38 0.1693
## Year1998 -0.7370 0.3115 -2.37 0.0192 *
## Year1999 -1.1588 0.3608 -3.21 0.0016 **
## Year2000 -0.8554 0.1138 -7.51 4.5e-12 ***
```

```

## Year2001          -0.8176      0.1536    -5.32   3.6e-07 ***
## Year2002          -0.6677      0.1906    -3.50   0.0006 ***
## Year2003          -0.9586      0.1810    -5.29   4.1e-07 ***
## Year2004          -0.8995      0.1908    -4.71   5.4e-06 ***
## Year2005          -0.9642      0.1847    -5.22   5.7e-07 ***
## Year2006          -1.1787      0.1581    -7.45   6.3e-12 ***
## Year2007          -1.2400      0.1584    -7.83   7.6e-13 ***
## Year2008          -1.0519      0.1641    -6.41   1.7e-09 ***
## Year2009          -0.8004      0.1205    -6.64   5.1e-10 ***
## Year2010          -1.0985      0.1124    -9.77   < 2e-16 ***
## Year2011          -1.0546      0.1631    -6.47   1.3e-09 ***
## Year2012          -0.7489      0.1738    -4.31   2.9e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.506
## Multiple R-squared:  0.156, Adjusted R-squared:  0.0406
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 159 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.118  0.885   0.959   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      5.71e-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 2.925 1      1.710
## LastAuthorFemale 2.109 1      1.452
## Year              5.772 15      1.060

```

Residuals from first and last author



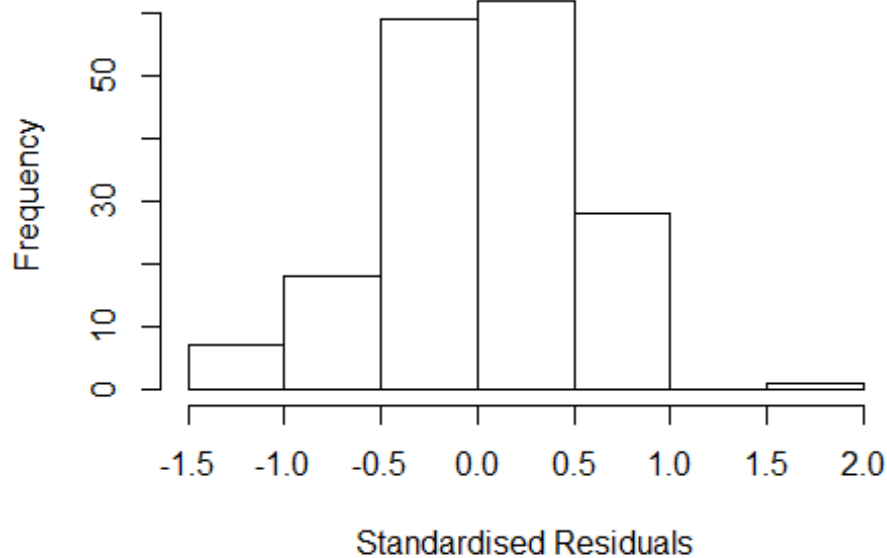
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3224 -0.3021 0.0176 0.3116 1.6218
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.2226 0.0773 28.74 < 2e-16 ***
## FirstAuthorFemale1 0.0284 0.0773 0.37 0.71418
## LastAuthorFemale1 -0.1057 0.0973 -1.09 0.27908
## Year1998 -0.7740 0.3285 -2.36 0.01970 *
## Year1999 -0.9944 0.3568 -2.79 0.00597 **
## Year2000 -0.8128 0.1017 -8.00 2.6e-13 ***
## Year2001 -0.7796 0.1461 -5.34 3.3e-07 ***
## Year2002 -0.6827 0.1526 -4.47 1.5e-05 ***
## Year2003 -0.9003 0.1527 -5.89 2.2e-08 ***
## Year2004 -0.8572 0.1671 -5.13 8.5e-07 ***
## Year2005 -0.9059 0.1823 -4.97 1.7e-06 ***
## Year2006 -1.0935 0.1509 -7.25 1.8e-11 ***
```

```

## Year2007          -1.0981      0.1577   -6.96  8.5e-11 ***
## Year2008          -0.9408      0.1529   -6.15  6.1e-09 ***
## Year2009          -0.7079      0.1125   -6.29  3.0e-09 ***
## Year2010          -0.9885      0.1160   -8.52  1.2e-14 ***
## Year2011          -0.9032      0.1254   -7.20  2.3e-11 ***
## Year2012          -0.5839      0.1612   -3.62  0.00039 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.503
## Multiple R-squared:  0.117, Adjusted R-squared:  0.0217
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 158 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.277  0.876  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      5.71e-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 2.822 1      1.680
## Year              2.822 15      1.035

```

Residuals from first author



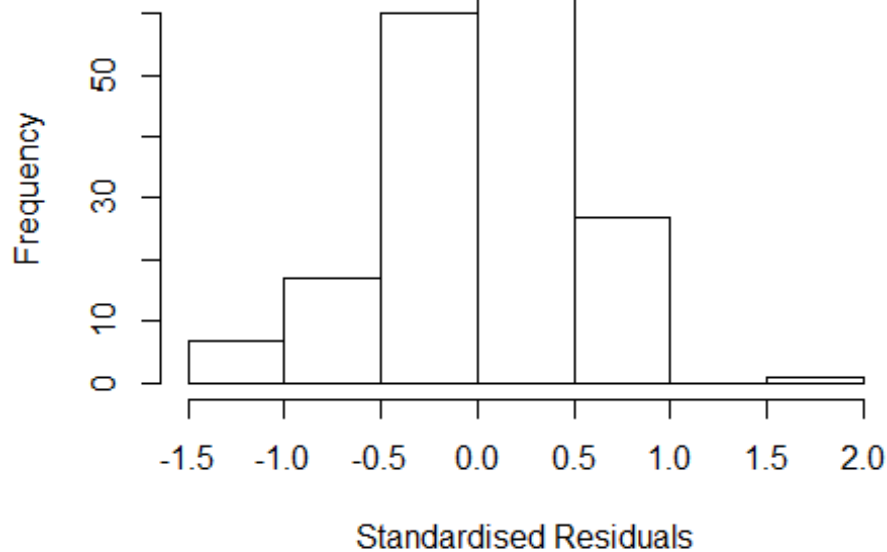
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3027 -0.2954 0.0118 0.3206 1.6788
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.2388 0.0767 29.19 < 2e-16 ***
## FirstAuthorFemale1 0.0122 0.0767 0.16 0.87428
## Year1998 -0.8033 0.3468 -2.32 0.02184 *
## Year1999 -1.0677 0.3242 -3.29 0.00122 **
## Year2000 -0.8370 0.1040 -8.05 1.9e-13 ***
## Year2001 -0.8018 0.1490 -5.38 2.6e-07 ***
## Year2002 -0.7083 0.1484 -4.77 4.1e-06 ***
## Year2003 -0.9361 0.1521 -6.16 5.9e-09 ***
## Year2004 -0.8691 0.1642 -5.29 4.0e-07 ***
## Year2005 -0.9380 0.1769 -5.30 3.8e-07 ***
## Year2006 -1.1400 0.1434 -7.95 3.3e-13 ***
## Year2007 -1.1068 0.1575 -7.03 6.0e-11 ***
```

```

## Year2008          -0.9688      0.1552   -6.24   3.8e-09 ***
## Year2009          -0.7470      0.1061   -7.04   5.5e-11 ***
## Year2010          -1.0415      0.0958  -10.87   < 2e-16 ***
## Year2011          -0.9367      0.1310   -7.15   3.0e-11 ***
## Year2012          -0.6289      0.1607   -3.91   0.00013 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.502
## Multiple R-squared:  0.112, Adjusted R-squared:  0.0217
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 157 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.241  0.875  0.959  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      5.71e-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 2.006 1          1.416
## Year            2.006 15          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
## -1.3332 -0.2998 0.0138 0.3049 1.6104
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.25e+00 1.03e-07 2.18e+07 < 2e-16 ***
## LastAuthorFemale1 -1.00e-01 9.67e-02 -1.04e+00 0.3018
## Year1998 -7.84e-01 3.25e-01 -2.41e+00 0.0171 *
## Year1999 -1.01e+00 3.66e-01 -2.77e+00 0.0064 **
## Year2000 -8.29e-01 9.87e-02 -8.41e+00 2.3e-14 ***
## Year2001 -7.98e-01 1.45e-01 -5.51e+00 1.4e-07 ***
## Year2002 -7.04e-01 1.49e-01 -4.73e+00 4.9e-06 ***
## Year2003 -9.18e-01 1.47e-01 -6.25e+00 3.7e-09 ***
## Year2004 -8.78e-01 1.50e-01 -5.84e+00 2.8e-08 ***
## Year2005 -9.31e-01 1.65e-01 -5.66e+00 7.1e-08 ***
## Year2006 -1.12e+00 1.29e-01 -8.71e+00 3.9e-15 ***
## Year2007 -1.11e+00 1.50e-01 -7.43e+00 6.5e-12 ***
```

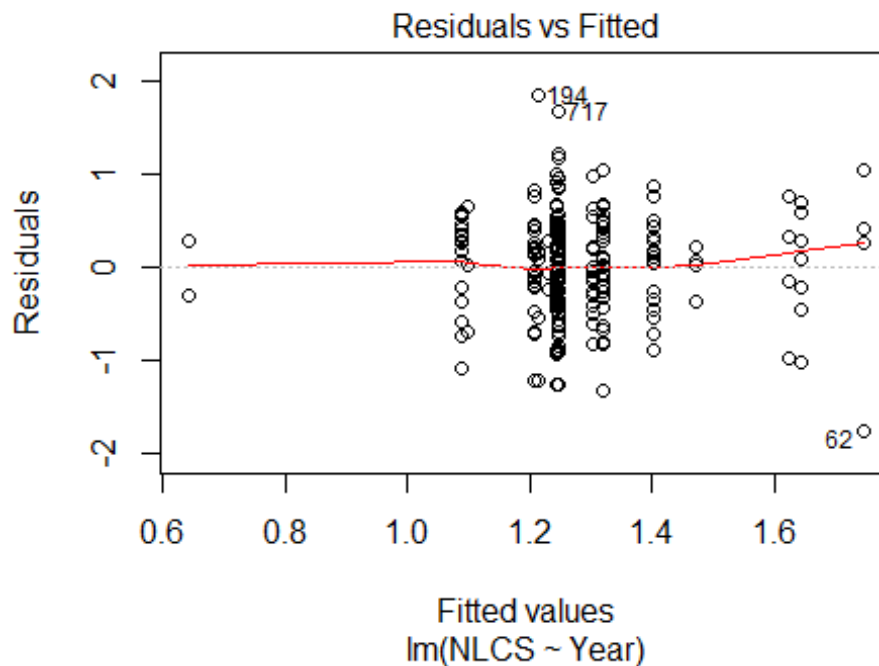
```

## Year2008          -9.64e-01    1.43e-01 -6.75e+00    2.7e-10 ***
## Year2009          -7.30e-01    1.02e-01 -7.15e+00    3.1e-11 ***
## Year2010          -1.01e+00    1.12e-01 -8.95e+00    9.2e-16 ***
## Year2011          -9.24e-01    1.18e-01 -7.82e+00    7.2e-13 ***
## Year2012          -6.06e-01    1.48e-01 -4.10e+00    6.7e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.499
## Multiple R-squared:  0.117, Adjusted R-squared:  0.0274
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 158 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.277  0.877  0.961  0.911  0.986  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      5.71e-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: 175"
## [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
##   18   25   22   24   27   39   27   28   28   43   42   47   47   66   50
## 2011 2012
##   54   72
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    3    4    2    4    7    7    5    8   17   17   19   21   32   23
## 2011 2012
##   29   40

```



```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    2    3    1    4    5    6    5    7   10   14   15   16   30   22
## 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 = 22, df = 16, p-value = 0.1
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.051, df = 1, p-value = 0.8

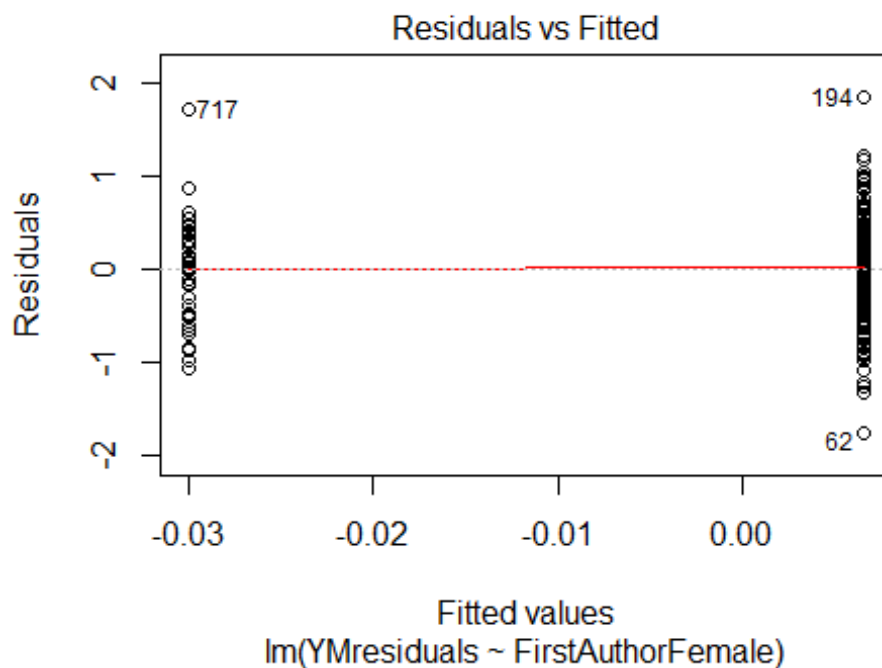
## [1] "Female first author team size 2018 geometric mean: 3.20753432999583"
## [1] "Male first author team size 2018 geometric mean: 3.45385694179826"

## Warning in wilcox.test.default(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: 3.3166247903554"
## [1] "Male last author team size 2018 geometric mean: 3.4318104759178"

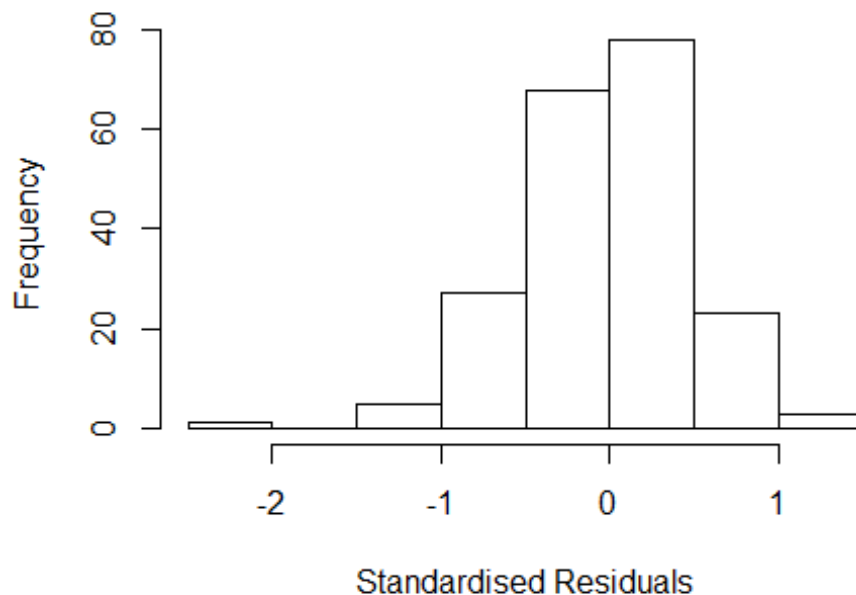
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## 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"
##
```

| | GVI | F | Df | GVI ^{1/(2*Df)} |
|-------------------|--------|----|----|-------------------------|
| FirstAuthorFemale | 1.817 | 1 | 1 | 1.348 |
| LastAuthorFemale | 1.930 | 1 | 1 | 1.389 |
| UniqueAuthors | 29.845 | 4 | 4 | 1.529 |
| Year | 60.578 | 16 | 16 | 1.137 |

Residuals from first and last author and team size



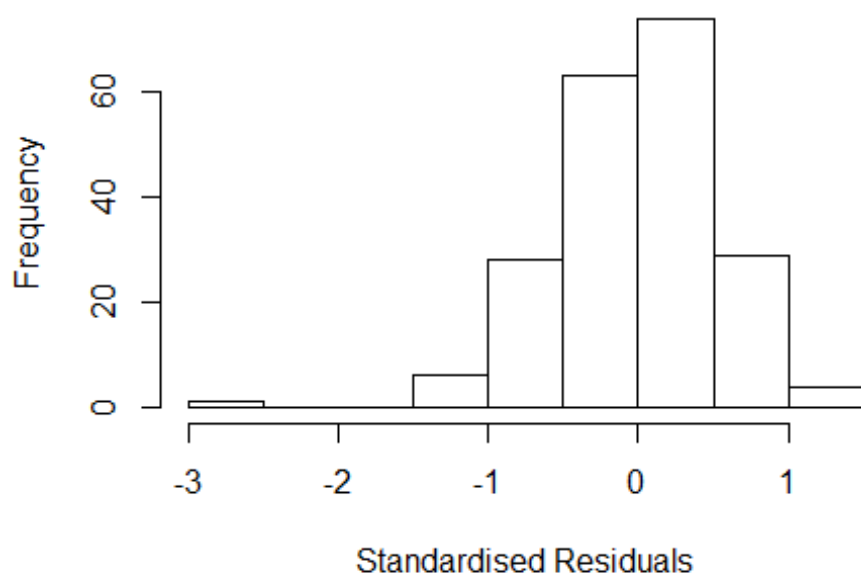
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -2.3108 -0.3520 0.0116 0.3493 1.2708
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.10336 0.15991 6.90 8.3e-11 ***
## FirstAuthorFemale1 -0.10489 0.08519 -1.23 0.21985
## LastAuthorFemale1 0.15673 0.13047 1.20 0.23121
## UniqueAuthors2 0.19137 0.15140 1.26 0.20783
## UniqueAuthors3 0.45799 0.16972 2.70 0.00762 **
## UniqueAuthors4 0.45228 0.17134 2.64 0.00902 **
## UniqueAuthors5 0.39601 0.16210 2.44 0.01552 *
## Year1997 0.34064 0.28582 1.19 0.23489
## Year1998 1.20740 0.34401 3.51 0.00057 ***
## Year1999 -0.16436 0.15991 -1.03 0.30540
```

```

## Year2000      0.20993      0.36150      0.58  0.56214
## Year2001     -0.39574      0.32764     -1.21  0.22868
## Year2002      0.34074      0.27181      1.25  0.21161
## Year2003      0.24008      0.16721      1.44  0.15278
## Year2004     -0.12745      0.28461     -0.45  0.65484
## Year2005     -0.20564      0.19561     -1.05  0.29453
## Year2006     -0.13133      0.19186     -0.68  0.49453
## Year2007      0.00107      0.18589      0.01  0.99541
## Year2008     -0.18843      0.18326     -1.03  0.30521
## Year2009     -0.11140      0.17693     -0.63  0.52975
## Year2010     -0.27919      0.19429     -1.44  0.15245
## Year2011     -0.20678      0.16780     -1.23  0.21944
## Year2012     -0.23560      0.18217     -1.29  0.19756
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.539
## Multiple R-squared:  0.189, Adjusted R-squared:  0.0912
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 181 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0266 0.8940 0.9530 0.9150 0.9810 0.9990
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          4.88e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.745 1 1.321
## LastAuthorFemale 2.133 1 1.460
## Year 3.614 16 1.041

```

Residuals from first and last author



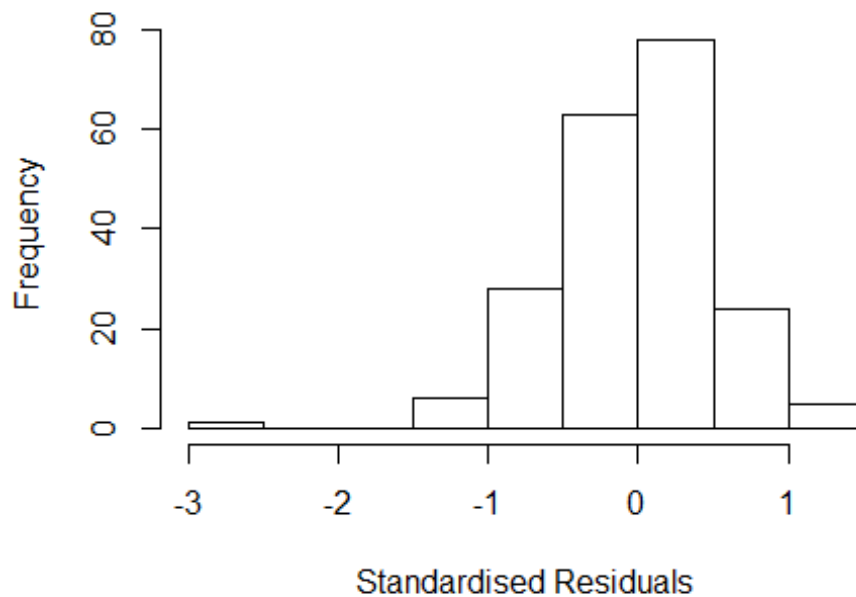
```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId NLCS Year OneField Fields residuals
## 62 11744333849  0 1998    2210      1    -2.501
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.5009 -0.3096  0.0281  0.3512  1.2814
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2280    0.1311   9.37 < 2e-16 ***
## FirstAuthorFemale1 -0.0384    0.0903  -0.43   0.671
## LastAuthorFemale1  0.1923    0.1374   1.40   0.163
## Year1997          0.2160    0.2707   0.80   0.426
## Year1998          1.2728    0.2715   4.69 5.3e-06 ***
## Year1999         -0.2890    0.1311  -2.21  0.029 *
## Year2000          0.4336    0.4520   0.96   0.339
## Year2001         -0.3795    0.2914  -1.30   0.195
## Year2002          0.3475    0.2773   1.25   0.212
## Year2003          0.2601    0.1636   1.59   0.114
## Year2004         -0.0709    0.2368  -0.30   0.765
## Year2005         -0.0687    0.1992  -0.35   0.730
```

```

## Year2006          -0.0244      0.1917   -0.13    0.899
## Year2007           0.1290      0.1721    0.75    0.454
## Year2008          -0.0461      0.1840   -0.25    0.803
## Year2009           0.1344      0.1637    0.82    0.413
## Year2010          -0.0865      0.1862   -0.46    0.643
## Year2011          -0.0173      0.1650   -0.11    0.916
## Year2012          -0.0252      0.1747   -0.14    0.885
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.541
## Multiple R-squared:  0.12,   Adjusted R-squared:  0.0352
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 190 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0006 0.8860 0.9570 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      4.88e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.704 1          1.305
## Year              1.704 16          1.017

```

Residuals from first author



```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId NLCS Year OneField Fields residuals
## 62 11744333849  0 1998    2210      1    -2.501
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.5053 -0.3315  0.0296  0.3685  1.2693
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.22801    0.13110   9.37  < 2e-16 ***
## FirstAuthorFemale1 -0.04586    0.08967  -0.51   0.610
## Year1997         0.21599    0.27085   0.80   0.426
## Year1998         1.27729    0.26178   4.88 2.3e-06 ***
## Year1999        -0.28901    0.13110  -2.20  0.029 *
## Year2000         0.43802    0.40534   1.08   0.281
## Year2001        -0.37728    0.29093  -1.30   0.196
## Year2002         0.35130    0.27715   1.27   0.207
## Year2003         0.26318    0.16382   1.61   0.110
## Year2004         0.05423    0.20493   0.26   0.792
## Year2005        -0.02110    0.19394  -0.11   0.913
## Year2006         0.00644    0.18827   0.03   0.973
```

```

## Year2007          0.18212      0.17318      1.05      0.294
## Year2008          -0.00851      0.18303     -0.05      0.963
## Year2009           0.13571      0.16376      0.83      0.408
## Year2010          -0.07431      0.18559     -0.40      0.689
## Year2011           0.01075      0.16991      0.06      0.950
## Year2012           0.00552      0.17311      0.03      0.975
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.537
## Multiple R-squared:  0.108, Adjusted R-squared:  0.0265
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## observation 6 is an outlier with |weight| <= 9.4e-05 ( < 0.00049);
## 22 weights are ~= 1. The remaining 182 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.499  0.881  0.953  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      4.88e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 2.13 1          1.460
## Year              2.13 16          1.024
##
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId NLCS Year OneField Fields residuals
## 62 11744333849      0 1998      2210      1      -2.501
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max

```



```

## -2.4815 -0.3092 0.0278 0.3399 1.2831
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2280    0.1311   9.37 < 2e-16 ***
## LastAuthorFemale1 0.1947    0.1373   1.42  0.158
## Year1997          0.2160    0.2709   0.80  0.426
## Year1998          1.2535    0.2835   4.42 1.7e-05 ***
## Year1999         -0.2890    0.1311  -2.20  0.029 *
## Year2000          0.4354    0.4538   0.96  0.339
## Year2001         -0.3890    0.2908  -1.34  0.183
## Year2002          0.3292    0.2773   1.19  0.237
## Year2003          0.2442    0.1567   1.56  0.121
## Year2004         -0.0844    0.2316  -0.36  0.716
## Year2005         -0.0798    0.1987  -0.40  0.689
## Year2006         -0.0318    0.1903  -0.17  0.867
## Year2007          0.1207    0.1704   0.71  0.480
## Year2008         -0.0566    0.1804  -0.31  0.754
## Year2009          0.1291    0.1629   0.79  0.429
## Year2010         -0.0881    0.1869  -0.47  0.638
## Year2011         -0.0208    0.1642  -0.13  0.899
## Year2012         -0.0368    0.1719  -0.21  0.830
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.537
## Multiple R-squared:  0.12, Adjusted R-squared:  0.04
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 192 ones are summarized as
##   Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.0007 0.8830 0.9590 0.9130 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      4.88e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 205"

```

```

## [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
##    15    10    23    18    26    32    19    31    24    38    24    30    33    53    44
## 2011 2012
##    41    54
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     2     1     1     1     3     9     5     7     9    15    11    11    15    28    21
## 2011 2012
##    18    23
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     2     1     1     1     3     9     5     7     8    10     9    10    13    24    18
## 2011 2012
##    17    17
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 5.74456264653803"
## [1] "Male first author team size 2018 geometric mean: 3.12796043033154"

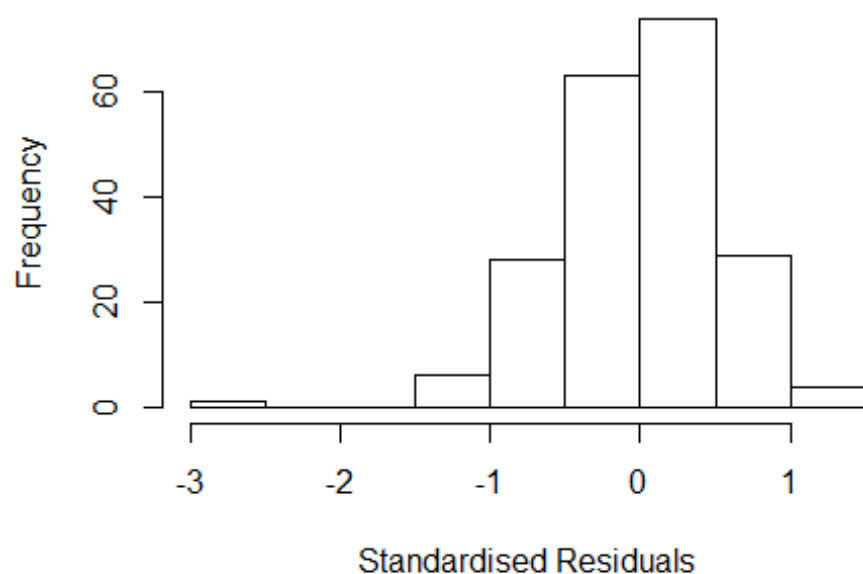
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 23, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 11"
## [1] "Male last author team size 2018 geometric mean: 3.12028448666954"

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

```

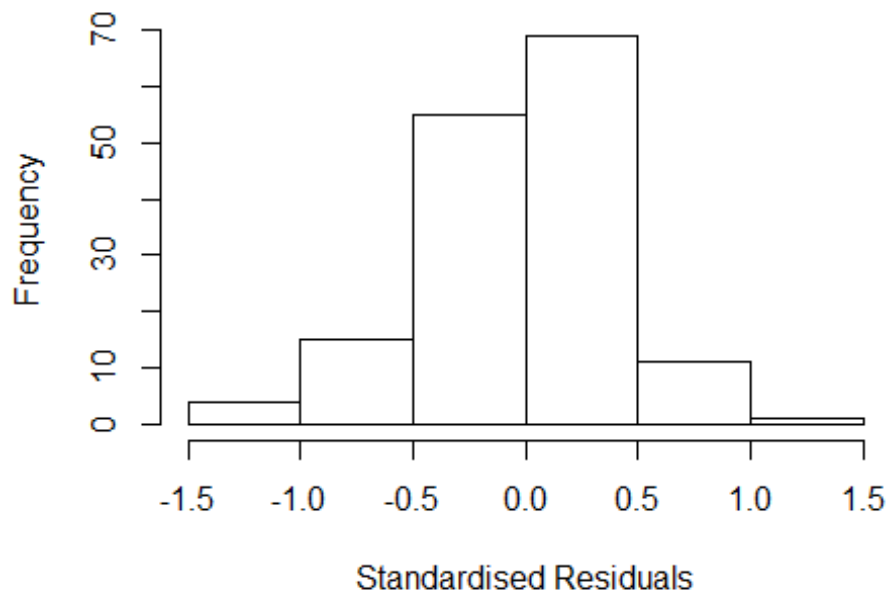
Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 17, 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 | 4.055e+13 | 1 | 6.368e+06 |
| ## LastAuthorFemale | 9.004e+13 | 1 | 9.489e+06 |
| ## UniqueAuthors | 8.632e+14 | 4 | 7.362e+01 |
| ## Year | 4.277e+28 | 16 | 7.847e+00 |

Residuals from first and last author and team size



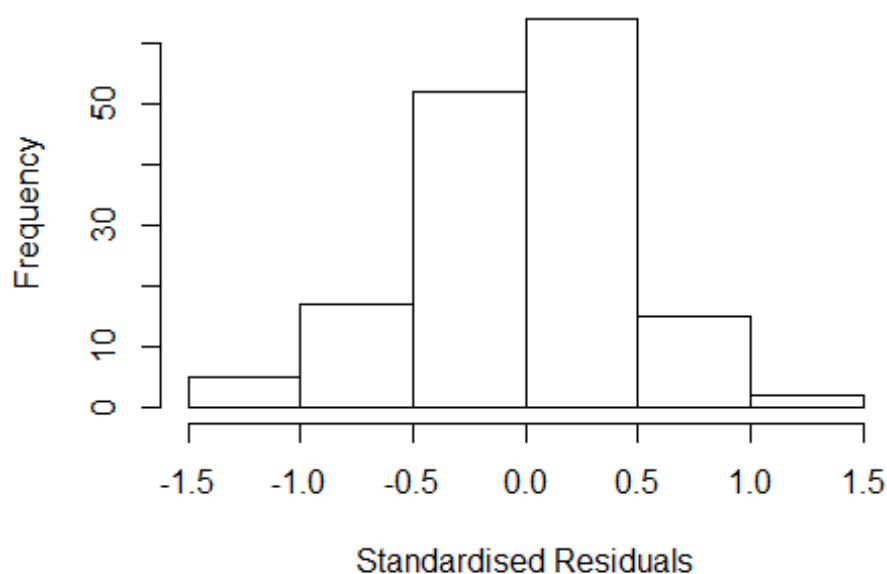
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.4163 -0.2461  0.0284  0.2242  1.2177
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.27170    0.09820   12.95 < 2e-16 ***
## FirstAuthorFemale1 -0.06768    0.10838   -0.62  0.53340
## LastAuthorFemale1 -0.07146    0.16097   -0.44  0.65783
## UniqueAuthors2     0.15359    0.13531    1.14  0.25837
## UniqueAuthors3     0.24611    0.16336    1.51  0.13432
## UniqueAuthors4     0.49523    0.19916    2.49  0.01415 *
## UniqueAuthors5     0.22636    0.15850    1.43  0.15562
## Year1997          0.24443    0.23431    1.04  0.29876
## Year1998          1.37070    0.09820   13.96 < 2e-16 ***
## Year1999         -0.33270    0.09820   -3.39  0.00093 ***
```

```

## Year2000      0.29634      0.26576      1.12      0.26684
## Year2001     -0.26409      0.27190     -0.97      0.33318
## Year2002     -0.00902      0.54570     -0.02      0.98683
## Year2003     -0.13726      0.11699     -1.17      0.24278
## Year2004     -0.11463      0.14376     -0.80      0.42666
## Year2005     -0.05943      0.11799     -0.50      0.61531
## Year2006     -0.34677      0.16862     -2.06      0.04170 *
## Year2007     -0.10531      0.15838     -0.66      0.50728
## Year2008     -0.24905      0.14187     -1.76      0.08149 .
## Year2009     -0.19031      0.13585     -1.40      0.16360
## Year2010     -0.39437      0.16835     -2.34      0.02065 *
## Year2011     -0.35351      0.14934     -2.37      0.01938 *
## Year2012     -0.37805      0.15382     -2.46      0.01528 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.435
## Multiple R-squared:  0.23, Adjusted R-squared:  0.102
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 138 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.268  0.889  0.962  0.907  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      6.45e-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 6.289e+12 1      2.508e+06
## LastAuthorFemale  1.954e+13 1      4.420e+06
## Year              4.144e+13 16      2.664e+00

```

Residuals from first and last author



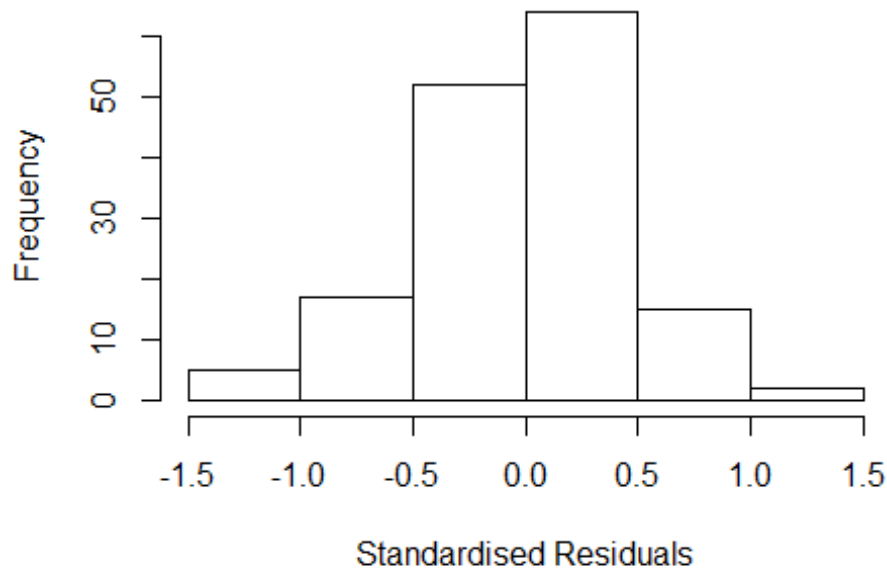
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2713 -0.2618  0.0351  0.2669  1.2532
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.34850    0.12855   10.49  <2e-16 ***
## FirstAuthorFemale1  0.04731    0.09165    0.52  0.6065
## LastAuthorFemale1 -0.00791    0.16155   -0.05  0.9610
## Year1997        -0.01091    0.21130   -0.05  0.9589
## Year1998         1.44750    0.12855   11.26  <2e-16 ***
## Year1999        -0.40950    0.12855   -3.19  0.0018 **
## Year2000         0.54679    0.32908    1.66  0.0989 .
## Year2001        -0.26552    0.30124   -0.88  0.3796
## Year2002        -0.07723    0.53431   -0.14  0.8853
## Year2003        -0.09531    0.16111   -0.59  0.5551
## Year2004        -0.12896    0.16751   -0.77  0.4427
## Year2005         0.03630    0.15340    0.24  0.8133
```

```

## Year2006      -0.31966      0.19392      -1.65      0.1016
## Year2007      -0.06428      0.20452      -0.31      0.7538
## Year2008      -0.11487      0.17432      -0.66      0.5110
## Year2009      -0.01033      0.15418      -0.07      0.9467
## Year2010      -0.25837      0.19468      -1.33      0.1867
## Year2011      -0.24577      0.16154      -1.52      0.1305
## Year2012      -0.18845      0.18448      -1.02      0.3088
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.443
## Multiple R-squared:  0.167, Adjusted R-squared:  0.0572
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 144 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.390  0.869  0.963  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      6.45e-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 8.009e+13 1      8.949e+06
## Year      8.009e+13 16      2.719e+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
## -1.2780 -0.2619 0.0381 0.2662 1.2557
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3485 0.1286 10.48 <2e-16 ***
## FirstAuthorFemale1 0.0462 0.0912 0.51 0.6132
## Year1997 -0.0177 0.1577 -0.11 0.9107
## Year1998 1.4475 0.1286 11.25 <2e-16 ***
## Year1999 -0.4095 0.1286 -3.18 0.0018 **
## Year2000 0.5471 0.3301 1.66 0.0997 .
## Year2001 -0.2643 0.3028 -0.87 0.3843
## Year2002 -0.0705 0.5383 -0.13 0.8959
## Year2003 -0.0950 0.1612 -0.59 0.5568
## Year2004 -0.1326 0.1505 -0.88 0.3798
## Year2005 0.0349 0.1487 0.23 0.8146
## Year2006 -0.3222 0.1915 -1.68 0.0948 .
```



```

## Year2007          -0.0645      0.2024   -0.32    0.7506
## Year2008          -0.1154      0.1804   -0.64    0.5234
## Year2009          -0.0100      0.1543   -0.06    0.9484
## Year2010          -0.2583      0.1951   -1.32    0.1877
## Year2011          -0.2451      0.1616   -1.52    0.1315
## Year2012          -0.1874      0.1842   -1.02    0.3107
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.438
## Multiple R-squared:  0.168, Adjusted R-squared:  0.0651
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 144 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.376  0.868  0.963  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      6.45e-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 5.439e+14 1      2.332e+07
## Year      5.439e+14 16      2.887e+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
## -1.3210 -0.2719  0.0293  0.2632  1.2484

```

```

##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.34850    0.12863   10.48 <2e-16 ***
## LastAuthorFemale1 -0.00373    0.16192   -0.02  0.9817
## Year1997       0.03223    0.20680    0.16  0.8764
## Year1998       1.44750    0.12863   11.25 <2e-16 ***
## Year1999      -0.40950    0.12863   -3.18  0.0018 **
## Year2000       0.54712    0.33006    1.66  0.0997 .
## Year2001      -0.24403    0.28803   -0.85  0.3983
## Year2002      -0.02747    0.50176   -0.05  0.9564
## Year2003      -0.08185    0.15908   -0.51  0.6077
## Year2004      -0.11962    0.16962   -0.71  0.4819
## Year2005       0.04512    0.15233    0.30  0.7675
## Year2006      -0.31490    0.19204   -1.64  0.1034
## Year2007      -0.04796    0.20361   -0.24  0.8141
## Year2008      -0.10160    0.17014   -0.60  0.5514
## Year2009      -0.00546    0.15405   -0.04  0.9718
## Year2010      -0.25071    0.19502   -1.29  0.2008
## Year2011      -0.23241    0.15794   -1.47  0.1435
## Year2012      -0.17386    0.18364   -0.95  0.3455
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.439
## Multiple R-squared:  0.167, Adjusted R-squared:  0.0638
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 142 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.344  0.866  0.961  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           6.45e-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: 155"
## [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]"
##
## 1997 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
## 11 1 2 3 1 3 1 2 1 1 2 4 6 3 6
##
## 1997 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
## 2 1 0 0 0 2 1 2 1 0 1 3 2 2 5
##
## 1997 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
## 1 0 0 0 0 2 1 2 1 0 1 3 2 2 4
## [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: 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 = 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: 2.47939698673123"

## Warning in wilcox.test.default(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 = 0.06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
## [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: 19"
## [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
##    6    6   11    7    6    7    5    1    2    9    3    7   10   15   18
## 2011 2012
##   14   25
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    3    3    2    3    4    3    1    2    9    1    4    8    9   12
## 2011 2012
##   11   17
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    3    2    2    3    4    2    1    2    8    1    4    6    6   10
## 2011 2012
##    8   14
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.36042145371267"
## [1] "Male first author team size 2018 geometric mean: 1.25992104989487"

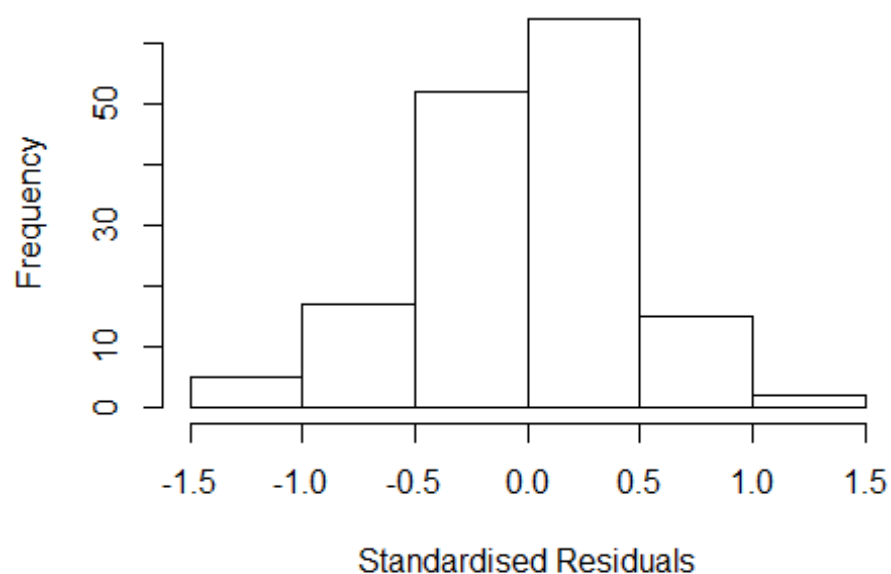
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 18, 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: 2.35930400063821"

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

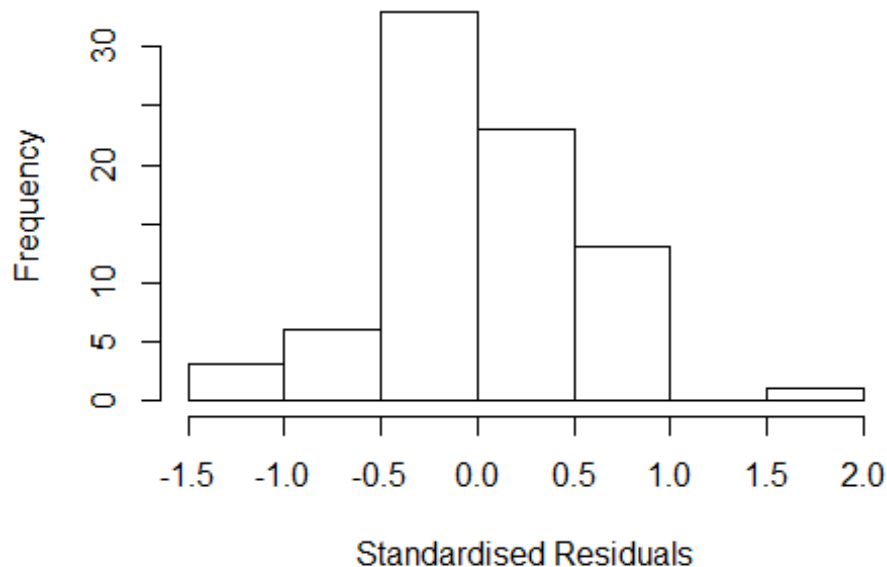
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4.5, 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  106.64  1         10.327
## LastAuthorFemale   41.09  1          6.410
## UniqueAuthors     5168.33  4          2.912
## Year              21439.59 16          1.366
```

Residuals from first and last author and team size



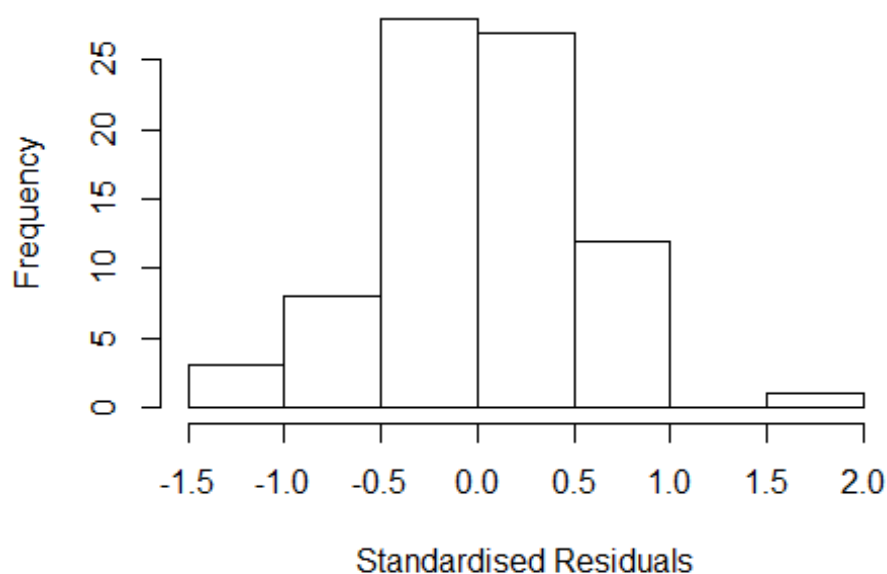
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1847 -0.2845 -0.0228 0.2911 1.8148
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.57343 0.11967 4.79 1.3e-05 ***
## FirstAuthorFemale1 -0.05188 0.33009 -0.16 0.876
## LastAuthorFemale1 0.44145 0.20712 2.13 0.037 *
## UniqueAuthors2 0.08115 0.29617 0.27 0.785
## UniqueAuthors3 -0.01419 0.22095 -0.06 0.949
## UniqueAuthors4 0.39149 0.16468 2.38 0.021 *
## UniqueAuthors5 0.34966 0.27089 1.29 0.202
## Year1997 0.46086 0.41746 1.10 0.274
## Year1998 0.72966 0.87080 0.84 0.406
## Year1999 0.00471 0.36615 0.01 0.990
```

```

## Year2000      0.19028    0.26421    0.72    0.474
## Year2001      0.35337    0.37815    0.93    0.354
## Year2002      0.25800    0.20716    1.25    0.218
## Year2003     -0.35158    0.30481   -1.15    0.254
## Year2004     -0.00862    0.22854   -0.04    0.970
## Year2005      0.21981    0.87712    0.25    0.803
## Year2006      0.06542    0.30481    0.21    0.831
## Year2007      0.19339    0.25263    0.77    0.447
## Year2008      0.48433    0.28661    1.69    0.097 .
## Year2009      0.02859    0.15693    0.18    0.856
## Year2010      0.43487    0.30480    1.43    0.159
## Year2011      0.43330    0.27291    1.59    0.118
## Year2012      0.42080    0.19043    2.21    0.031 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.495
## Multiple R-squared:  0.262, Adjusted R-squared:  -0.0282
## Convergence in 34 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 68 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.151  0.868  0.956  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      1.27e-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 20.86 1      4.567
## LastAuthorFemale  46.44 1      6.815
## Year              12.92 16      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
## -1.1575 -0.3571 0.0396 0.3780 1.6799
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7060 0.1393 5.07 4.1e-06 ***
## FirstAuthorFemale1 -0.0400 0.2270 -0.18 0.8608
## LastAuthorFemale1 0.3817 0.2007 1.90 0.0620 .
## Year1997 0.3500 0.3792 0.92 0.3597
## Year1998 0.5900 0.6774 0.87 0.3872
## Year1999 -0.0634 0.2469 -0.26 0.7983
## Year2000 0.2318 0.2591 0.89 0.3746
## Year2001 0.3748 0.3089 1.21 0.2297
## Year2002 0.1660 0.1942 0.86 0.3959
## Year2003 -0.4030 0.1393 -2.89 0.0053 **
## Year2004 -0.1244 0.1780 -0.70 0.4876
## Year2005 0.2221 0.4190 0.53 0.5981
```

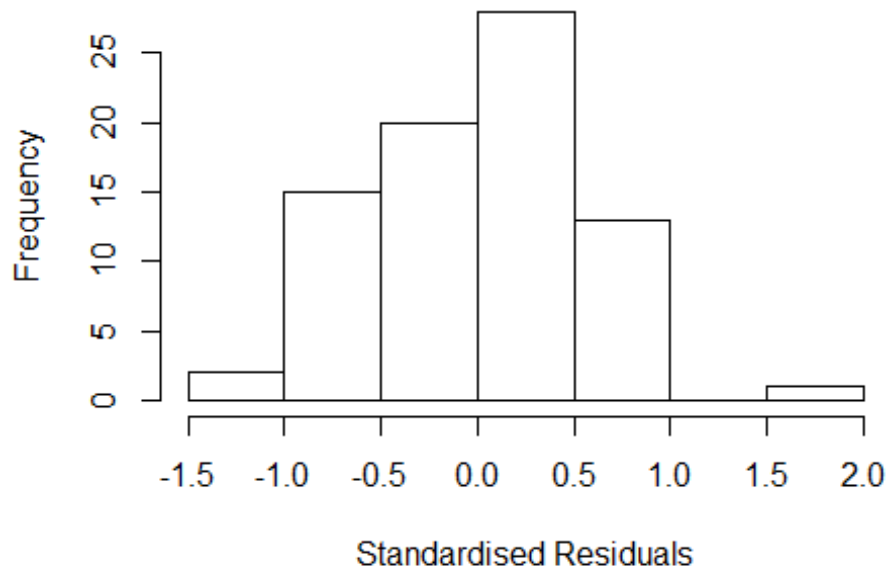


```

## Year2006          0.0140      0.1393      0.10      0.9202
## Year2007          0.1822      0.2410      0.76      0.4525
## Year2008          0.4844      0.2916      1.66      0.1019
## Year2009          0.0317      0.2012      0.16      0.8752
## Year2010          0.3765      0.2287      1.65      0.1049
## Year2011          0.4280      0.2455      1.74      0.0864 .
## Year2012          0.5278      0.2051      2.57      0.0126 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.619
## Multiple R-squared:  0.179, Adjusted R-squared:  -0.0679
## 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.
##  0.441  0.923  0.962  0.934  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 5.023 1      2.241
## Year              5.023 16      1.052

```

Residuals from first author



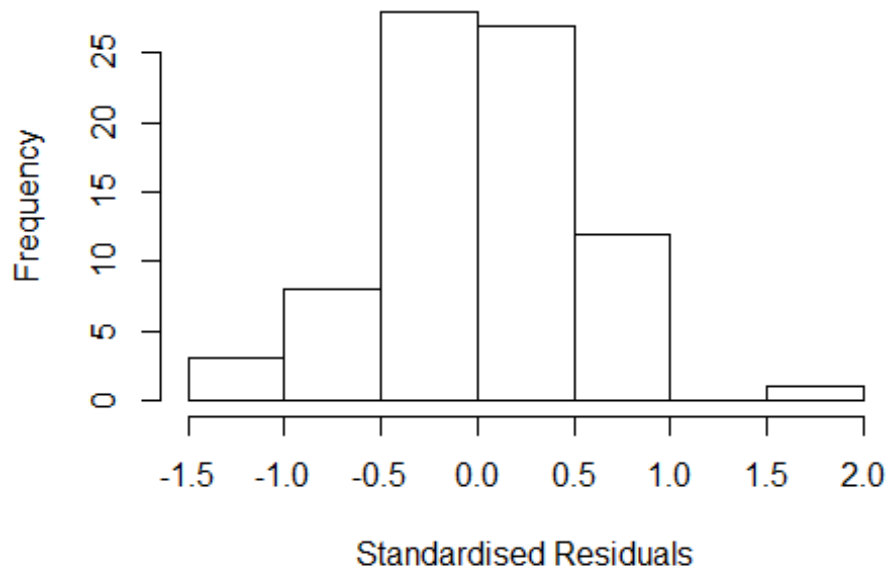
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.289 -0.349 0.053 0.378 1.648
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.70590 0.13921 5.07 4e-06 ***
## FirstAuthorFemale1 0.10529 0.21423 0.49 0.6249
## Year1997 0.34913 0.37728 0.93 0.3584
## Year1998 0.59010 0.66712 0.88 0.3799
## Year1999 0.05496 0.19578 0.28 0.7799
## Year2000 0.30652 0.31044 0.99 0.3274
## Year2001 0.37616 0.30803 1.22 0.2267
## Year2002 0.16610 0.19400 0.86 0.3952
## Year2003 -0.40290 0.13921 -2.89 0.0053 **
## Year2004 -0.00604 0.20919 -0.03 0.9771
## Year2005 0.25379 0.46349 0.55 0.5860
## Year2006 0.01410 0.13921 0.10 0.9196
```

```

## Year2007          0.24629    0.23528    1.05    0.2993
## Year2008          0.47787    0.30573    1.56    0.1232
## Year2009          0.00724    0.20611    0.04    0.9721
## Year2010          0.34509    0.22794    1.51    0.1352
## Year2011          0.49625    0.23969    2.07    0.0427 *
## Year2012          0.58311    0.20939    2.78    0.0071 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.633
## Multiple R-squared:  0.142, Adjusted R-squared:  -0.0975
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 69 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.478  0.918  0.957  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      1.27e-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 17.19 1      4.146
## Year              17.19 16      1.093

```

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
## -1.1879 -0.3644 0.0304 0.3730 1.6999
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7059 0.1393 5.07 4e-06 ***
## LastAuthorFemale1 0.3603 0.1718 2.10 0.0401 *
## Year1997 0.3497 0.3784 0.92 0.3591
## Year1998 0.5901 0.6732 0.88 0.3842
## Year1999 -0.0726 0.2437 -0.30 0.7669
## Year2000 0.2258 0.2529 0.89 0.3756
## Year2001 0.3754 0.3085 1.22 0.2283
## Year2002 0.1661 0.1941 0.86 0.3956
## Year2003 -0.4029 0.1393 -2.89 0.0053 **
## Year2004 -0.1336 0.1654 -0.81 0.4223
## Year2005 0.2022 0.3583 0.56 0.5747
## Year2006 0.0141 0.1393 0.10 0.9200
```

```

## Year2007          0.1773      0.2382      0.74      0.4596
## Year2008          0.4820      0.2972      1.62      0.1099
## Year2009          0.0249      0.1977      0.13      0.9001
## Year2010          0.3677      0.2227      1.65      0.1039
## Year2011          0.4228      0.2425      1.74      0.0864 .
## Year2012          0.5223      0.2077      2.51      0.0146 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.624
## Multiple R-squared:  0.178, Adjusted R-squared:  -0.0507
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 71 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.438  0.924  0.959  0.934  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] "Sample size for the above analysis: 79"
## [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]"
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    7    4    5    2    3    1    4    4    3    5    5    9    4    8
## 2012
##    5
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    7    2    4    0    3    1    4    3    3    5    4    5    2    6
## 2012

```

```
##      3
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##      1      7      2      4      0      3      0      3      2      3      3      4      4      1      6
## 2012
##      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: 2"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale      1.500  1          1.225
## LastAuthorFemale       4.573  1          2.138
## UniqueAuthors        9947.035  4          3.160
## Year                 31141.489 13          1.489

## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##      Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##      k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.7654 -0.1500 -0.0102  0.1378  1.5293
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.52e+00   1.60e-01   9.51e+00  6.0e-10 ***
## FirstAuthorFemale1  8.88e-01   3.47e-01   2.56e+00  0.01674 *
## LastAuthorFemale1 -2.19e-01   2.78e-01  -7.90e-01  0.43785
## UniqueAuthors2    6.39e-01   1.60e-01   3.99e+00  0.00048 ***
## UniqueAuthors3    9.27e-01   2.13e-01   4.35e+00  0.00018 ***
## UniqueAuthors4    5.50e-01   2.76e-01   1.99e+00  0.05680 .
## UniqueAuthors5    7.87e-01   2.77e-01   2.84e+00  0.00868 **
## Year1998          -1.14e+00   4.72e-01  -2.41e+00  0.02350 *
## Year1999          -6.36e-01   1.01e-01  -6.29e+00  1.2e-06 ***
## Year2000          -1.37e+00   1.38e-01  -9.97e+00  2.3e-10 ***
## Year2002          -1.72e-01   1.23e-01  -1.40e+00  0.17202
## Year2004          -4.54e-01   1.18e-01  -3.85e+00  0.00070 ***
## Year2005          -7.03e-01   8.94e-02  -7.86e+00  2.5e-08 ***
## Year2006          -7.54e-01   1.65e-01  -4.57e+00  0.00010 ***
## Year2007          -4.61e-01   1.67e-01  -2.75e+00  0.01065 *
```

```

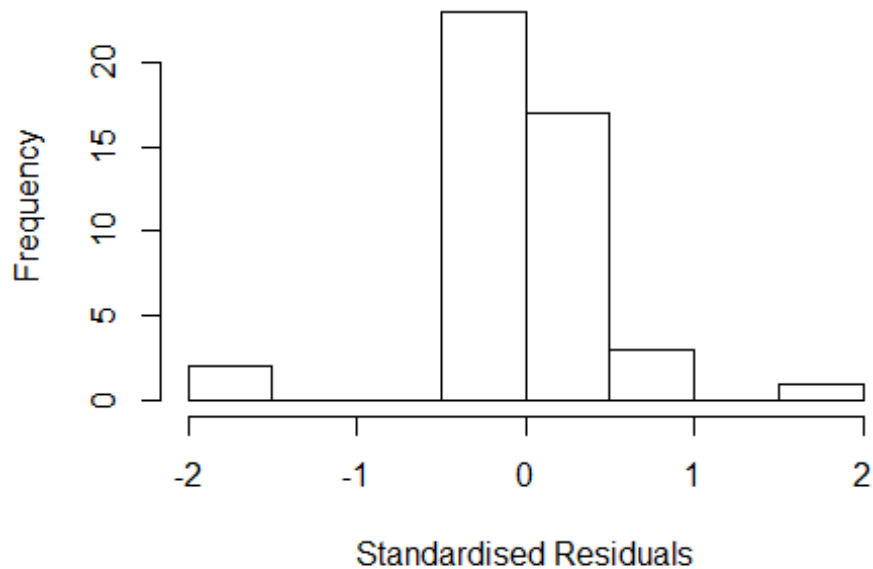
## Year2008          -7.66e-01   1.44e-01 -5.34e+00   1.4e-05 ***
## Year2009          -2.44e-01   1.30e-01 -1.88e+00   0.07079 .
## Year2010          -2.98e-01   1.88e-08 -1.59e+07   < 2e-16 ***
## Year2011          -9.70e-01   2.12e-01 -4.58e+00   0.00010 ***
## Year2012          -3.97e-01   1.26e-01 -3.15e+00   0.00405 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.352
## Multiple R-squared:  0.808, Adjusted R-squared:  0.668
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 2 observations c(16,46) are outliers with |weight| = 0 ( < 0.0022);
## 7 weights are ~ = 1. The remaining 37 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.019  0.918  0.983   0.911  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      2.17e-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 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
## -1.58e+00 -2.70e-01 -3.19e-15  2.01e-01  1.10e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.1620     0.0000      Inf < 2e-16 ***
## FirstAuthorFemale1  0.8250     0.1740      4.74 4.8e-05 ***
## LastAuthorFemale1 -0.4693     0.1746     -2.69 0.01161 *
## Year1998          -1.0846     0.2972     -3.65 0.00099 ***
## Year1999          -0.9560     0.3185     -3.00 0.00537 **
## Year2000          -1.8642     0.1961     -9.51 1.5e-10 ***
## Year2002          -0.8831     0.3803     -2.32 0.02722 *
```



```

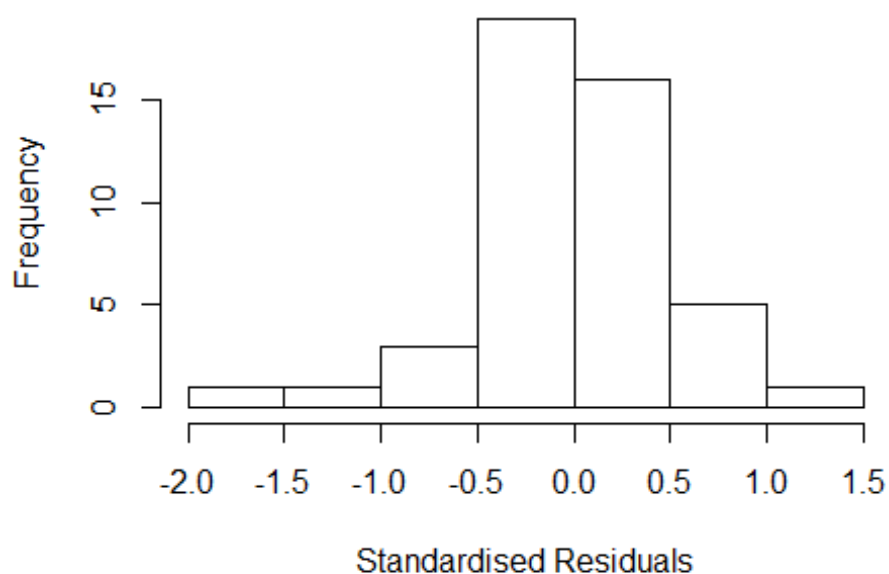
## Year2004          -0.5696      0.3365    -1.69   0.10093
## Year2005          -0.5585      0.1467    -3.81   0.00065 ***
## Year2006          -0.8596      0.1073    -8.01   6.0e-09 ***
## Year2007          -0.3270      0.1362    -2.40   0.02274 *
## Year2008          -0.6341      0.1790    -3.54   0.00132 **
## Year2009          -0.1205      0.0804    -1.50   0.14437
## Year2010          -0.2980      0.0000     -Inf   < 2e-16 ***
## Year2011          -0.9310      0.2029    -4.59   7.4e-05 ***
## Year2012          -0.5822      0.6985    -0.83   0.41118
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.559
## Multiple R-squared:  0.496, Adjusted R-squared:  0.244
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 38 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.404  0.891  0.972   0.924   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          2.17e-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 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
## -1.59187 -0.22468 -0.00525  0.28496  1.10532
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.1620     0.0000      Inf < 2e-16 ***
## FirstAuthorFemale1  0.6021     0.1185     5.08 1.7e-05 ***
## Year1998          -1.0853     0.2986    -3.63 0.00100 ***
## Year1999          -0.9560     0.3193    -2.99 0.00536 **
## Year2000          -1.8645     0.1966   -9.49 1.1e-10 ***
## Year2002          -0.8828     0.3819    -2.31 0.02762 *
## Year2004          -0.5696     0.3376    -1.69 0.10162
## Year2005          -0.5585     0.1468   -3.80 0.00063 ***
```

```

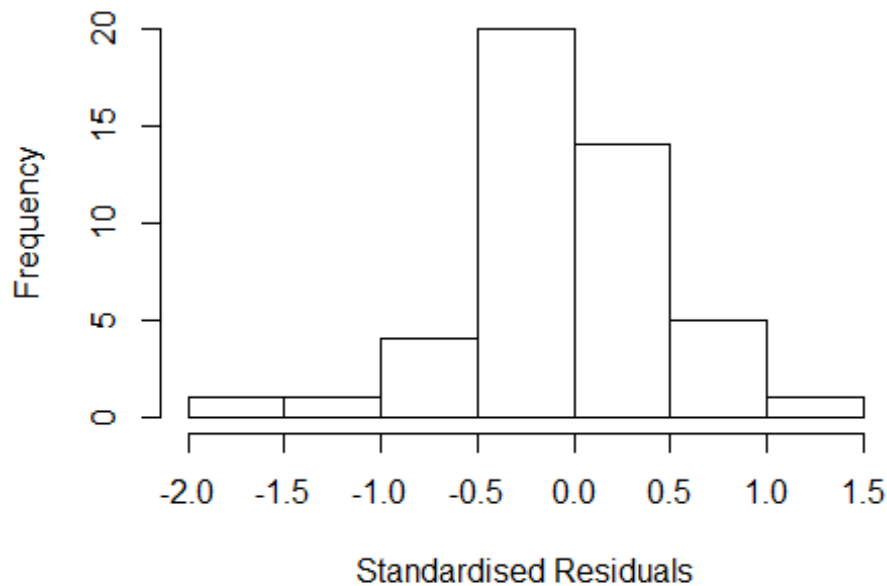
## Year2006          -0.9427      0.0941  -10.01  3.1e-11 ***
## Year2007          -0.4863      0.1057   -4.60  6.7e-05 ***
## Year2008          -0.5745      0.1711   -3.36  0.00209 **
## Year2009          -0.2195      0.2011   -1.09  0.28342
## Year2010          -0.2980      0.0000    -Inf < 2e-16 ***
## Year2011          -0.9309      0.2031   -4.58  7.1e-05 ***
## Year2012          -0.5701      0.7091   -0.80  0.42751
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.553
## Multiple R-squared:  0.479, Adjusted R-squared:  0.243
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~ = 1. The remaining 37 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.387  0.885  0.968  0.918  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      2.17e-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 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
## -1.4823 -0.2832 -0.0022  0.3263  1.0981
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.162      0.000    Inf < 2e-16 ***
## LastAuthorFemale1 -0.206      0.277   -0.74  0.46206
## Year1998         -1.078      0.288   -3.75  0.00073 ***
## Year1999         -0.956      0.312   -3.07  0.00445 **
## Year2000         -1.862      0.194   -9.62  8.1e-11 ***
## Year2002         -0.885      0.367   -2.41  0.02199 *
## Year2004         -0.569      0.327   -1.74  0.09180 .
## Year2005         -0.559      0.146   -3.83  0.00059 ***
```

```

## Year2006          -0.678      0.211   -3.21  0.00304 **
## Year2007          -0.416      0.141   -2.95  0.00604 **
## Year2008          -0.422      0.212   -1.99  0.05584 .
## Year2009          -0.181      0.129   -1.41  0.16887
## Year2010          -0.298      0.000    -Inf < 2e-16 ***
## Year2011          -0.932      0.201   -4.64  6.1e-05 ***
## Year2012          -0.680      0.701   -0.97  0.33973
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.624
## Multiple R-squared:  0.432, Adjusted R-squared:  0.175
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.552  0.913  0.964  0.936  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      2.17e-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 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
##    1    1    1    2    4    5    5    1    4    7    4    8    9    9    19
## 2011 2012
##   13   15
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    1    0    0    1    3    2    1    1    2    2    1    5    5    8

```

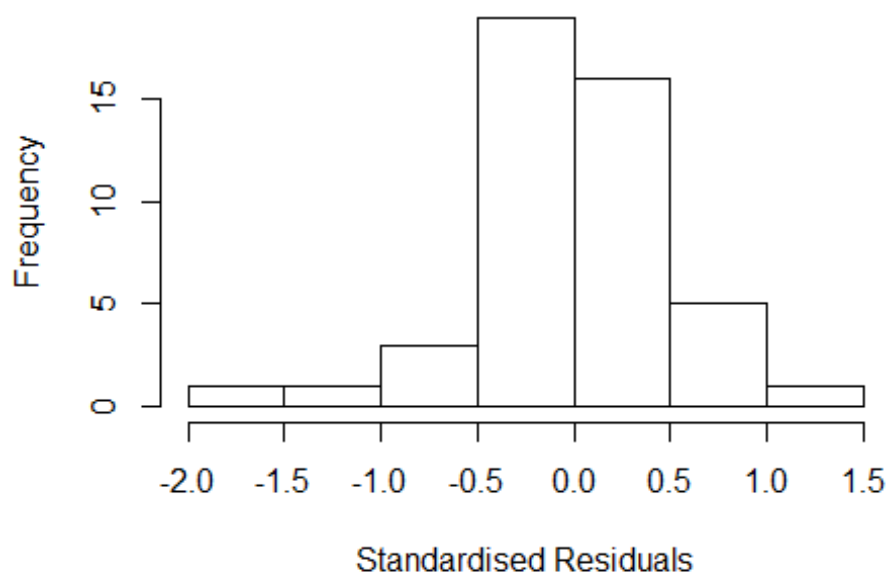
```
## 2011 2012
##    6    9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    1    0    0    1    1    1    1    1    1    1    1    4    5    7
## 2011 2012
##    6    8
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3"
## [1] "Male first author team size 2018 geometric mean: 3.03505497531656"

## 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 = 4, 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: 3.03505497531656"

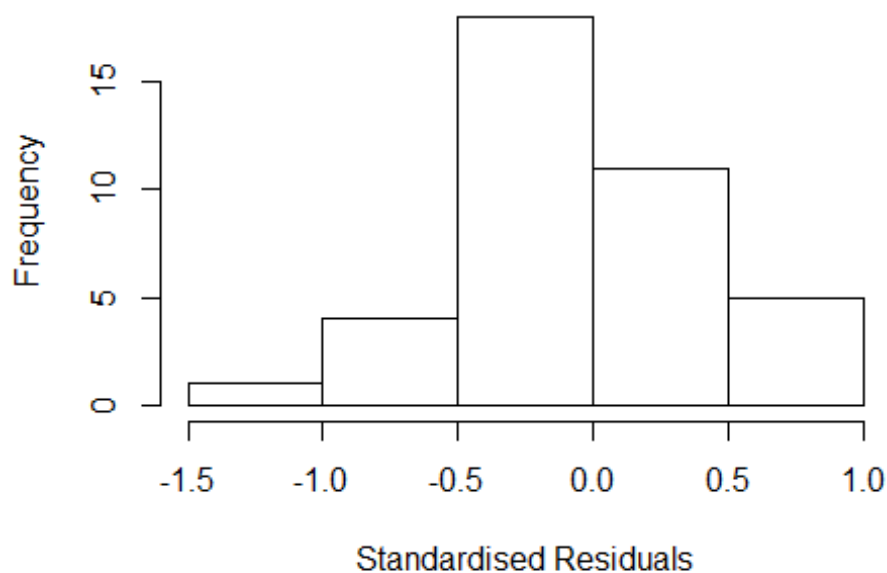
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

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] "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 7.047e+15  1      8.395e+07
## LastAuthorFemale  7.313e+15  1      8.552e+07
## Year              3.109e+16 13      4.309e+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
```

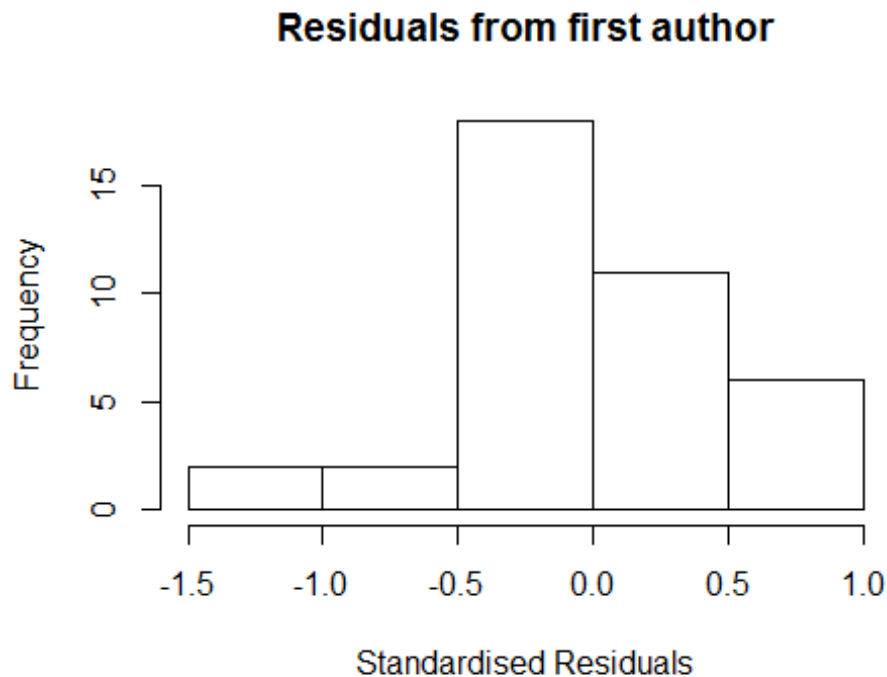
```

## -1.24e+00 -1.12e-01 4.44e-16 2.09e-01 8.77e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.4236      0.4151   3.43  0.0023 **
## FirstAuthorFemale1 -0.5621      0.2977  -1.89  0.0717 .
## LastAuthorFemale1  0.5155      0.3107   1.66  0.1107
## Year2000           0.0134      0.4151   0.03  0.9745
## Year2001           0.6994      0.4151   1.69  0.1055
## Year2002          -1.0656      0.4151  -2.57  0.0172 *
## Year2003          -0.2606      0.4151  -0.63  0.5363
## Year2004          -0.0786      0.4151  -0.19  0.8514
## Year2005           0.0784      0.4151   0.19  0.8519
## Year2006          -0.3696      0.4151  -0.89  0.3824
## Year2007          -0.4626      0.4151  -1.11  0.2765
## Year2008           0.0193      0.3943   0.05  0.9614
## Year2009          -0.2242      0.4494  -0.50  0.6225
## Year2010           0.1376      0.4476   0.31  0.7613
## Year2011           0.1284      0.3844   0.33  0.7414
## Year2012           0.1678      0.4210   0.40  0.6939
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.512
## Multiple R-squared:  0.363, Adjusted R-squared:  -0.0525
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 23 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.536  0.859  0.951  0.902  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      2.56e-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 -1.449e+14 1 NaN
## Year -1.449e+14 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
## -1.34e+00 -1.63e-01 1.78e-15 1.36e-01 8.86e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.872 0.230 8.13 2.3e-08 ***
## FirstAuthorFemale1 -0.495 0.230 -2.15 0.04178 *
## Year2000 -0.435 0.230 -1.89 0.07090 .
## Year2001 0.251 0.230 1.09 0.28620
## Year2002 -1.514 0.230 -6.58 8.4e-07 ***
## Year2003 -0.709 0.230 -3.08 0.00512 **
## Year2004 -0.527 0.230 -2.29 0.03113 *
## Year2005 -0.370 0.230 -1.61 0.12101
## Year2006 -0.818 0.230 -3.55 0.00161 **
```

```

## Year2007          -0.911      0.230   -3.96  0.00059 ***
## Year2008          -0.343      0.135   -2.53  0.01838 *
## Year2009          -0.531      0.569   -0.93  0.35975
## Year2010          -0.310      0.288   -1.08  0.29181
## Year2011          -0.222      0.283   -0.78  0.44216
## Year2012          -0.290      0.259   -1.12  0.27406
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.498
## Multiple R-squared:  0.316, Adjusted R-squared:  -0.0838
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 24 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.449  0.847  0.937  0.892  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      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 -3.833e+15  1      NaN
## Year              -3.833e+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 ~ 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 -4.84e-02  2.22e-16  2.49e-01  9.30e-01

```

```

##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.904      0.434    2.08  0.0483 *
## LastAuthorFemale1 0.473      0.434    1.09  0.2871
## Year2000          0.533      0.434    1.23  0.2318
## Year2001          1.219      0.434    2.81  0.0098 **
## Year2002         -0.546      0.434   -1.26  0.2210
## Year2003          0.259      0.434    0.60  0.5566
## Year2004          0.441      0.434    1.02  0.3202
## Year2005          0.598      0.434    1.38  0.1814
## Year2006          0.150      0.434    0.35  0.7328
## Year2007          0.057      0.434    0.13  0.8966
## Year2008          0.290      0.562    0.52  0.6109
## Year2009          0.309      0.416    0.74  0.4647
## Year2010          0.657      0.469    1.40  0.1740
## Year2011          0.554      0.393    1.41  0.1717
## Year2012          0.634      0.465    1.36  0.1857
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.508
## Multiple R-squared:  0.291, Adjusted R-squared:  -0.123
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 24 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.589  0.815   0.947   0.894   0.975   0.999
## Algorithmic parameters:
##           tuning.chi           bb           tuning.psi           refine.tol
##           1.55e+00           5.00e-01           4.69e+00           1.00e-07
##           rel.tol           solve.tol           eps.outlier           eps.x
##           1.00e-07           1.00e-07           2.56e-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 2216"

```

```

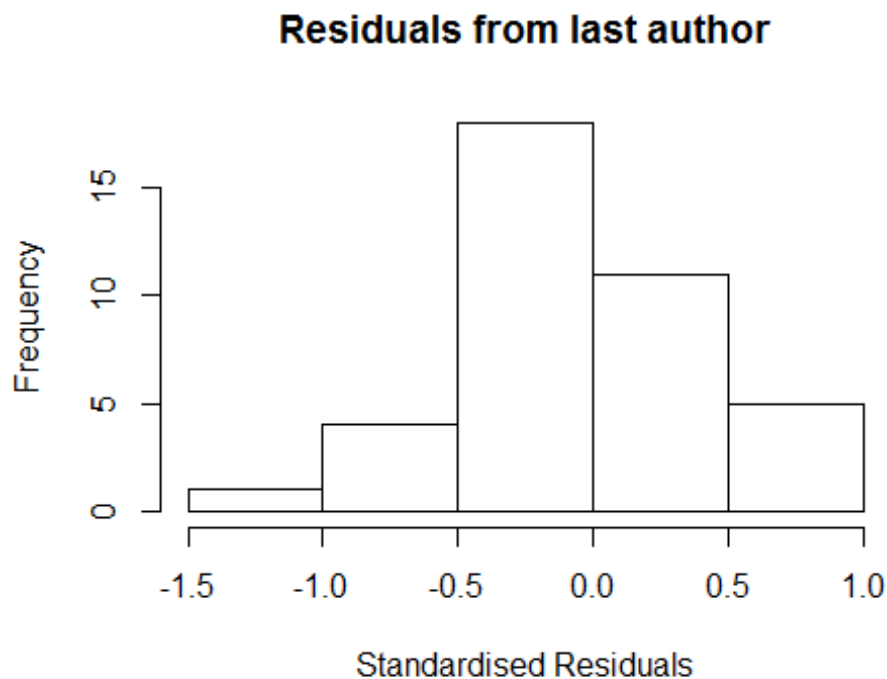
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1998 1999 2000 2003 2006 2007 2008 2009 2011
##    1    2    2    1    2    1    1    1    1
##
## 1998 1999 2000 2003 2006 2007 2008 2009 2011
##    0    1    1    1    2    1    1    1    1
##
## 1998 1999 2000 2003 2006 2007 2008 2009 2011
##    0    1    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: 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: 9"
## [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
##   15   27   13    8   16    7   17   19   18   21   33   42   38   50   16
## 2011 2012
##   41   47
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    8   14    7    6    1    4    9   10   12   16   18   20   18   38   10
## 2011 2012
##   28   30
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5   12    7    5    1    2    9   10   10   11   16   16   17   34    9
## 2011 2012
##   28   26
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.69601230919456"
## [1] "Male first author team size 2018 geometric mean: 2.39492974229681"

```

```
## Warning in wilcox.test.default(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 = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.6207413942089"
## [1] "Male last author team size 2018 geometric mean: 2.48441690882754"

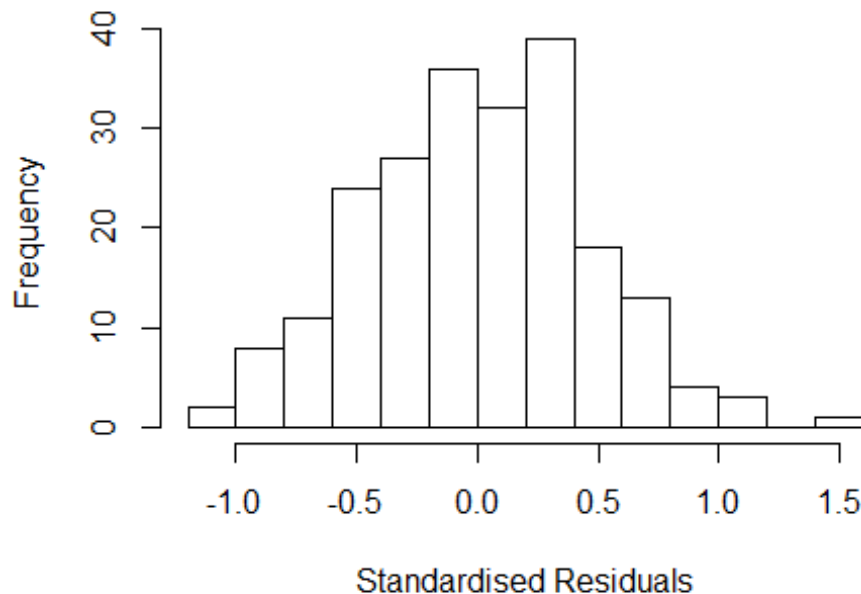
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 21, p-value = 0.7
## 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 7.43 1 2.726
```

| | | | |
|---------------------|--------|----|-------|
| ## LastAuthorFemale | 16.30 | 1 | 4.038 |
| ## UniqueAuthors | 18.05 | 4 | 1.436 |
| ## Year | 152.49 | 16 | 1.170 |

Residuals from first and last author and team size



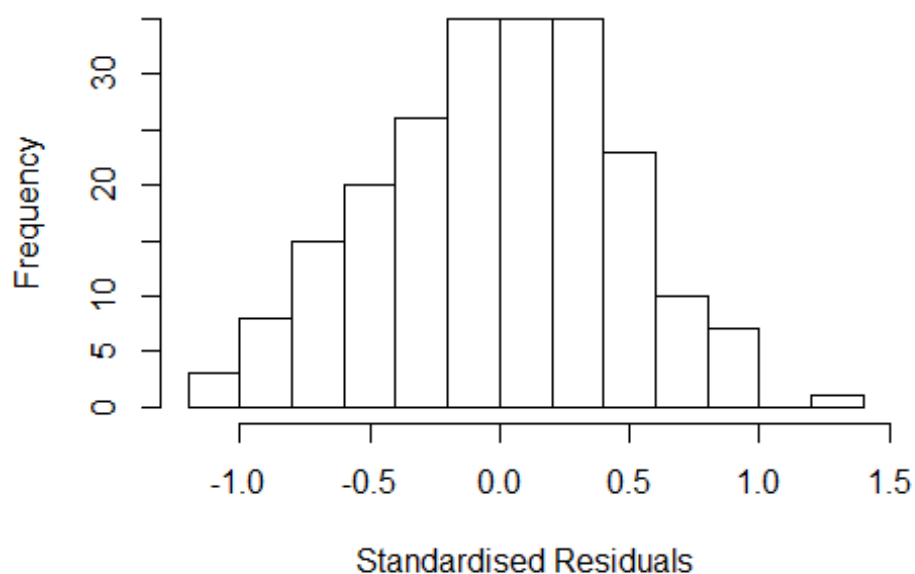
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.047853 -0.317749  0.000344  0.322063  1.439256
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1367    0.2197   5.17 5.7e-07 ***
## FirstAuthorFemale1  0.0357    0.0830   0.43  0.668
## LastAuthorFemale1 -0.1494    0.1077  -1.39  0.167
## UniqueAuthors2     0.0324    0.1004   0.32  0.747
## UniqueAuthors3     0.2435    0.1084   2.25  0.026 *
## UniqueAuthors4     0.0737    0.1287   0.57  0.567
```

```

## UniqueAuthors5      0.1429      0.1208      1.18      0.238
## Year1997             -0.0582      0.2508     -0.23      0.817
## Year1998             -0.1378      0.2483     -0.55      0.580
## Year1999             -0.2531      0.2284     -1.11      0.269
## Year2000              0.1950      0.2599      0.75      0.454
## Year2001              0.0763      0.2372      0.32      0.748
## Year2002             -0.5405      0.2654     -2.04      0.043 *
## Year2003             -0.2562      0.2673     -0.96      0.339
## Year2004             -0.1561      0.2711     -0.58      0.565
## Year2005             -0.2011      0.2731     -0.74      0.462
## Year2006             -0.3191      0.2521     -1.27      0.207
## Year2007             -0.1238      0.2531     -0.49      0.625
## Year2008             -0.1213      0.2475     -0.49      0.625
## Year2009             -0.1552      0.2361     -0.66      0.512
## Year2010             -0.0996      0.2585     -0.39      0.701
## Year2011              0.0226      0.2585      0.09      0.930
## Year2012              0.0660      0.2741      0.24      0.810
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.484
## Multiple R-squared:  0.137, Adjusted R-squared:  0.0394
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 199 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.356  0.889   0.954   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      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 7.222 1      2.687
## LastAuthorFemale 12.962 1      3.600
## Year              26.702 16      1.108

```

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
## -1.1900 -0.3041  0.0219  0.3256  1.3247
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.21055    0.28039   4.32 2.5e-05 ***
## FirstAuthorFemale1  0.04044    0.08565   0.47  0.637
## LastAuthorFemale1 -0.15243    0.10387  -1.47  0.144
## Year1997         -0.09468    0.30405  -0.31  0.756
## Year1998         -0.13359    0.29968  -0.45  0.656
## Year1999         -0.26572    0.28939  -0.92  0.360
## Year2000          0.11944    0.31335   0.38  0.703
## Year2001          0.00245    0.29433   0.01  0.993
## Year2002         -0.56847    0.32448  -1.75  0.081 .
## Year2003         -0.29473    0.31673  -0.93  0.353
## Year2004         -0.14996    0.33164  -0.45  0.652
## Year2005         -0.21866    0.32065  -0.68  0.496
```

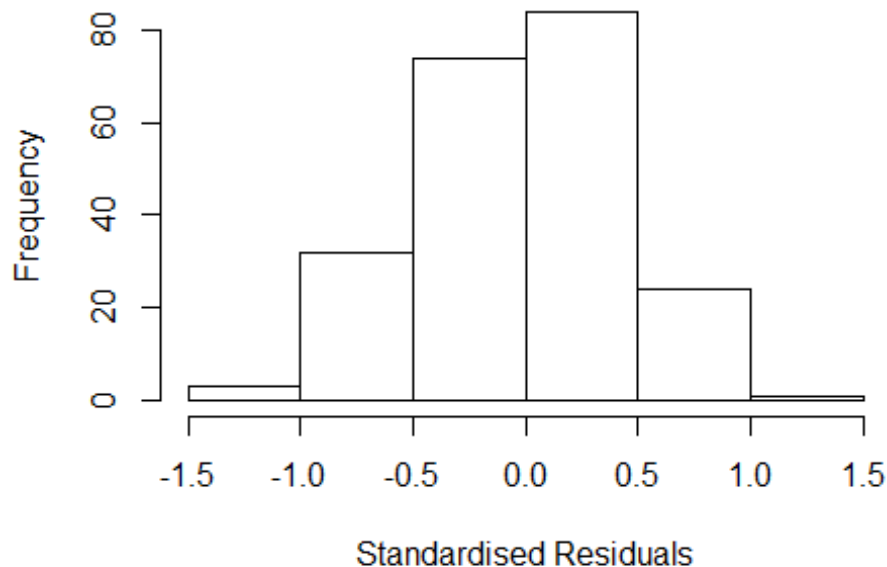


```

## Year2006      -0.30212      0.30026      -1.01      0.316
## Year2007      -0.11411      0.29771      -0.38      0.702
## Year2008      -0.10163      0.30213      -0.34      0.737
## Year2009      -0.14749      0.29391      -0.50      0.616
## Year2010      -0.09112      0.30804      -0.30      0.768
## Year2011       0.09141      0.30569       0.30      0.765
## Year2012       0.10679      0.32409       0.33      0.742
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.478
## Multiple R-squared:  0.109, Adjusted R-squared:  0.0289
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 198 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.423  0.878  0.949  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      4.59e-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 7.713 1      2.777
## Year              7.713 16      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
## -1.31287 -0.29754 0.00882 0.32721 1.35276
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.18965 0.25793 4.61 7.1e-06 ***
## FirstAuthorFemale1 0.02174 0.08830 0.25 0.806
## Year1997 -0.10966 0.27655 -0.40 0.692
## Year1998 -0.13331 0.27732 -0.48 0.631
## Year1999 -0.24112 0.26793 -0.90 0.369
## Year2000 0.00661 0.27178 0.02 0.981
## Year2001 0.02335 0.27303 0.09 0.932
## Year2002 -0.54316 0.30574 -1.78 0.077 .
## Year2003 -0.28226 0.29764 -0.95 0.344
## Year2004 -0.15247 0.30672 -0.50 0.620
## Year2005 -0.22102 0.29610 -0.75 0.456
## Year2006 -0.28999 0.27791 -1.04 0.298
```

```

## Year2007          -0.11917    0.27497   -0.43    0.665
## Year2008          -0.10162    0.27924   -0.36    0.716
## Year2009          -0.14946    0.27144   -0.55    0.583
## Year2010          -0.06465    0.28763   -0.22    0.822
## Year2011           0.10148    0.28366    0.36    0.721
## Year2012           0.09959    0.31020    0.32    0.749
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.472
## Multiple R-squared:  0.0998, Adjusted R-squared:  0.0233
## 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.
##   0.392  0.873  0.948  0.909  0.984  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      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 12.72  1          3.566
## Year              12.72 16          1.083
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.1757 -0.3079  0.0177  0.3372  1.3125

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.21562    0.28238   4.30 2.6e-05 ***
## LastAuthorFemale1 -0.14649    0.10397  -1.41  0.160
## Year1997        -0.09877    0.30621  -0.32  0.747
## Year1998        -0.13383    0.30095  -0.44  0.657
## Year1999        -0.26283    0.29179  -0.90  0.369
## Year2000         0.14887    0.30557   0.49  0.627
## Year2001        -0.00262    0.29620  -0.01  0.993
## Year2002        -0.56544    0.32684  -1.73  0.085 .
## Year2003        -0.28452    0.31432  -0.91  0.366
## Year2004        -0.13284    0.32829  -0.40  0.686
## Year2005        -0.20190    0.31567  -0.64  0.523
## Year2006        -0.30384    0.30113  -1.01  0.314
## Year2007        -0.10297    0.29905  -0.34  0.731
## Year2008        -0.09964    0.30259  -0.33  0.742
## Year2009        -0.14827    0.29529  -0.50  0.616
## Year2010        -0.08446    0.30946  -0.27  0.785
## Year2011         0.10655    0.30379   0.35  0.726
## Year2012         0.11384    0.32488   0.35  0.726
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.494
## Multiple R-squared:  0.108, Adjusted R-squared:  0.0321
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 197 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.460  0.886  0.952  0.916  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      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 2301"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1997 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    3    1    1    4    4    3    2    5    5    3    4    6    6    6
##
## 1997 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    0    1    2    4    3    2    5    4    3    2    4    5    5
##
## 1997 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    0    1    2    4    3    2    5    4    3    2    4    5    5
## [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: 1.60935392754712"

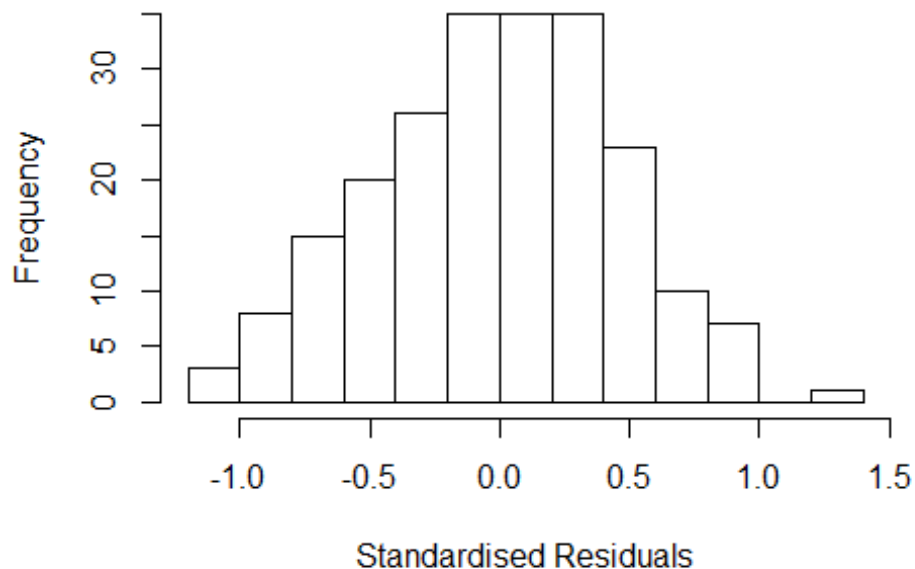
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## 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: 1"
## [1] "Male last author team size 2018 geometric mean: 1.60935392754712"

## 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 = 0.6
## 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 -3.436e+15  1      NaN
## LastAuthorFemale  -2.293e+15  1      NaN
## UniqueAuthors    -5.519e+14  3      NaN
## Year              -5.633e+29 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 +
UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
```

```

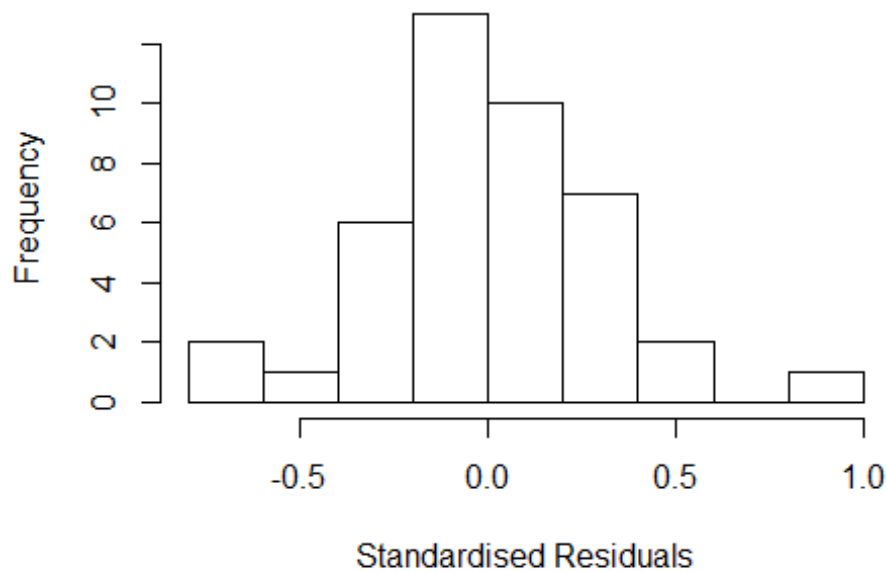
##           Min           1Q           Median           3Q           Max
## -7.88e-01 -1.90e-01  4.16e-17  1.98e-01  9.99e-01
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.1775     0.1623   13.42 2.3e-12 ***
## FirstAuthorFemale1 -0.3116     0.1804   -1.73 0.09754 .
## LastAuthorFemale1  0.0419     0.1476    0.28 0.77885
## UniqueAuthors2     -0.1025     0.1623   -0.63 0.53397
## UniqueAuthors3     -0.0751     0.1408   -0.53 0.59877
## UniqueAuthors4     -0.0781     0.2856   -0.27 0.78695
## Year1998           -1.1035     0.1623   -6.80 6.2e-07 ***
## Year2000            -0.6988     0.2340   -2.99 0.00660 **
## Year2002            -0.7280     0.2395   -3.04 0.00583 **
## Year2003            -0.8988     0.2047   -4.39 0.00021 ***
## Year2004            -0.7965     0.2248   -3.54 0.00173 **
## Year2005            -0.5239     0.1743   -3.01 0.00630 **
## Year2006            -0.8339     0.1717   -4.86 6.7e-05 ***
## Year2007            -0.9926     0.2445   -4.06 0.00049 ***
## Year2008            -0.6727     0.1930   -3.49 0.00199 **
## Year2009            -0.9526     0.2383   -4.00 0.00057 ***
## Year2010            -0.7937     0.1682   -4.72 9.4e-05 ***
## Year2011            -0.5561     0.2588   -2.15 0.04239 *
## Year2012            -1.1203     0.4194   -2.67 0.01365 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.433
## Multiple R-squared:  0.367, Adjusted R-squared:  -0.129
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 37 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.574  0.951  0.973   0.947  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      2.38e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"

```

```
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"

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

Residuals from first and last author and team size



```
##               GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -3.074e+15  1      NaN
## LastAuthorFemale -2.084e+15  1      NaN
## Year              -9.325e+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
## -7.55e-01 -2.08e-01  1.61e-15  2.12e-01  1.06e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
```



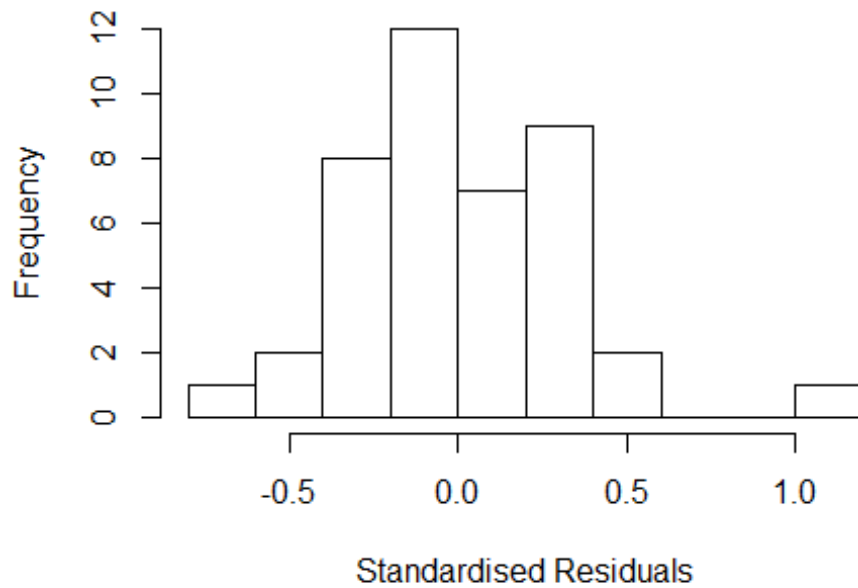
```

## (Intercept)          2.08e+00  0.00e+00      Inf < 2e-16 ***
## FirstAuthorFemale1 -3.32e-01  1.54e-01 -2.15e+00  0.04089 *
## LastAuthorFemale1   8.90e-02  1.27e-01  7.00e-01  0.48962
## Year1998            -1.00e+00  2.06e-08 -4.85e+07 < 2e-16 ***
## Year2000            -6.23e-01  1.77e-01 -3.52e+00  0.00162 **
## Year2002            -6.25e-01  1.79e-01 -3.50e+00  0.00171 **
## Year2003            -8.62e-01  1.47e-01 -5.88e+00  3.3e-06 ***
## Year2004            -7.10e-01  1.62e-01 -4.40e+00  0.00016 ***
## Year2005            -4.86e-01  1.84e-01 -2.64e+00  0.01375 *
## Year2006            -7.48e-01  6.63e-02 -1.13e+01  1.6e-11 ***
## Year2007            -9.39e-01  2.18e-01 -4.31e+00  0.00021 ***
## Year2008            -6.41e-01  1.73e-01 -3.71e+00  0.00099 ***
## Year2009            -9.05e-01  1.96e-01 -4.62e+00  9.3e-05 ***
## Year2010            -7.51e-01  1.78e-01 -4.22e+00  0.00026 ***
## Year2011            -5.05e-01  2.05e-01 -2.47e+00  0.02041 *
## Year2012            -1.08e+00  5.02e-01 -2.15e+00  0.04146 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.386
## Multiple R-squared:  0.375, Adjusted R-squared:  0.0151
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 36 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.433  0.939  0.964  0.931  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      2.38e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##   nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##   trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"

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

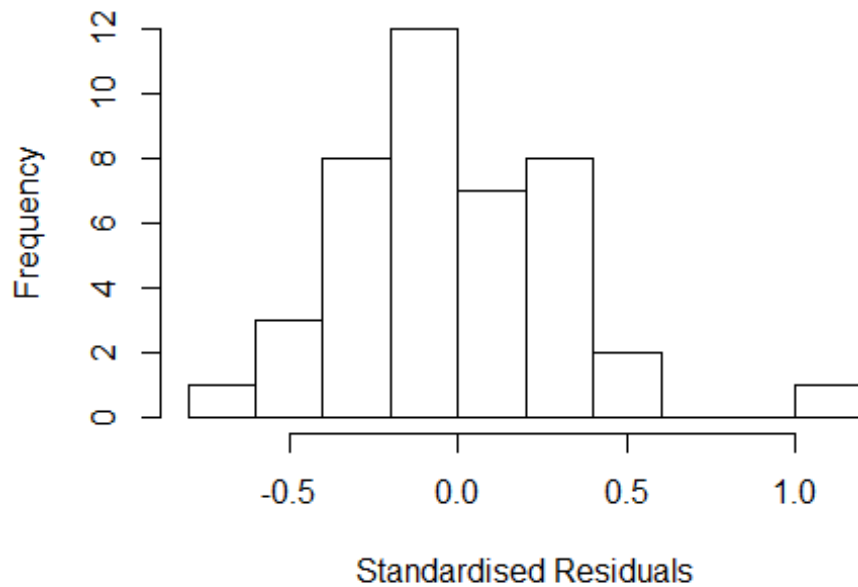
```

Residuals from first and last author



| ## | | GVIF | Df | $GVIF^{(1/(2*Df))}$ |
|----|-------------------|-----------|----|---------------------|
| ## | FirstAuthorFemale | 2.185e+15 | 1 | 4.674e+07 |
| ## | Year | 2.185e+15 | 13 | 3.890e+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
## -0.7087 -0.2242 -0.0053 0.2104 1.0736
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.08e+00 0.00e+00 Inf < 2e-16 ***
## FirstAuthorFemale1 -2.74e-01 1.61e-01 -1.70e+00 0.09988 .
## Year1998 -1.00e+00 1.90e-08 -5.27e+07 < 2e-16 ***
## Year2000 -5.92e-01 1.61e-01 -3.69e+00 0.00101 **
## Year2002 -6.25e-01 1.81e-01 -3.46e+00 0.00182 **
## Year2003 -8.19e-01 1.46e-01 -5.60e+00 6.1e-06 ***
## Year2004 -6.90e-01 1.53e-01 -4.51e+00 0.00011 ***
## Year2005 -4.71e-01 1.94e-01 -2.43e+00 0.02208 *
## Year2006 -7.54e-01 6.94e-02 -1.08e+01 2.4e-11 ***
## Year2007 -9.50e-01 2.42e-01 -3.93e+00 0.00053 ***
## Year2008 -6.40e-01 1.66e-01 -3.85e+00 0.00066 ***
## Year2009 -9.18e-01 1.83e-01 -5.01e+00 3.0e-05 ***
## Year2010 -7.60e-01 1.85e-01 -4.10e+00 0.00034 ***
## Year2011 -5.09e-01 2.10e-01 -2.43e+00 0.02227 *
## Year2012 -1.09e+00 5.46e-01 -2.00e+00 0.05560 .
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.36
## Multiple R-squared: 0.383, Adjusted R-squared: 0.0625
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~ = 1. The remaining 36 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.354 0.931 0.961 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 2.38e-03 1.82e-12
## warn.limit.reject warn.limit.meanrw
## 5.00e-01 5.00e-01
## nResample max.it best.r.s k.fast.s k.max maxit.scale
## 500 50 2 1 1000 200
```

```

##      trace.lev      mts  compute.rd
##           0          1000          0
##           psi          subsampling          cov
##           "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.333e+14 1 1.155e+07
## Year 1.333e+14 13 3.494e+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
## -0.9235 -0.2012  0.0362  0.1989  1.0525
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    2.08e+00   1.70e-08  1.22e+08 < 2e-16 ***
## LastAuthorFemale1 -8.00e-02   1.56e-01 -5.10e-01  0.61317
## Year1998        -1.00e+00   2.05e-08 -4.88e+07 < 2e-16 ***
## Year2000        -7.86e-01   1.56e-01 -5.02e+00  2.9e-05 ***
## Year2002        -6.25e-01   1.79e-01 -3.50e+00  0.00165 **
## Year2003        -7.80e-01   1.69e-01 -4.63e+00  8.3e-05 ***
## Year2004        -8.16e-01   1.82e-01 -4.47e+00  0.00013 ***
## Year2005        -5.67e-01   1.24e-01 -4.57e+00  9.7e-05 ***
## Year2006        -8.45e-01   5.84e-02 -1.45e+01  3.1e-14 ***
## Year2007        -1.04e+00   2.57e-01 -4.04e+00  0.00040 ***
## Year2008        -8.60e-01   1.07e-01 -8.04e+00  1.2e-08 ***
## Year2009        -1.15e+00   9.77e-02 -1.18e+01  3.7e-12 ***
## Year2010        -8.82e-01   2.21e-01 -4.00e+00  0.00044 ***
## Year2011        -6.18e-01   1.48e-01 -4.18e+00  0.00027 ***
## Year2012        -1.07e+00   6.51e-01 -1.65e+00  0.11147
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.385
## Multiple R-squared:  0.301, Adjusted R-squared:  -0.0608
## Convergence in 27 IRWLS iterations
##
## Robustness weights:

```

```

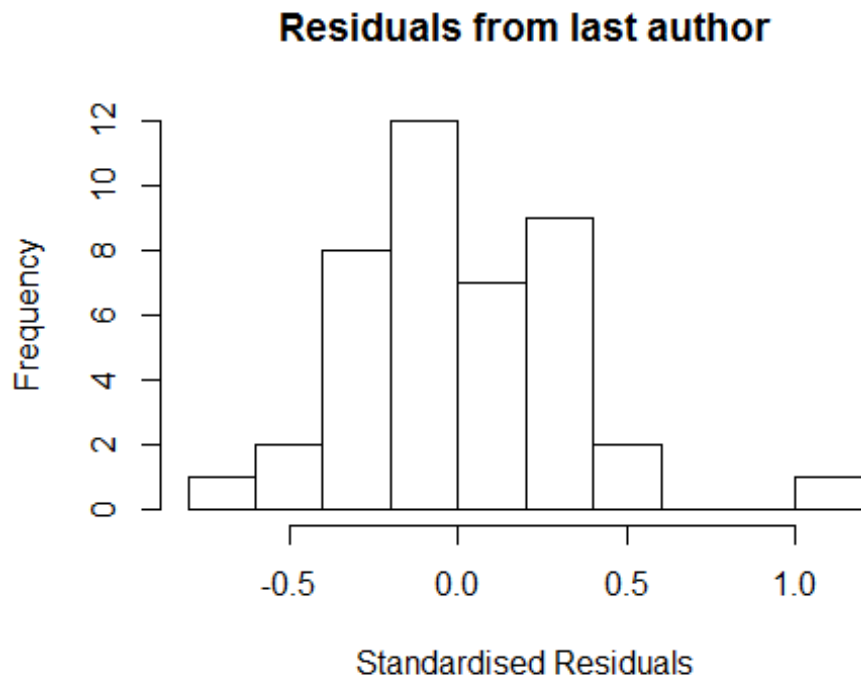
## 5 weights are ~= 1. The remaining 37 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.435  0.915  0.970  0.925  0.992  0.999
## Algorithmic parameters:
##       tuning.chi          bb          tuning.psi          refine.tol
##       1.55e+00          5.00e-01          4.69e+00          1.00e-07
##       rel.tol          solve.tol          eps.outlier          eps.x
##       1.00e-07          1.00e-07          2.38e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##       5.00e-01          5.00e-01
##   nResample          max.it          best.r.s          k.fast.s          k.max maxit.scale
##       500          50          2          1          1000          200
##   trace.lev          mts          compute.rd
##       0          1000          0
##           psi          subsampling          cov
##           "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 42"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2302"
## [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
##    1    3    1    3    3    1    2    4    3    2    5    5    10    6    9
## 2012
##    12
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    0    0    0    2    1    0    0    3    1    1    3    1    4    1    8
## 2012
##     9
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    0    0    0    1    1    0    0    2    1    1    3    1    3    0    5
## 2012
##     7
## [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"

## 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
```



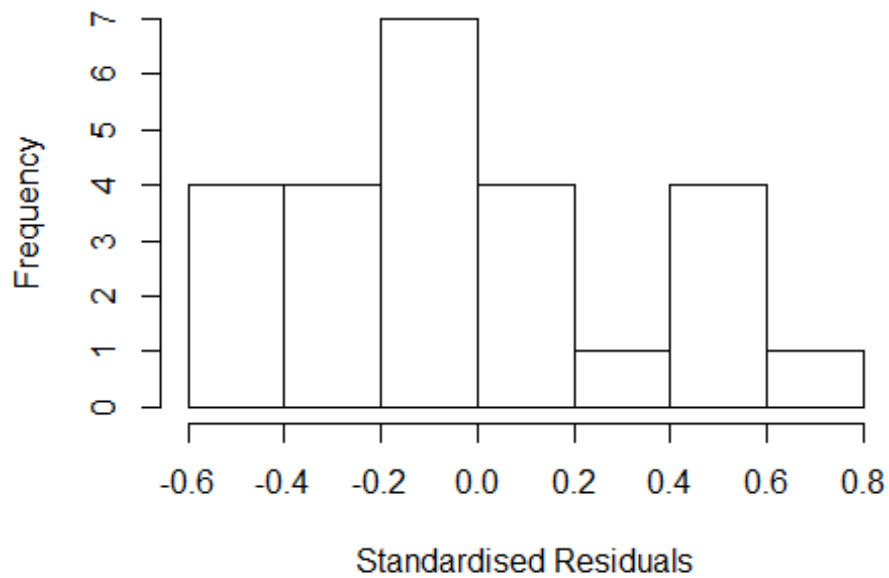
```
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1          NaN
## Year           NaN 9          NaN

## [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

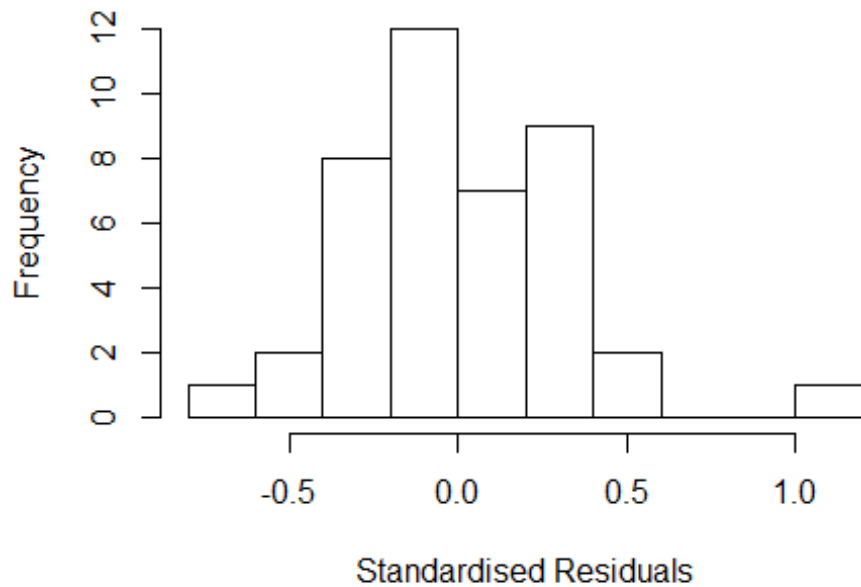
## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

Residuals from first author



| | | | |
|---------------------|------|----|-------------------------|
| ## | GVIF | Df | $GVIF^{1/(2 \cdot Df)}$ |
| ## LastAuthorFemale | NaN | 1 | NaN |
| ## Year | NaN | 9 | NaN |

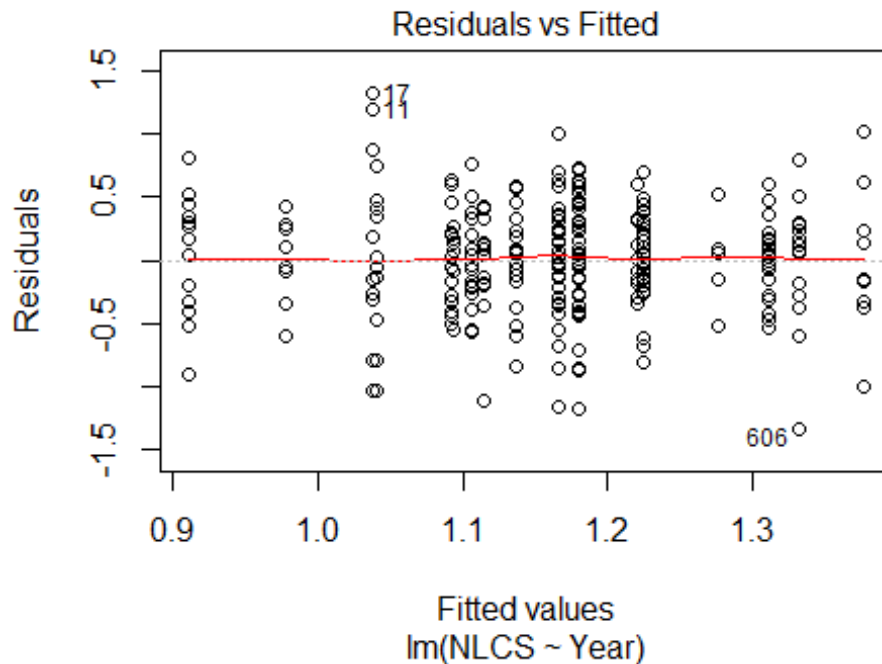
Residuals from last author



```

## [1] "Sample size for the above analysis: 25"
## [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
##    26    23    28    39    42    33    41    29    27    51    48    44    35    55    37
## 2011 2012
##    64    49
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    10     9     5    14     6     8    15    13    11    11    23    22    15    35    22
## 2011 2012
##    40    32
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     9     7     4    12     6     7    11     9     8    10    20    19    13    30    18
## 2011 2012
##    36    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 = 40, df = 16, p-value = 8e-04

```

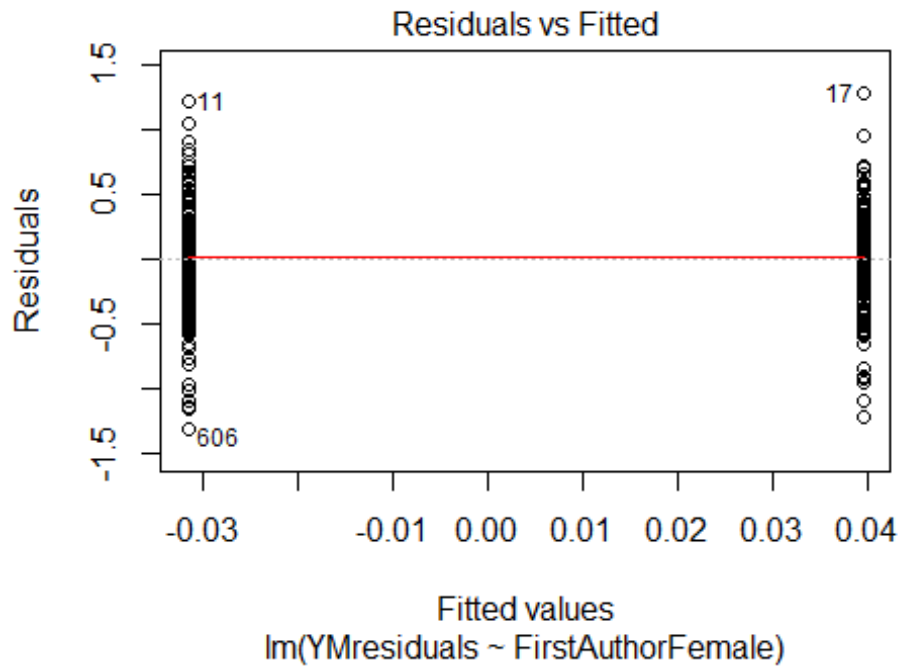
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.1, df = 1, p-value = 0.3

## [1] "Female first author team size 2018 geometric mean: 2.76899210019734"
## [1] "Male first author team size 2018 geometric mean: 2.22064303492292"

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

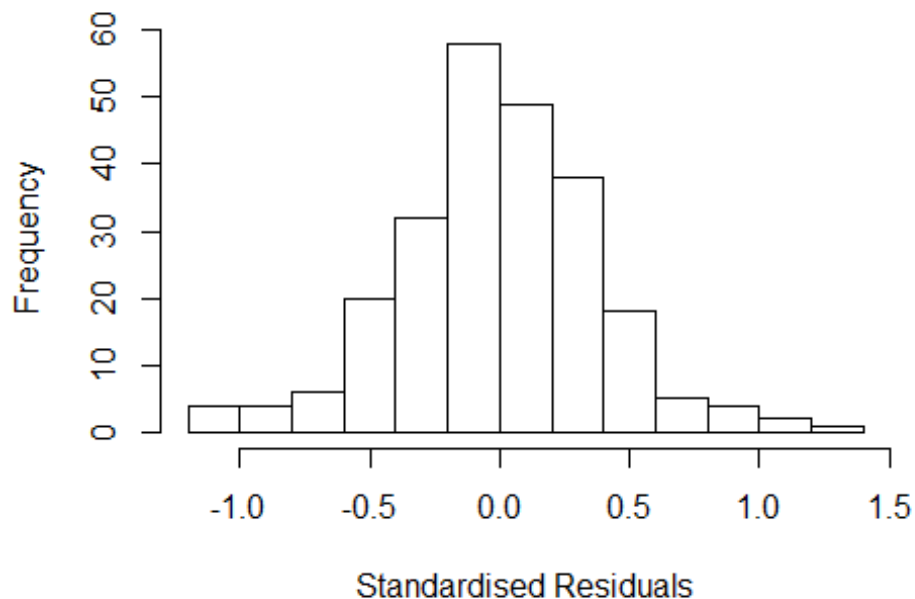
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 26, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2"
## [1] "Male last author team size 2018 geometric mean: 2.73393226647865"

## Warning in wilcox.test.default(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 1.353 1      1.163
## LastAuthorFemale  1.455 1      1.206
## UniqueAuthors    3.435 4      1.167
## Year              5.315 16     1.054
```

Residuals from first and last author and team size



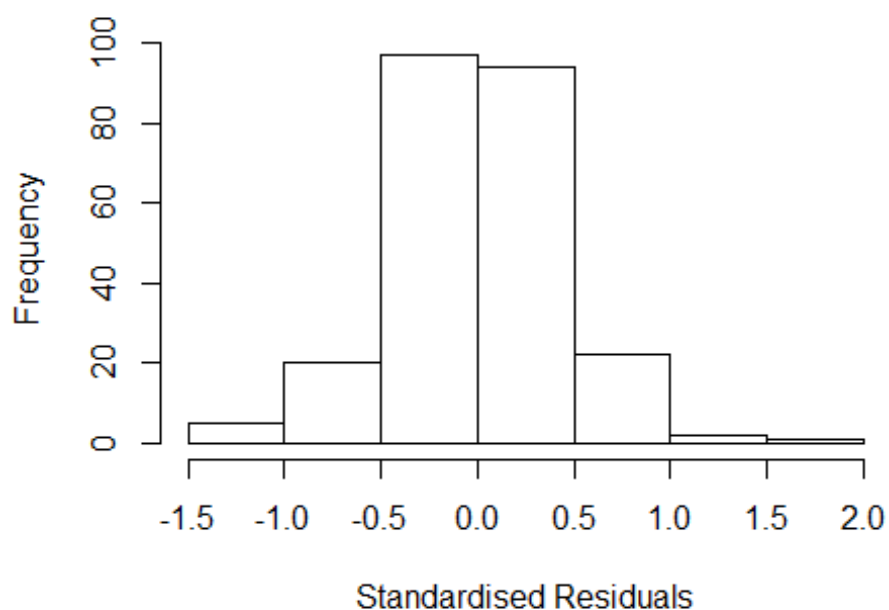
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1659 -0.2299 -0.0112 0.2348 1.3504
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.57064 0.34276 1.66 0.09738 .
## FirstAuthorFemale1 0.00464 0.05129 0.09 0.92801
## LastAuthorFemale1 0.01745 0.07277 0.24 0.81071
## UniqueAuthors2 0.31635 0.12027 2.63 0.00914 **
## UniqueAuthors3 0.49311 0.13066 3.77 0.00021 ***
## UniqueAuthors4 0.37801 0.13576 2.78 0.00583 **
## UniqueAuthors5 0.57097 0.13416 4.26 3.1e-05 ***
## Year1997 0.35147 0.39522 0.89 0.37482
## Year1998 0.28885 0.41910 0.69 0.49142
## Year1999 -0.01019 0.38850 -0.03 0.97910
```

```

## Year2000      0.16035      0.39393      0.41  0.68436
## Year2001      0.17105      0.38048      0.45  0.65347
## Year2002      0.22692      0.37534      0.60  0.54609
## Year2003      0.19689      0.38683      0.51  0.61128
## Year2004      0.22769      0.39772      0.57  0.56758
## Year2005      0.18756      0.40975      0.46  0.64759
## Year2006      0.16708      0.37971      0.44  0.66036
## Year2007      0.04133      0.37871      0.11  0.91320
## Year2008      0.41784      0.39153      1.07  0.28706
## Year2009      0.22250      0.37835      0.59  0.55709
## Year2010      0.28115      0.38521      0.73  0.46626
## Year2011      0.21264      0.39070      0.54  0.58682
## Year2012      0.14414      0.38519      0.37  0.70862
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.374
## Multiple R-squared:  0.224, Adjusted R-squared:  0.145
## Convergence in 28 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 220 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.164  0.875  0.958  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      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 1.305 1      1.143
## LastAuthorFemale  1.407 1      1.186
## Year              1.796 16      1.018

```

Residuals from first and last author

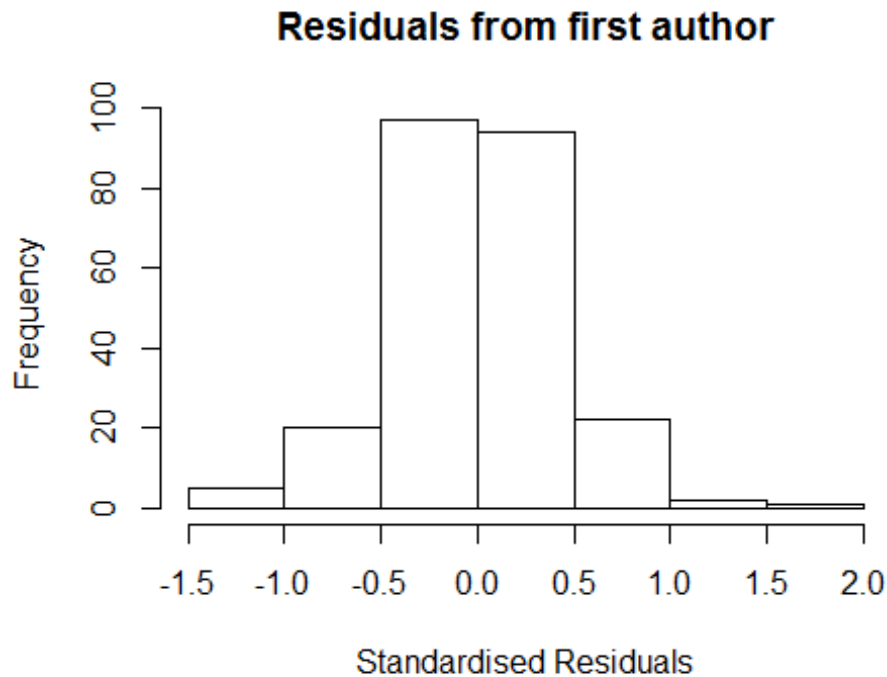


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.38500 -0.22603 -0.00676 0.23208 1.58708
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6479 0.3106 2.09 0.038 *
## FirstAuthorFemale1 0.0513 0.0530 0.97 0.334
## LastAuthorFemale1 -0.0257 0.0759 -0.34 0.735
## Year1997 0.5036 0.3661 1.38 0.170
## Year1998 0.5787 0.3715 1.56 0.121
## Year1999 0.1654 0.3493 0.47 0.636
## Year2000 0.4370 0.3328 1.31 0.190
## Year2001 0.4029 0.3272 1.23 0.220
## Year2002 0.4739 0.3282 1.44 0.150
## Year2003 0.4614 0.3526 1.31 0.192
## Year2004 0.5399 0.3288 1.64 0.102
## Year2005 0.5254 0.3573 1.47 0.143
```

```

## Year2006          0.4599      0.3227      1.42      0.156
## Year2007          0.4221      0.3217      1.31      0.191
## Year2008          0.7371      0.3459      2.13      0.034 *
## Year2009          0.5068      0.3178      1.60      0.112
## Year2010          0.6223      0.3214      1.94      0.054 .
## Year2011          0.5514      0.3326      1.66      0.099 .
## Year2012          0.5339      0.3241      1.65      0.101
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.393
## Multiple R-squared:  0.115, Adjusted R-squared:  0.0434
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 217 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0667 0.8720 0.9520 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      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 1.29 1      1.136
## Year              1.29 16      1.008

```

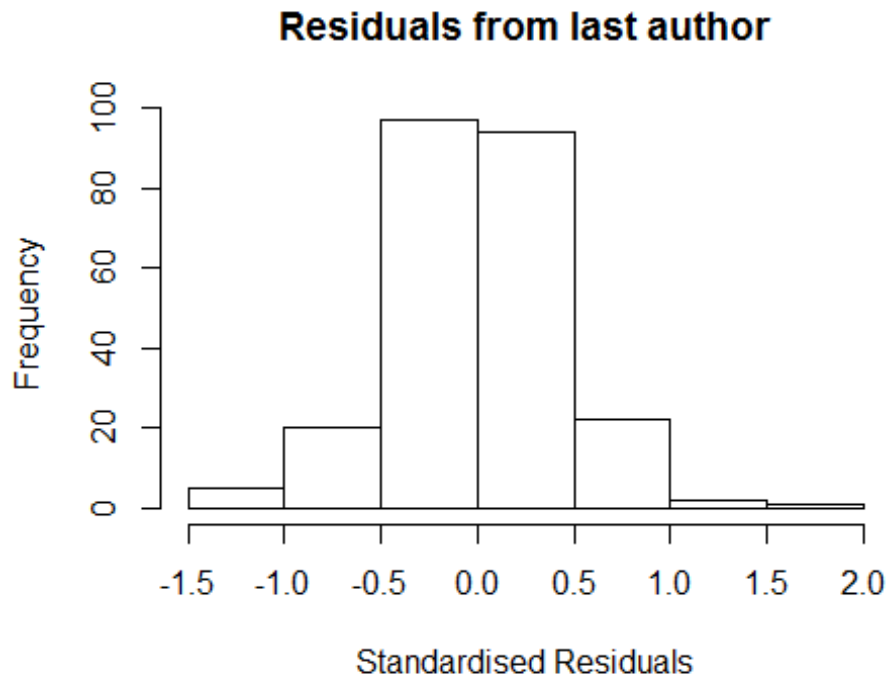


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.37270 -0.22510 -0.00309 0.23091 1.58428
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6507 0.3114 2.09 0.038 *
## FirstAuthorFemale1 0.0503 0.0539 0.93 0.352
## Year1997 0.5024 0.3667 1.37 0.172
## Year1998 0.5763 0.3737 1.54 0.124
## Year1999 0.1630 0.3502 0.47 0.642
## Year2000 0.4345 0.3326 1.31 0.193
## Year2001 0.4004 0.3273 1.22 0.223
## Year2002 0.4716 0.3278 1.44 0.152
## Year2003 0.4550 0.3472 1.31 0.191
## Year2004 0.5377 0.3293 1.63 0.104
## Year2005 0.5216 0.3558 1.47 0.144
## Year2006 0.4544 0.3225 1.41 0.160
```

```

## Year2007          0.4114      0.3189      1.29      0.198
## Year2008          0.7220      0.3409      2.12      0.035 *
## Year2009          0.5001      0.3167      1.58      0.116
## Year2010          0.6169      0.3206      1.92      0.056 .
## Year2011          0.5435      0.3294      1.65      0.100
## Year2012          0.5247      0.3222      1.63      0.105
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.395
## Multiple R-squared:  0.114, Adjusted R-squared:  0.0459
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 219 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0715 0.8770 0.9550 0.8970 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      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 1.364 1      1.168
## Year      1.364 16      1.010

```

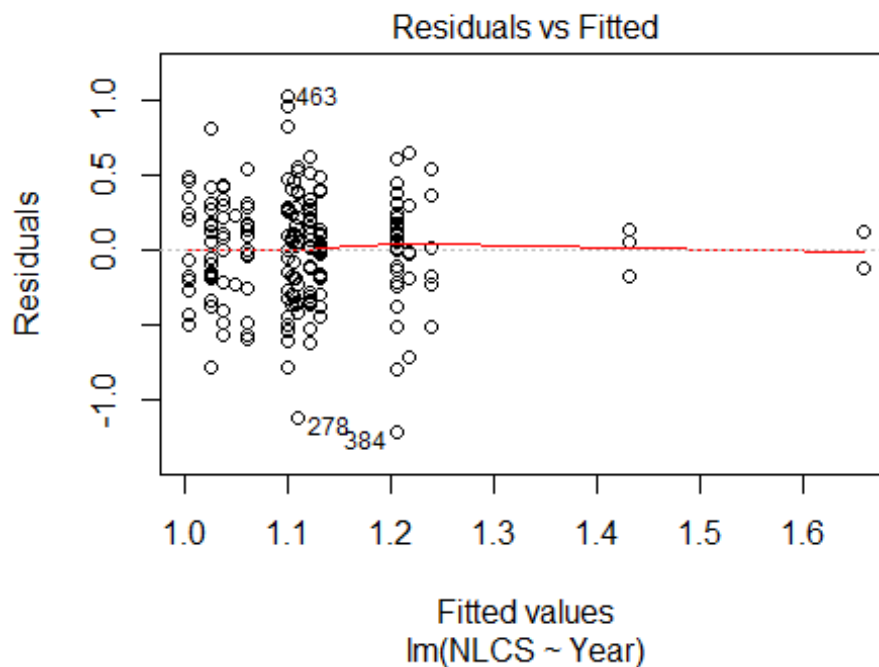
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.39886 -0.22591 0.00919 0.24585 1.58790
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6471 0.3255 1.99 0.048 *
## LastAuthorFemale1 -0.0217 0.0758 -0.29 0.775
## Year1997 0.5387 0.3762 1.43 0.154
## Year1998 0.5981 0.3849 1.55 0.122
## Year1999 0.1770 0.3621 0.49 0.625
## Year2000 0.4612 0.3461 1.33 0.184
## Year2001 0.4191 0.3407 1.23 0.220
## Year2002 0.4968 0.3424 1.45 0.148
## Year2003 0.4937 0.3612 1.37 0.173
## Year2004 0.5609 0.3410 1.64 0.101
## Year2005 0.5537 0.3734 1.48 0.140
## Year2006 0.4766 0.3370 1.41 0.159
```

```

## Year2007          0.4481      0.3329      1.35      0.180
## Year2008          0.7518      0.3608      2.08      0.038 *
## Year2009          0.5370      0.3311      1.62      0.106
## Year2010          0.6527      0.3343      1.95      0.052 .
## Year2011          0.5735      0.3494      1.64      0.102
## Year2012          0.5512      0.3357      1.64      0.102
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.393
## Multiple R-squared:  0.112, Adjusted R-squared:  0.0445
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 216 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0653 0.8730 0.9520 0.8940 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      4.15e-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 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
##   11   13   6   15   10   10   10   15   19   12   25   20   19   29   23
## 2011 2012
##   32   39
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    6    2    6    5    3    6   15   17    8   18   13   10   22   12
## 2011 2012

```

```
## 28 21
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2 5 2 6 5 3 5 14 13 7 16 13 10 21 10
## 2011 2012
## 26 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 = 18, df = 16, p-value = 0.4
```



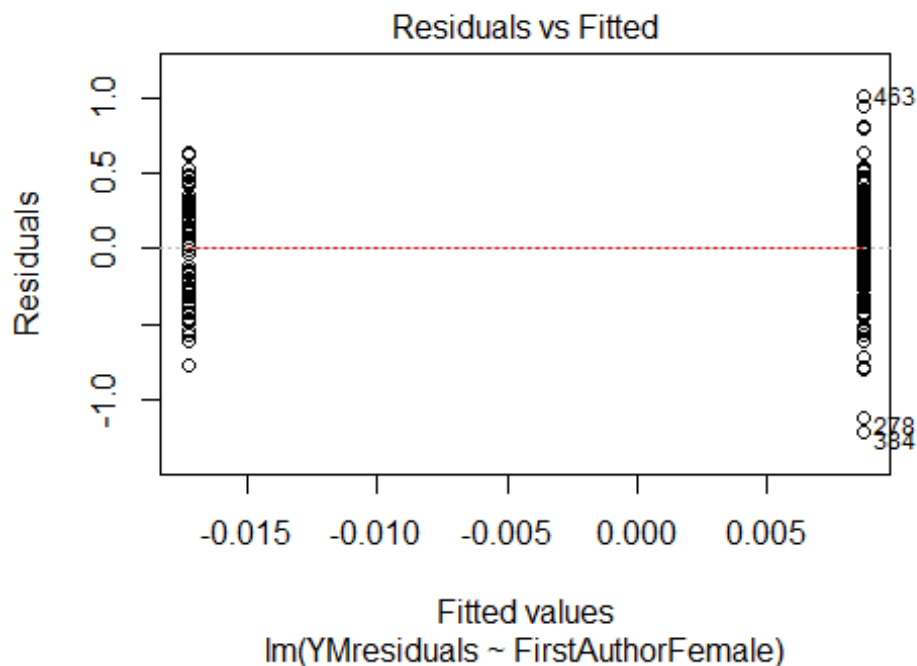
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.47, df = 1, p-value = 0.5

## [1] "Female first author team size 2018 geometric mean: 3.46410161513775"
## [1] "Male first author team size 2018 geometric mean: 4.04282321707986"

## Warning in wilcox.test.default(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: 6"
## [1] "Male last author team size 2018 geometric mean: 3.24534222319921"

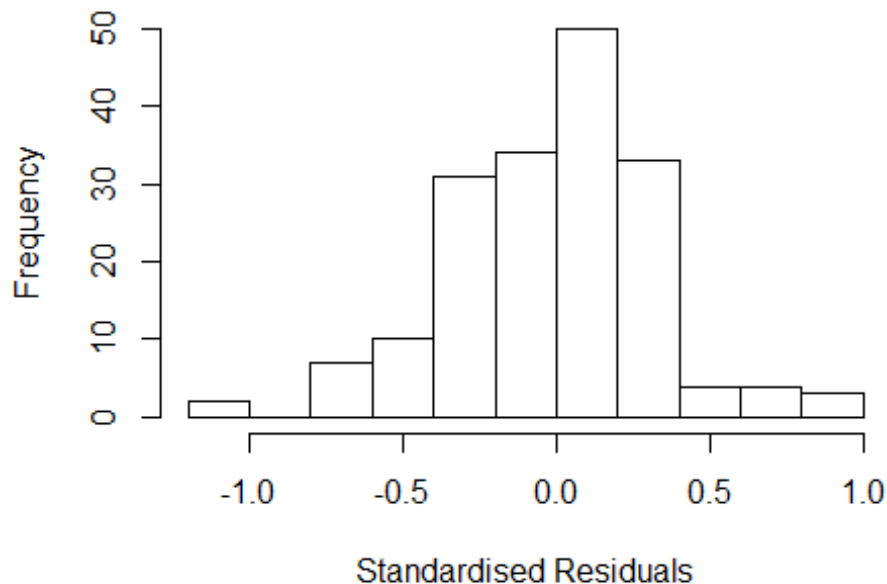
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## 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] "Regression 1: First author gender, last author gender, team size,
## Year as factors"
##
```

| | GVIF | Df | GVIF^(1/(2*Df)) |
|-------------------|--------|----|-----------------|
| FirstAuthorFemale | 1.343 | 1 | 1.159 |
| LastAuthorFemale | 2.042 | 1 | 1.429 |
| UniqueAuthors | 12.970 | 4 | 1.378 |
| Year | 23.965 | 16 | 1.104 |

Residuals from first and last author and team size



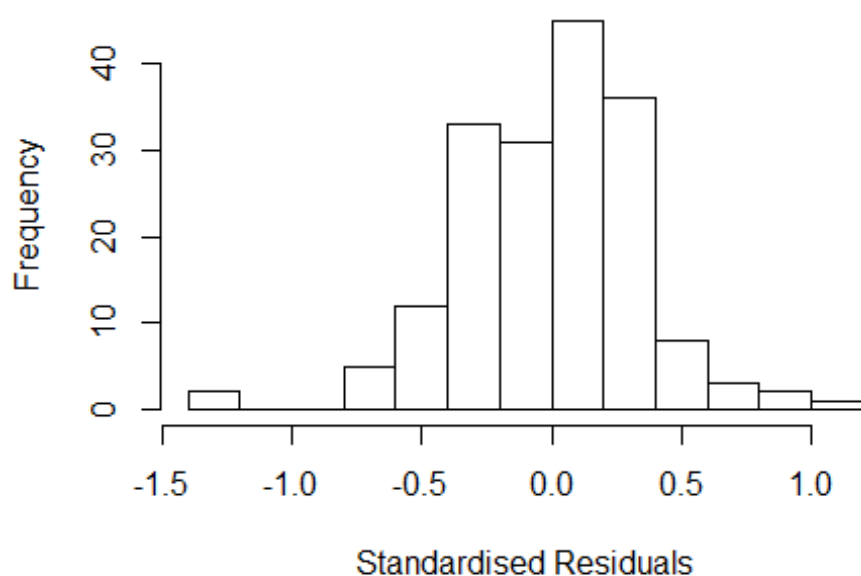
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1381 -0.2226 0.0209 0.1991 0.9248
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4420 0.1988 7.26 1.8e-11 ***
## FirstAuthorFemale1 0.0596 0.0598 1.00 0.32099
## LastAuthorFemale1 -0.1504 0.1157 -1.30 0.19556
## UniqueAuthors2 0.0561 0.1449 0.39 0.69927
## UniqueAuthors3 0.1051 0.1372 0.77 0.44489
## UniqueAuthors4 0.2626 0.1417 1.85 0.06576 .
## UniqueAuthors5 0.2601 0.1297 2.01 0.04656 *
## Year1997 -0.1440 0.2201 -0.65 0.51387
## Year1998 -0.4037 0.2119 -1.91 0.05857 .
## Year1999 -0.4358 0.1764 -2.47 0.01458 *
```

```

## Year2000          -0.3661      0.1745    -2.10   0.03753 *
## Year2001          -0.1552      0.1697    -0.91   0.36190
## Year2002          -0.2698      0.4089    -0.66   0.51034
## Year2003          -0.5282      0.1686    -3.13   0.00207 **
## Year2004          -0.5796      0.1795    -3.23   0.00152 **
## Year2005          -0.4841      0.1690    -2.86   0.00476 **
## Year2006          -0.5269      0.1609    -3.27   0.00131 **
## Year2007          -0.6291      0.1693    -3.72   0.00028 ***
## Year2008          -0.3600      0.1817    -1.98   0.04932 *
## Year2009          -0.5027      0.1528    -3.29   0.00124 **
## Year2010          -0.6973      0.1811    -3.85   0.00017 ***
## Year2011          -0.3542      0.1537    -2.30   0.02251 *
## Year2012          -0.5125      0.1936    -2.65   0.00897 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.311
## Multiple R-squared:  0.202, Adjusted R-squared:  0.089
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 167 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.153  0.891  0.949   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          5.62e-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 1.241 1          1.114
## LastAuthorFemale 1.782 1          1.335
## Year              1.844 16          1.019

```

Residuals from first and last author



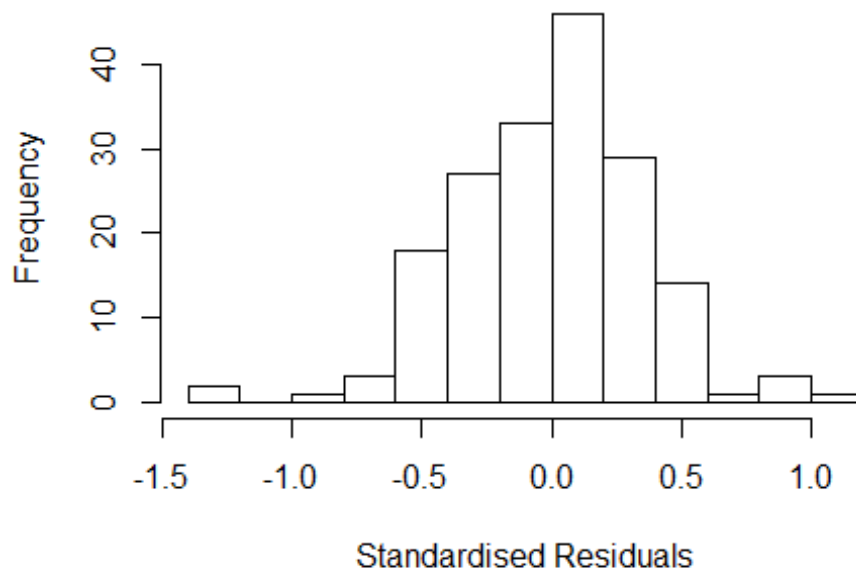
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2718 -0.2202 0.0283 0.2265 1.0423
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.7278 0.1510 11.44 < 2e-16 ***
## FirstAuthorFemale1 0.0471 0.0599 0.79 0.43227
## LastAuthorFemale1 -0.1868 0.0997 -1.87 0.06294 .
## Year1997 -0.3489 0.2123 -1.64 0.10236
## Year1998 -0.6090 0.1866 -3.26 0.00135 **
## Year1999 -0.5989 0.1782 -3.36 0.00097 ***
## Year2000 -0.5731 0.1823 -3.14 0.00199 **
## Year2001 -0.3104 0.1761 -1.76 0.07987 .
## Year2002 -0.4413 0.3818 -1.16 0.24947
## Year2003 -0.6947 0.1707 -4.07 7.4e-05 ***
## Year2004 -0.6806 0.1763 -3.86 0.00016 ***
## Year2005 -0.6091 0.1900 -3.21 0.00163 **
```

```

## Year2006          -0.6087      0.1708    -3.56  0.00048 ***
## Year2007          -0.7033      0.1820    -3.86  0.00016 ***
## Year2008          -0.4591      0.2126    -2.16  0.03233 *
## Year2009          -0.5936      0.1647    -3.60  0.00042 ***
## Year2010          -0.8346      0.1773    -4.71  5.4e-06 ***
## Year2011          -0.4560      0.1610    -2.83  0.00521 **
## Year2012          -0.6531      0.1868    -3.50  0.00061 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.331
## Multiple R-squared:  0.147, Adjusted R-squared:  0.0505
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 163 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.108  0.881  0.951  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      5.62e-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 1.205 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
## -1.2579 -0.2502 0.0313 0.2469 1.0477
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6490 0.0852 19.36 < 2e-16 ***
## FirstAuthorFemale1 0.0179 0.0589 0.30 0.76118
## Year1997 -0.3161 0.1746 -1.81 0.07205 .
## Year1998 -0.6090 0.2024 -3.01 0.00305 **
## Year1999 -0.5450 0.1231 -4.43 1.8e-05 ***
## Year2000 -0.5605 0.1463 -3.83 0.00018 ***
## Year2001 -0.2226 0.1161 -1.92 0.05705 .
## Year2002 -0.4147 0.2543 -1.63 0.10499
## Year2003 -0.6173 0.1192 -5.18 6.7e-07 ***
## Year2004 -0.6498 0.1306 -4.98 1.7e-06 ***
## Year2005 -0.5569 0.1358 -4.10 6.6e-05 ***
## Year2006 -0.5257 0.1068 -4.92 2.1e-06 ***
```

```

## Year2007          -0.6117      0.1296   -4.72  5.1e-06 ***
## Year2008          -0.4467      0.1906   -2.34  0.02033 *
## Year2009          -0.5153      0.1107   -4.65  6.8e-06 ***
## Year2010          -0.7474      0.1226   -6.10  7.8e-09 ***
## Year2011          -0.3911      0.1044   -3.74  0.00025 ***
## Year2012          -0.5797      0.1422   -4.08  7.2e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.359
## Multiple R-squared:  0.113, Adjusted R-squared:  0.0187
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 160 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.195  0.882  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      5.62e-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 1.594 1      1.263
## Year      1.594 16      1.015
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.2827 -0.2314  0.0256  0.2245  1.0233

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.7441    0.1638   10.65 < 2e-16 ***
## LastAuthorFemale1 -0.1722    0.0955   -1.80  0.07318 .
## Year1997          -0.3689    0.2197   -1.68  0.09510 .
## Year1998          -0.6090    0.1900   -3.21  0.00163 **
## Year1999          -0.6013    0.1875   -3.21  0.00162 **
## Year2000          -0.5785    0.2007   -2.88  0.00449 **
## Year2001          -0.3117    0.1815   -1.72  0.08786 .
## Year2002          -0.4464    0.3958   -1.13  0.26111
## Year2003          -0.6911    0.1819   -3.80  0.00021 ***
## Year2004          -0.6814    0.1900   -3.59  0.00044 ***
## Year2005          -0.6215    0.1965   -3.16  0.00187 **
## Year2006          -0.6130    0.1823   -3.36  0.00097 ***
## Year2007          -0.6955    0.1947   -3.57  0.00047 ***
## Year2008          -0.4792    0.2213   -2.17  0.03186 *
## Year2009          -0.5929    0.1775   -3.34  0.00104 **
## Year2010          -0.8380    0.1878   -4.46  1.5e-05 ***
## Year2011          -0.4614    0.1736   -2.66  0.00866 **
## Year2012          -0.6504    0.1982   -3.28  0.00127 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.331
## Multiple R-squared:  0.143, Adjusted R-squared:  0.0523
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 162 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0998 0.8750 0.9500 0.8980 0.9830 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      5.62e-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: 178"
## [1] ""

```

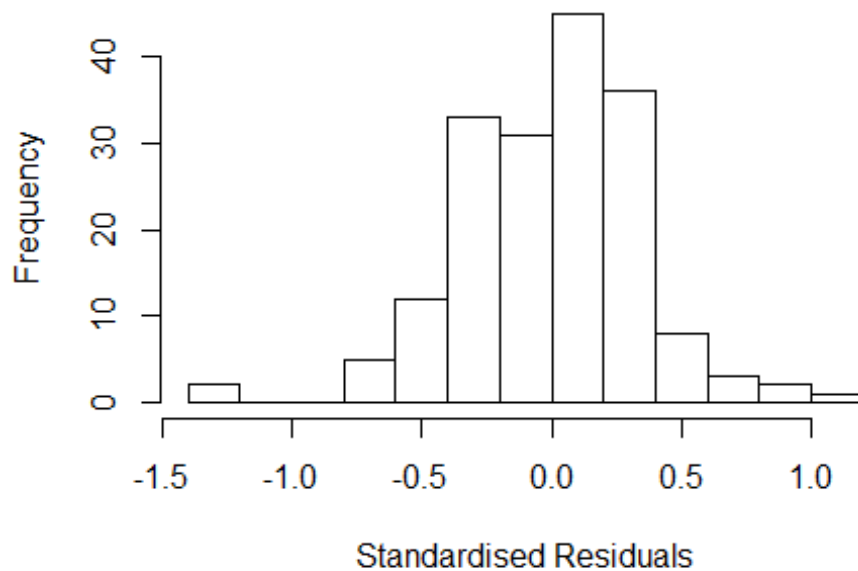
```

## [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
##    3    4    3    4    3    6    7   12   15    5   18    9   17   19   32
## 2011 2012
##   33   31
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    1    1    2    1    3    2    5    8    0   11    3    4    7   13
## 2011 2012
##   18   16
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    0    1    0    1    3    2    5    4    0    9    3    3    6    8
## 2011 2012
##   15   11
## [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: 2.49288287167843"

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

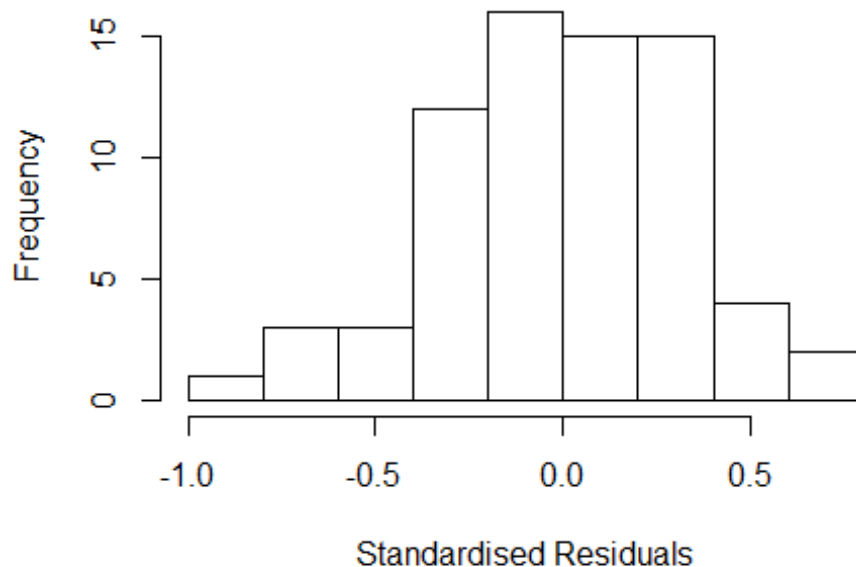
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7, p-value = 0.9
## 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: 2.61099908058295"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.904e+13  1      4.363e+06
## LastAuthorFemale  1.726e+13  1      4.155e+06
## UniqueAuthors    6.750e+02  4      2.258e+00
## Year              1.640e+16 12      4.738e+00
```

Residuals from first and last author and team size



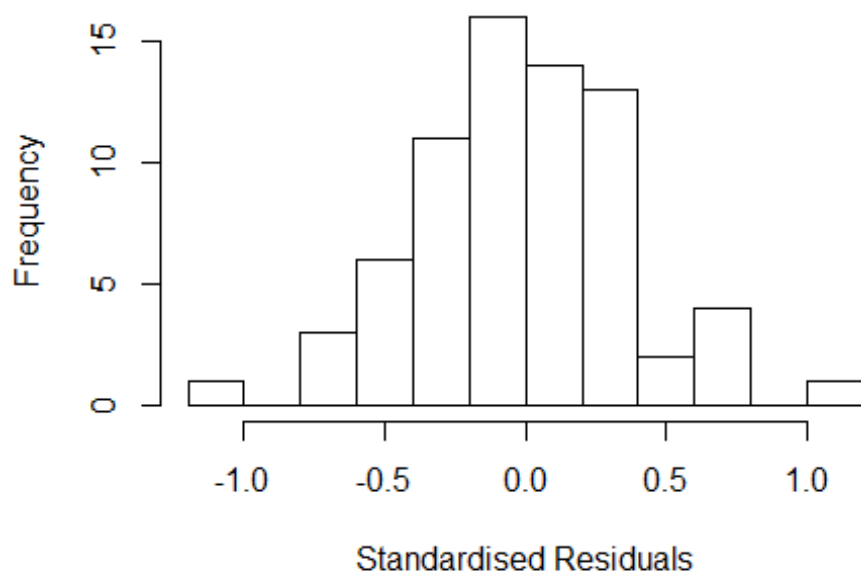
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.992 -0.210 0.014 0.217 0.789
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0281 0.1060 9.70 2.9e-13 ***
## FirstAuthorFemale1 0.0785 0.0968 0.81 0.42135
## LastAuthorFemale1 -0.1836 0.0922 -1.99 0.05174 .
## UniqueAuthors2 0.7032 0.3742 1.88 0.06585 .
## UniqueAuthors3 0.4071 0.4066 1.00 0.32129
## UniqueAuthors4 0.3195 0.4079 0.78 0.43703
## UniqueAuthors5 0.5910 0.4220 1.40 0.16731
## Year2000 -0.1061 0.1060 -1.00 0.32128
## Year2001 -0.9975 0.2725 -3.66 0.00059 ***
## Year2002 -0.3252 0.4560 -0.71 0.47895
```

```

## Year2003          -0.5258      0.4274    -1.23   0.22418
## Year2004          -0.4455      0.4284    -1.04   0.30313
## Year2006          -0.4591      0.4318    -1.06   0.29254
## Year2007          -0.2938      0.3154    -0.93   0.35590
## Year2008          -0.6964      0.4256    -1.64   0.10779
## Year2009          -0.5636      0.4157    -1.36   0.18102
## Year2010          -0.4227      0.3334    -1.27   0.21041
## Year2011          -0.3520      0.4434    -0.79   0.43083
## Year2012          -0.3914      0.4812    -0.81   0.41972
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.339
## Multiple R-squared:  0.326, Adjusted R-squared:  0.0921
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 65 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.372  0.889  0.955   0.910   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          1.41e-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 7.976e+13 1      8.931e+06
## LastAuthorFemale  8.034e+13 1      8.963e+06
## Year              3.993e+14 12      4.059e+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
## -1.11e+00 -2.36e-01 -4.16e-17 2.17e-01 1.08e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9498 0.1307 7.26 1.3e-09 ***
## FirstAuthorFemale1 0.1334 0.1100 1.21 0.231
## LastAuthorFemale1 -0.1602 0.1104 -1.45 0.153
## Year2000 -0.0278 0.1307 -0.21 0.832
## Year2001 -0.5280 0.3661 -1.44 0.155
## Year2002 0.2370 0.1151 2.06 0.044 *
## Year2003 0.0923 0.1994 0.46 0.645
## Year2004 0.1305 0.1657 0.79 0.434
## Year2006 0.1186 0.1447 0.82 0.416
## Year2007 0.0382 0.3703 0.10 0.918
## Year2008 0.0320 0.1193 0.27 0.789
## Year2009 0.0919 0.1385 0.66 0.510
```

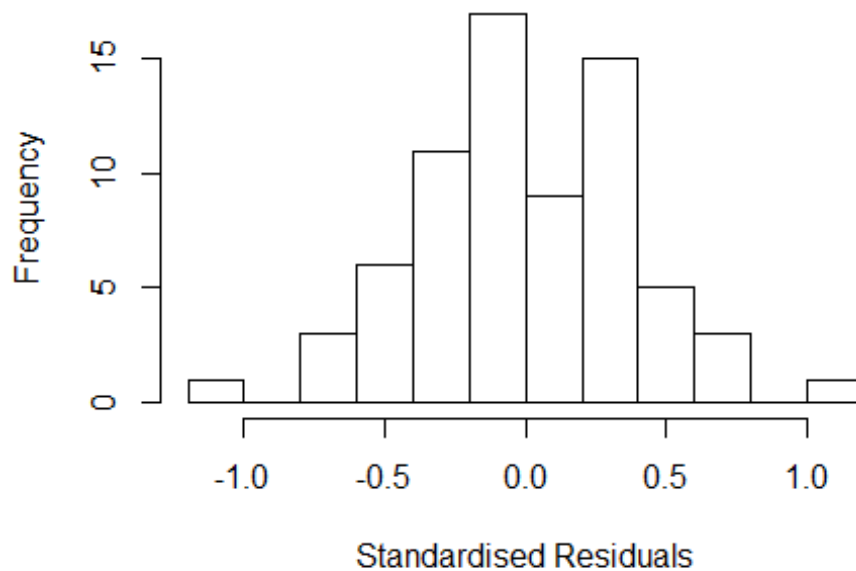


```

## Year2010          0.0730      0.2330      0.31      0.755
## Year2011          0.1576      0.1207      1.31      0.197
## Year2012          0.1633      0.1572      1.04      0.303
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.389
## Multiple R-squared:  0.147, Adjusted R-squared:  -0.0663
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 3 weights are ~= 1. The remaining 68 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.398  0.905  0.965  0.918  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 3: First author gender, Year as factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.325e+14 1      1.151e+07
## Year              1.325e+14 12      3.876e+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
## -1.09456 -0.26994 -0.00249 0.24082 1.08329
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8202 0.1083 7.57 3.5e-10 ***
## FirstAuthorFemale1 0.1028 0.1083 0.95 0.347
## Year2000 0.1018 0.1083 0.94 0.351
## Year2001 -0.3777 0.3579 -1.06 0.296
## Year2002 0.3819 0.0746 5.12 3.8e-06 ***
## Year2003 0.2095 0.2272 0.92 0.360
## Year2004 0.1943 0.1512 1.29 0.204
## Year2006 0.2160 0.1235 1.75 0.086 .
## Year2007 0.1372 0.2764 0.50 0.622
## Year2008 0.0546 0.1407 0.39 0.700
## Year2009 0.2085 0.1416 1.47 0.146
## Year2010 0.1945 0.2467 0.79 0.434
```

```

## Year2011          0.2744      0.1121      2.45      0.017 *
## Year2012          0.2636      0.1593      1.65      0.104
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.393
## Multiple R-squared:  0.121, Adjusted R-squared:  -0.079
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 61 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.417  0.900   0.952   0.909   0.979   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 4: Last author gender, Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 2.649e+14  1          1.628e+07
## Year              2.649e+14 12          3.990e+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
## -1.16e+00 -2.66e-01 -2.78e-17  2.03e-01  1.01e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0531      0.1071   9.83    7e-14 ***

```

```

## LastAuthorFemale1 -0.1301    0.1071   -1.21    0.23
## Year2000          -0.1311    0.1071   -1.22    0.23
## Year2001          -0.5468    0.4251   -1.29    0.20
## Year2002           0.2004    0.1393    1.44    0.16
## Year2003           0.0866    0.2349    0.37    0.71
## Year2004           0.0393    0.1746    0.23    0.82
## Year2006           0.0244    0.1190    0.21    0.84
## Year2007          -0.0711    0.3551   -0.20    0.84
## Year2008          -0.0913    0.0710   -1.29    0.20
## Year2009           0.0452    0.1458    0.31    0.76
## Year2010           0.0392    0.2250    0.17    0.86
## Year2011           0.1037    0.1228    0.84    0.40
## Year2012           0.0964    0.1564    0.62    0.54
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.384
## Multiple R-squared:  0.129, Adjusted R-squared:  -0.0702
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 2 weights are ~= 1. The remaining 69 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.344  0.906  0.959  0.916  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 2306"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 2001 2002 2003 2004 2005 2006 2008 2009 2010 2011 2012

```

```

##      2      2      1      1      1      2      3      2      4      4      1      3      8      5
##
## 1996 1997 1998 2001 2002 2003 2004 2005 2006 2008 2009 2010 2011 2012
##      1      2      1      0      1      1      2      1      3      4      0      2      6      2
##
## 1996 1997 1998 2001 2002 2003 2004 2005 2006 2008 2009 2010 2011 2012
##      1      2      1      0      1      1      1      1      3      3      0      2      6      2
## [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: 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 = 3, p-value = 0.4
## 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: 2.93015605158352"

## Warning in wilcox.test.default(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 = 0.07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
## [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 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

```

```

##      4      7      3      8      7      7      4      3     10      3      6      7      7     18     14
## 2011 2012
##     13      8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      4      1      3      4      3      2      1      4      2      1      3      2     13      6
## 2011 2012
##      9      5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      3      1      3      4      3      2      1      3      1      1      2      1     11      5
## 2011 2012
##      9      3
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.81094429095204"
## [1] "Male first author team size 2018 geometric mean: 2.21336383940064"

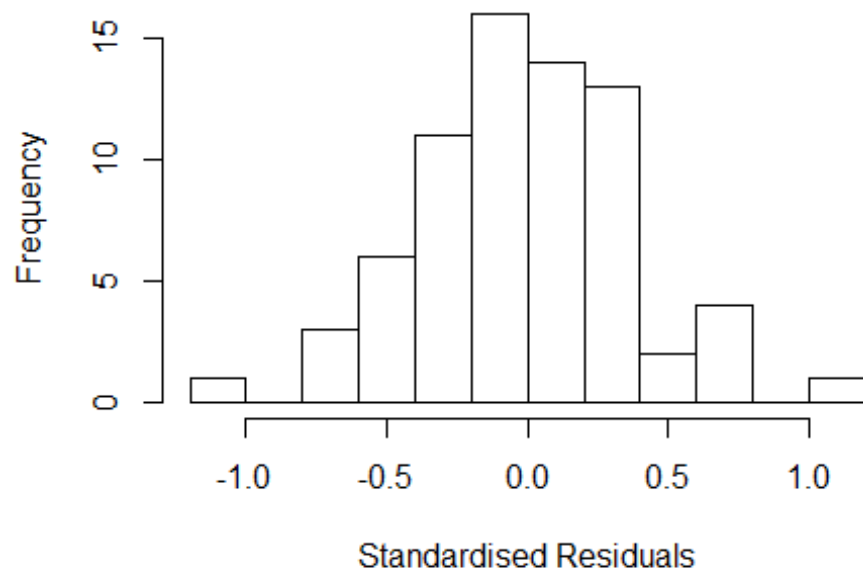
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 32, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.41995876569005"
## [1] "Male last author team size 2018 geometric mean: 3.19284598267482"

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

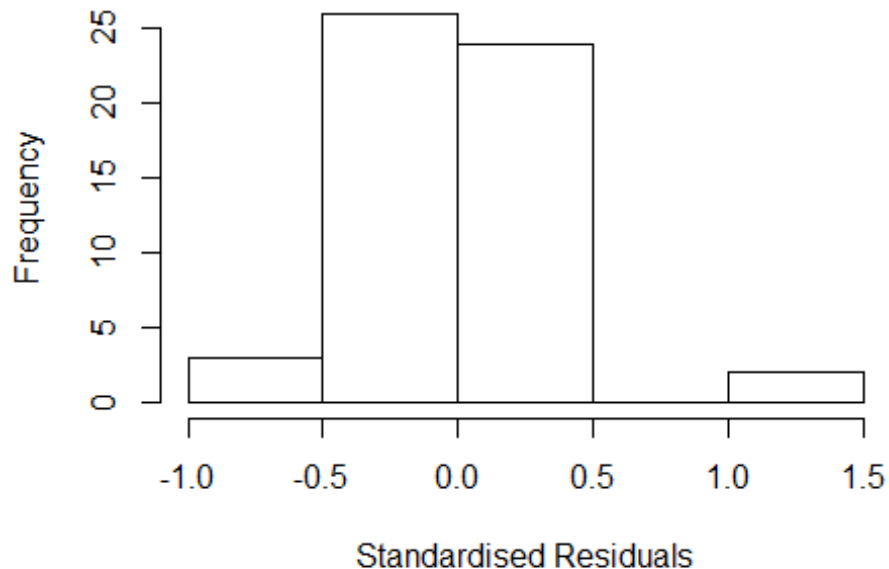
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 21, 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 1.162e+14 1      10781592
## LastAuthorFemale -1.198e+14 1           NaN
## UniqueAuthors    -2.186e+43 4           NaN
## Year              -8.300e+58 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
## -7.74e-01 -1.06e-01 1.67e-16 1.54e-01 1.22e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1989 0.4793 2.50 0.0177 *
## FirstAuthorFemale1 0.0970 0.2968 0.33 0.7460
## LastAuthorFemale1 0.2571 0.2427 1.06 0.2973
## UniqueAuthors2 -0.1898 0.3558 -0.53 0.5974
## UniqueAuthors3 -0.1045 0.4848 -0.22 0.8307
## UniqueAuthors4 -0.5513 0.7270 -0.76 0.4538
## UniqueAuthors5 -0.3460 0.3784 -0.91 0.3673
## Year1997 0.0338 0.4122 0.08 0.9352
## Year1998 0.2719 0.1510 1.80 0.0813 .
## Year1999 -0.2726 0.1836 -1.48 0.1474
```

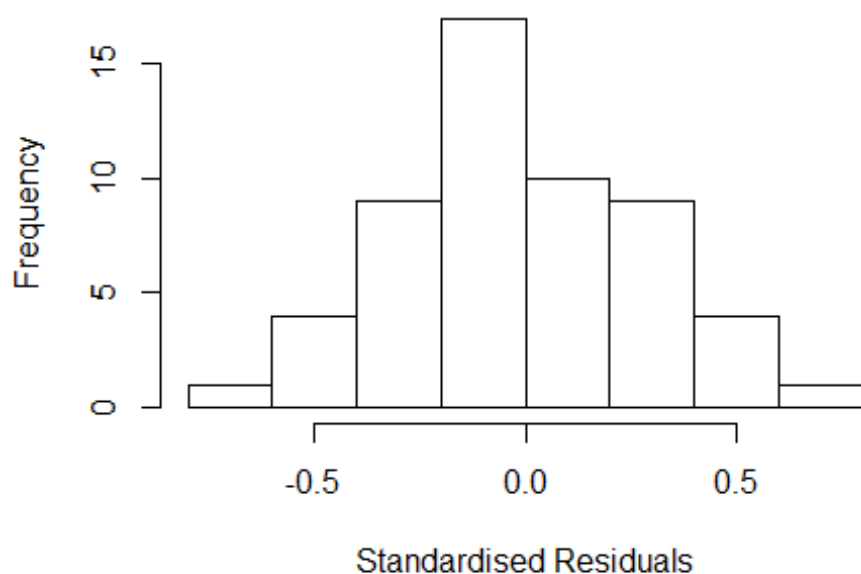


```

## Year2000          -0.2769      0.4793   -0.58   0.5675
## Year2001          -0.1522      0.2692   -0.57   0.5756
## Year2002          -0.2585      0.3230   -0.80   0.4293
## Year2003          -1.1061      0.4050   -2.73   0.0102 *
## Year2004          -0.4032      0.6451   -0.62   0.5364
## Year2005           0.4283      0.1342    3.19   0.0032 **
## Year2006           0.1319      0.1510    0.87   0.3890
## Year2007          -0.0101      0.2353   -0.04   0.9660
## Year2008           0.3559      0.1510    2.36   0.0247 *
## Year2009           0.1307      0.3462    0.38   0.7082
## Year2010           0.0727      0.2417    0.30   0.7655
## Year2011           0.2670      0.1215    2.20   0.0353 *
## Year2012          -0.2012      0.1654   -1.22   0.2326
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.279
## Multiple R-squared:  0.609, Adjusted R-squared:  0.34
## Convergence in 39 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 42 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.016  0.885   0.962   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      1.82e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -1.733e+15  1      NaN
## LastAuthorFemale  -5.252e+15  1      NaN
## Year              -3.403e+30 16      NaN

```

Residuals from first and last author



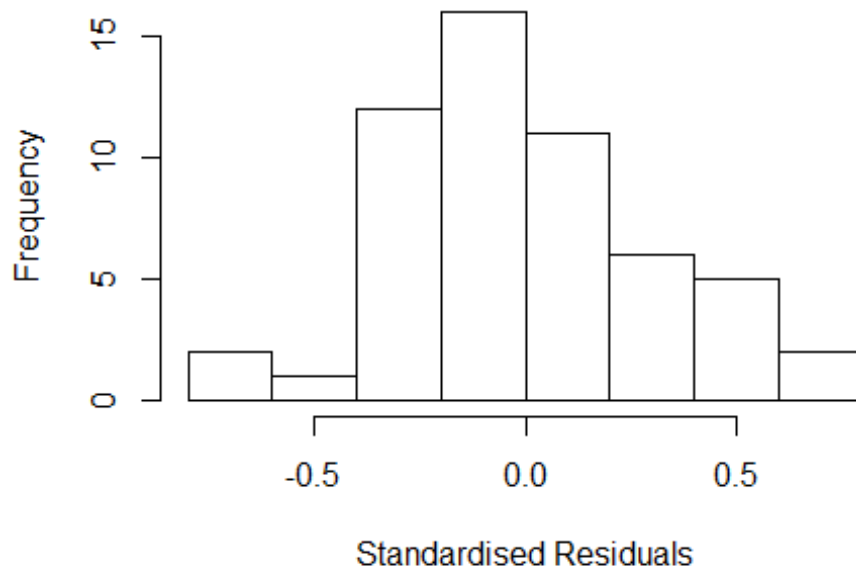
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -7.54e-01 -1.92e-01 -2.50e-16 2.06e-01 7.78e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9708 0.1263 7.69 4.2e-09 ***
## FirstAuthorFemale1 -0.1027 0.1556 -0.66 0.5135
## LastAuthorFemale1 0.1593 0.1675 0.95 0.3479
## Year1997 0.2656 0.3102 0.86 0.3975
## Year1998 0.3102 0.1263 2.46 0.0190 *
## Year1999 -0.3121 0.1806 -1.73 0.0925 .
## Year2000 -0.1336 0.2380 -0.56 0.5782
## Year2001 -0.0647 0.1835 -0.35 0.7266
## Year2002 -0.2495 0.2137 -1.17 0.2507
## Year2003 -0.8682 0.1263 -6.87 4.8e-08 ***
## Year2004 0.1651 0.6309 0.26 0.7951
## Year2005 0.2029 0.1787 1.13 0.2639
```

```

## Year2006          0.1702      0.1263      1.35      0.1863
## Year2007         -0.2335      0.1146     -2.04      0.0491 *
## Year2008          0.3942      0.1263      3.12      0.0035 **
## Year2009          0.2469      0.1592      1.55      0.1298
## Year2010          0.1198      0.1652      0.73      0.4729
## Year2011          0.2330      0.1673      1.39      0.1721
## Year2012         -0.3577      0.1732     -2.06      0.0462 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.325
## Multiple R-squared:  0.441, Adjusted R-squared:  0.162
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 50 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.545  0.919  0.960  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      1.82e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -2.825e+15  1      NaN
## Year              -2.825e+15 16      NaN

```

Residuals from first author



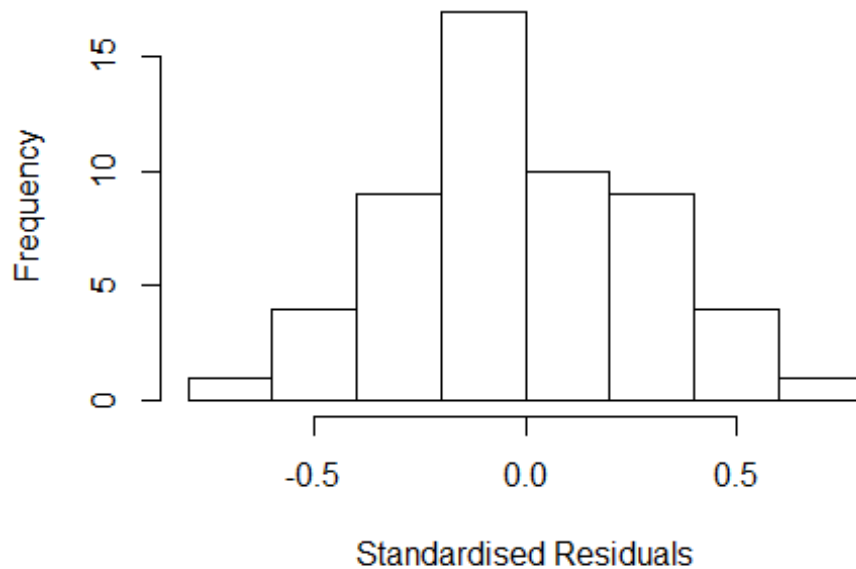
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -6.67e-01 -2.20e-01 -5.55e-17 1.93e-01 7.07e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.98155 0.11101 8.84 1.2e-10 ***
## FirstAuthorFemale1 -0.12409 0.12860 -0.96 0.3408
## Year1997 0.27205 0.28859 0.94 0.3520
## Year1998 0.29945 0.11101 2.70 0.0105 *
## Year1999 -0.26826 0.21180 -1.27 0.2132
## Year2000 -0.05109 0.17373 -0.29 0.7703
## Year2001 -0.00581 0.17002 -0.03 0.9729
## Year2002 -0.18055 0.26440 -0.68 0.4989
## Year2003 -0.85745 0.11101 -7.72 3.1e-09 ***
## Year2004 0.30424 0.39786 0.76 0.4493
## Year2005 0.35145 0.11101 3.17 0.0031 **
## Year2006 0.15945 0.11101 1.44 0.1593
```

```

## Year2007          -0.23350    0.10293   -2.27    0.0292 *
## Year2008          0.38345    0.11101    3.45    0.0014 **
## Year2009          0.30805    0.14622    2.11    0.0420 *
## Year2010          0.18293    0.16550    1.11    0.2762
## Year2011          0.25502    0.15889    1.61    0.1170
## Year2012          -0.31102    0.15367   -2.02    0.0502 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.382
## Multiple R-squared:  0.401, Adjusted R-squared:  0.126
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 46 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.713  0.931  0.955  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      1.82e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 5.305e+14 1      2.303e+07
## Year            5.305e+14 16      2.885e+00

```

Residuals from last author



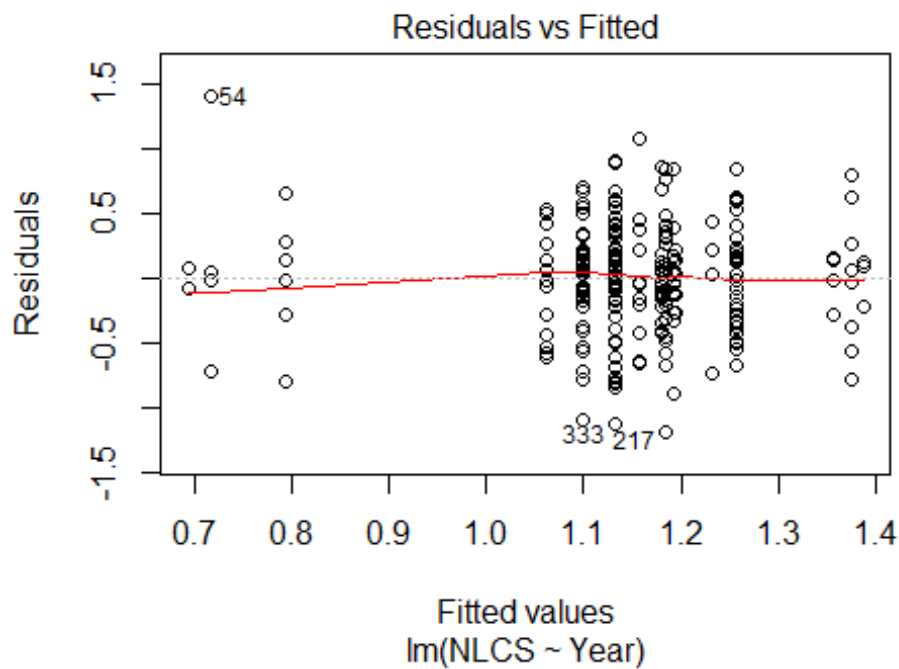
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -6.89e-01 -1.92e-01 -2.78e-17 1.86e-01 8.93e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.919 0.140 6.55 1.1e-07 ***
## LastAuthorFemale1 0.173 0.151 1.15 0.258
## Year1997 0.324 0.308 1.05 0.299
## Year1998 0.361 0.140 2.58 0.014 *
## Year1999 -0.296 0.222 -1.33 0.191
## Year2000 -0.105 0.229 -0.46 0.649
## Year2001 -0.084 0.208 -0.40 0.689
## Year2002 -0.205 0.227 -0.90 0.372
## Year2003 -0.919 0.140 -6.55 1.1e-07 ***
## Year2004 0.102 0.603 0.17 0.867
## Year2005 0.240 0.206 1.17 0.251
## Year2006 0.222 0.140 1.58 0.123
```

```

## Year2007          -0.234      0.170    -1.38     0.177
## Year2008           0.446      0.140     3.17     0.003 **
## Year2009           0.219      0.175     1.26     0.217
## Year2010           0.123      0.178     0.69     0.494
## Year2011           0.237      0.198     1.20     0.239
## Year2012          -0.311      0.194    -1.60     0.118
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.343
## Multiple R-squared:  0.425, Adjusted R-squared:  0.161
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 49 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.478  0.907   0.965   0.925   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      1.82e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 55"
## [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
##   11   4  11   9  11  10  14  13  16  23  21  27  34  59  38
## 2011 2012
##   59  61
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    3    6    4    5    5   11    5    8   15   15   18   22   43   22
## 2011 2012

```

```
## 42 32
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2 3 4 4 4 4 9 4 5 14 14 17 20 37 20
## 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 = 20, df = 16, p-value = 0.2
```

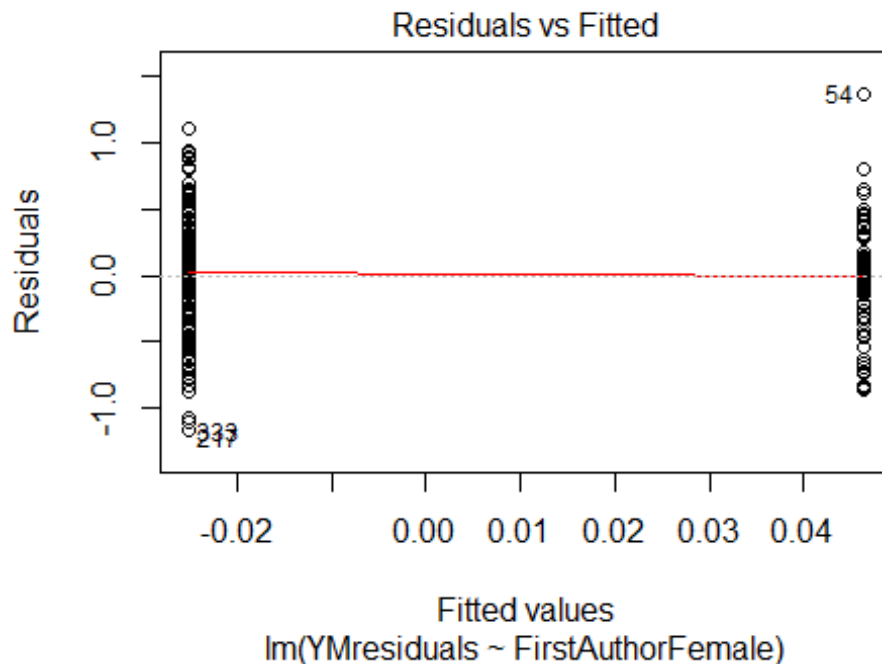


```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.1, df = 1, p-value = 0.1
## [1] "Female first author team size 2018 geometric mean: 2.28406018812534"
## [1] "Male first author team size 2018 geometric mean: 2.37086087999419"
## Warning in wilcox.test.default(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 = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.55880819475895"
## [1] "Male last author team size 2018 geometric mean: 2.20623298950442"

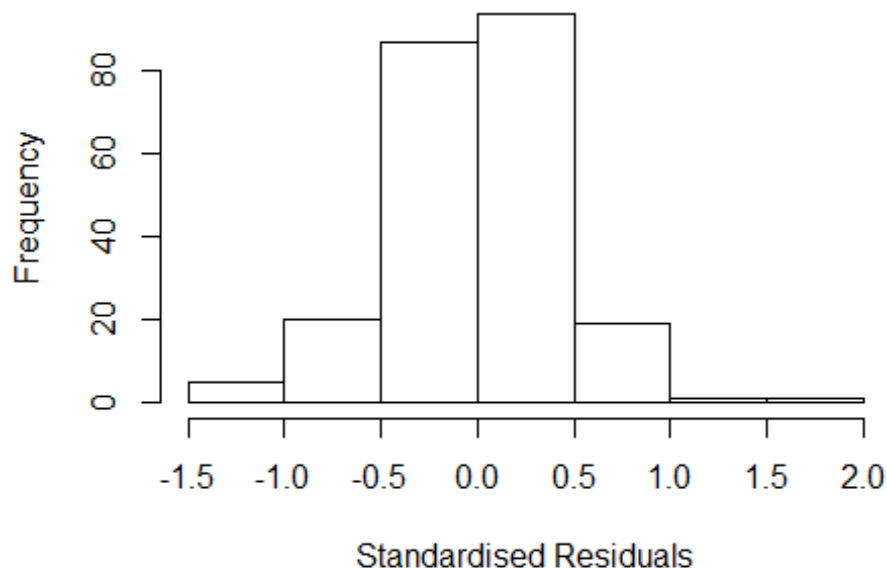
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 150, p-value = 0.5
## 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 | 2.399 | 1 | 1.549 |
| LastAuthorFemale | 3.193 | 1 | 1.787 |
| UniqueAuthors | 5.820 | 4 | 1.246 |
| Year | 7.804 | 16 | 1.066 |

Residuals from first and last author and team size



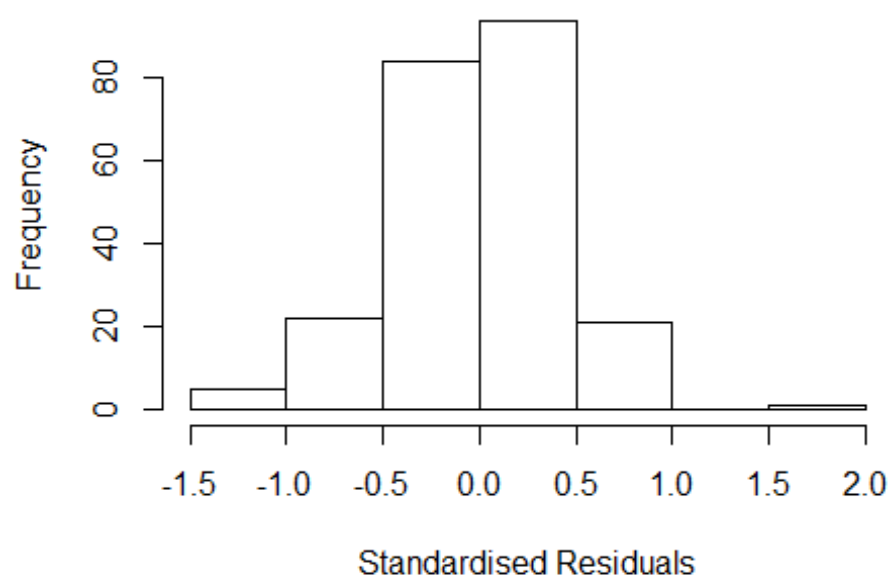
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2671 -0.2313 0.0108 0.2221 1.8512
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6557 0.0443 14.81 < 2e-16 ***
## FirstAuthorFemale1 0.0666 0.0740 0.90 0.36931
## LastAuthorFemale1 0.0091 0.0847 0.11 0.91450
## UniqueAuthors2 0.0537 0.0712 0.76 0.45098
## UniqueAuthors3 0.1681 0.1117 1.50 0.13398
## UniqueAuthors4 0.0587 0.1152 0.51 0.61065
## UniqueAuthors5 0.1051 0.1136 0.93 0.35591
## Year1997 0.6915 0.1151 6.01 8.4e-09 ***
## Year1998 0.2709 0.1730 1.57 0.11900
## Year1999 0.6735 0.0907 7.43 3.0e-12 ***
```

```

## Year2000          0.4215      0.1059      3.98  9.5e-05 ***
## Year2001         -0.4576      0.1977     -2.31  0.02166 *
## Year2002          0.2563      0.1277      2.01  0.04610 *
## Year2003          0.5079      0.1959      2.59  0.01022 *
## Year2004          0.6780      0.3256      2.08  0.03856 *
## Year2005          0.3743      0.1333      2.81  0.00547 **
## Year2006          0.4521      0.0932      4.85  2.4e-06 ***
## Year2007          0.3884      0.0669      5.81  2.4e-08 ***
## Year2008          0.5295      0.1220      4.34  2.2e-05 ***
## Year2009          0.3776      0.0967      3.91  0.00013 ***
## Year2010          0.5063      0.0999      5.07  9.0e-07 ***
## Year2011          0.4092      0.0819      4.99  1.3e-06 ***
## Year2012          0.4949      0.0998      4.96  1.5e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.358
## Multiple R-squared:  0.154, Adjusted R-squared:  0.0628
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## observation 18 is an outlier with |weight| = 0 ( < 0.00044);
## 32 weights are ~= 1. The remaining 194 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.184  0.847  0.950  0.881  0.985  0.999
## Algorithmic parameters:
##           tuning.chi          bb          tuning.psi          refine.tol
##           1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol          solve.tol          eps.outlier          eps.x
##           1.00e-07          1.00e-07          4.41e-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 2.032 1          1.425
## LastAuthorFemale  2.554 1          1.598
## Year              1.906 16          1.020

```

Residuals from first and last author



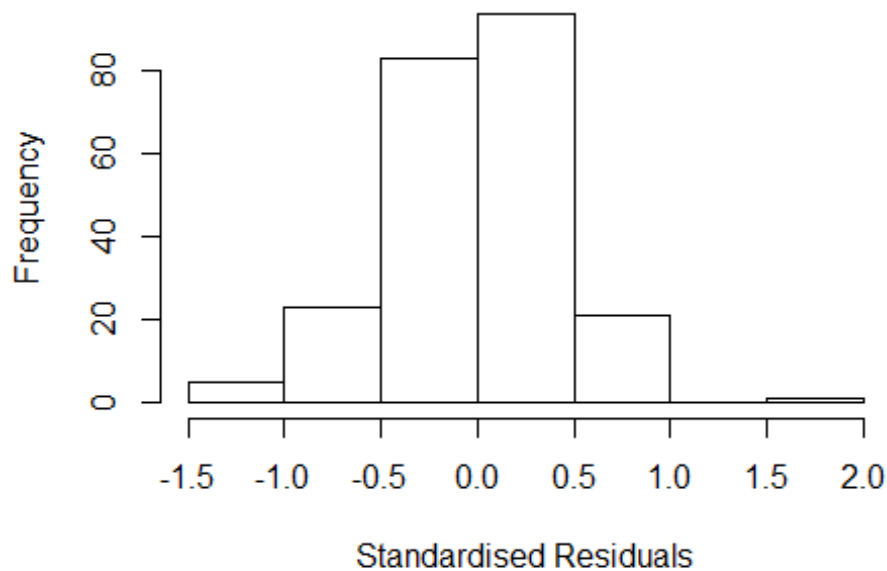
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2511 -0.2442 0.0056 0.2274 1.8639
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.66924 0.04893 13.68 < 2e-16 ***
## FirstAuthorFemale1 0.05840 0.06955 0.84 0.4021
## LastAuthorFemale1 -0.00987 0.07715 -0.13 0.8983
## Year1997 0.70476 0.11464 6.15 4.0e-09 ***
## Year1998 0.26969 0.18715 1.44 0.1511
## Year1999 0.67597 0.09714 6.96 4.4e-11 ***
## Year2000 0.44506 0.09742 4.57 8.4e-06 ***
## Year2001 -0.45671 0.21864 -2.09 0.0379 *
## Year2002 0.29503 0.12368 2.39 0.0180 *
## Year2003 0.49090 0.19691 2.49 0.0134 *
## Year2004 0.65772 0.30202 2.18 0.0305 *
## Year2005 0.40463 0.12799 3.16 0.0018 **
```

```

## Year2006          0.48840      0.09861      4.95  1.5e-06 ***
## Year2007          0.42731      0.06795      6.29  1.9e-09 ***
## Year2008          0.58183      0.11244      5.17  5.4e-07 ***
## Year2009          0.43265      0.08458      5.12  7.1e-07 ***
## Year2010          0.52010      0.10272      5.06  9.1e-07 ***
## Year2011          0.47129      0.07166      6.58  3.8e-10 ***
## Year2012          0.56680      0.09764      5.81  2.4e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.37
## Multiple R-squared:  0.131, Adjusted R-squared:  0.0561
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## observation 18 is an outlier with |weight| = 0 ( < 0.00044);
## 27 weights are ~= 1. The remaining 199 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.230  0.847   0.955   0.889   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      4.41e-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 1.23 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
## -1.24801 -0.24463 0.00271 0.22907 1.85853
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6667 0.0446 14.95 < 2e-16 ***
## FirstAuthorFemale1 0.0536 0.0547 0.98 0.3279
## Year1997 0.7057 0.1152 6.13 4.4e-09 ***
## Year1998 0.2732 0.1876 1.46 0.1469
## Year1999 0.6797 0.0924 7.36 4.2e-12 ***
## Year2000 0.4477 0.0947 4.73 4.2e-06 ***
## Year2001 -0.4538 0.2170 -2.09 0.0377 *
## Year2002 0.2963 0.1237 2.40 0.0175 *
## Year2003 0.4928 0.1963 2.51 0.0128 *
## Year2004 0.6598 0.3007 2.19 0.0294 *
## Year2005 0.4051 0.1273 3.18 0.0017 **
## Year2006 0.4915 0.0923 5.32 2.6e-07 ***
```

```

## Year2007          0.4285      0.0661      6.48  6.5e-10 ***
## Year2008          0.5813      0.1122      5.18  5.1e-07 ***
## Year2009          0.4340      0.0839      5.17  5.4e-07 ***
## Year2010          0.5203      0.1031      5.05  9.7e-07 ***
## Year2011          0.4720      0.0713      6.62  3.0e-10 ***
## Year2012          0.5695      0.0920      6.19  3.1e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.373
## Multiple R-squared:  0.13,   Adjusted R-squared:  0.0595
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## observation 18 is an outlier with |weight| = 0 ( < 0.00044);
## 27 weights are ~= 1. The remaining 199 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.240  0.847   0.956   0.890   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      4.41e-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 1.503  1          1.226
## Year              1.503 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

```

```

## -1.25446 -0.23715 0.00909 0.21952 1.88385
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.6798    0.0538   12.64 < 2e-16 ***
## LastAuthorFemale1 0.0274    0.0601    0.46  0.6486
## Year1997        0.7007    0.1112    6.30  1.7e-09 ***
## Year1998        0.2702    0.2012    1.34  0.1808
## Year1999        0.6803    0.1047    6.50  5.9e-10 ***
## Year2000        0.4547    0.0993    4.58  8.0e-06 ***
## Year2001       -0.4661    0.2186   -2.13  0.0342 *
## Year2002        0.2977    0.1298    2.29  0.0228 *
## Year2003        0.4780    0.1981    2.41  0.0167 *
## Year2004        0.6467    0.2834    2.28  0.0235 *
## Year2005        0.4068    0.1309    3.11  0.0021 **
## Year2006        0.4972    0.1005    4.95  1.5e-06 ***
## Year2007        0.4267    0.0746    5.72  3.7e-08 ***
## Year2008        0.5747    0.1155    4.97  1.4e-06 ***
## Year2009        0.4366    0.0874    4.99  1.2e-06 ***
## Year2010        0.5172    0.1079    4.79  3.1e-06 ***
## Year2011        0.4750    0.0754    6.30  1.7e-09 ***
## Year2012        0.5693    0.0990    5.75  3.1e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.38
## Multiple R-squared:  0.126, Adjusted R-squared:  0.0544
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## observation 18 is an outlier with |weight| = 0 ( < 0.00044);
## 26 weights are ~ = 1. The remaining 200 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.253  0.853  0.954  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           4.41e-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)

```



```

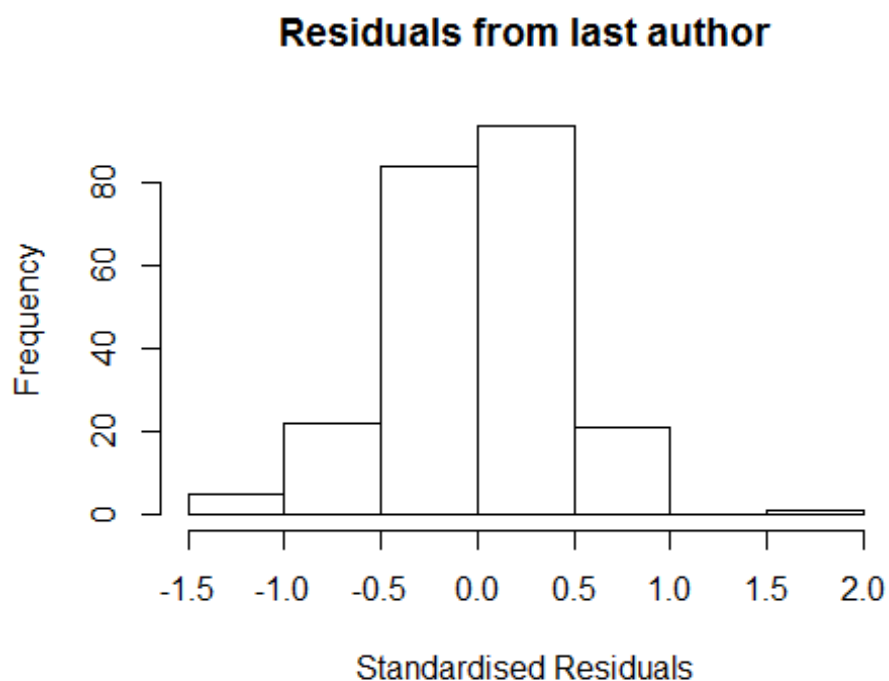
## [1] "Sample size for the above analysis: 227"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2309"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    2    4    5    2    4    9    7    2    5    7    7   13   12   22
## 2011 2012
##   13   16
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    2    2    1    2    6    2    1    3    7    5   12   10   15
## 2011 2012
##   10    9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    2    2    0    2    6    1    0    3    7    5   11   10   14
## 2011 2012
##   10    9
## [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: 2.56979658685065"

## Warning in wilcox.test.default(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 = 0.4
## 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: 2.28942848510666"

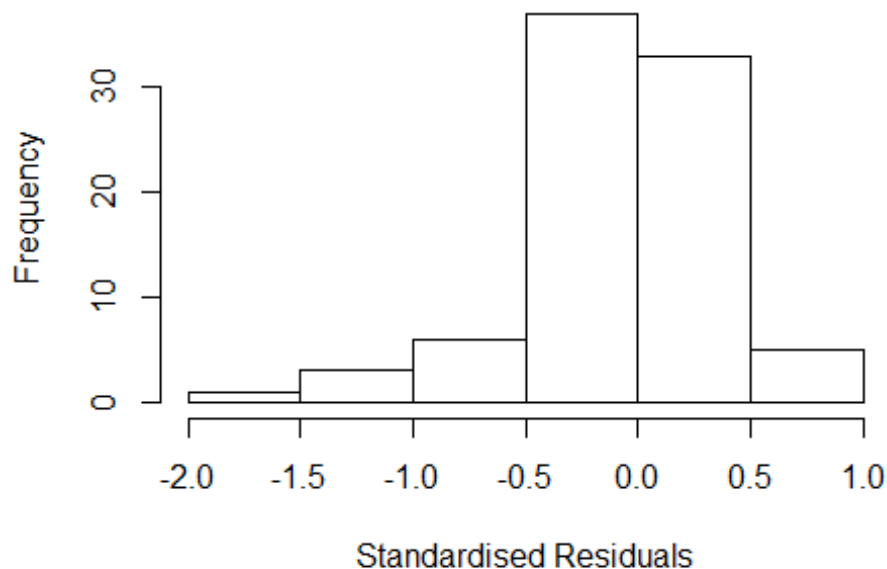
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7.5, 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 4.184e+00  1          2.045
## LastAuthorFemale  7.345e+00  1          2.710
## UniqueAuthors    1.684e+17  4         142.333
## Year              1.020e+18 14          4.397
```

Residuals from first and last author and team size



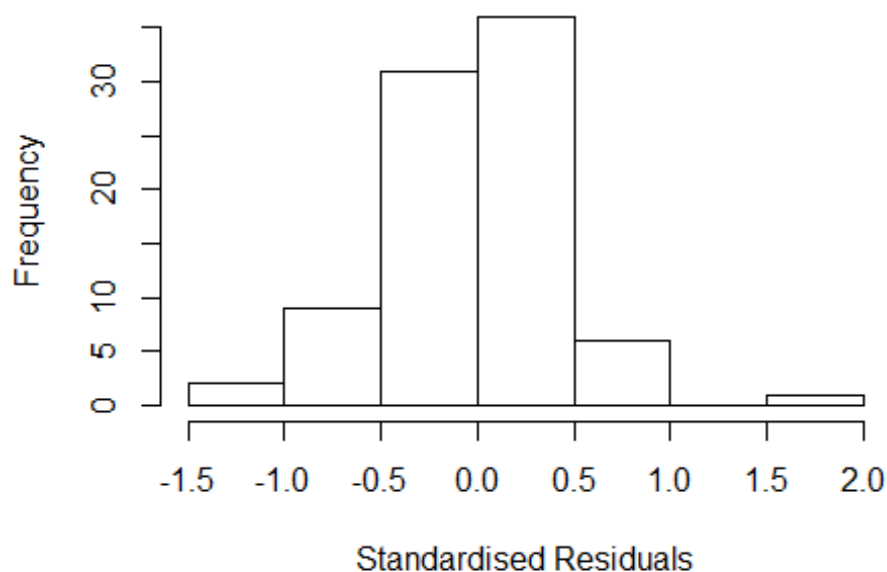
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.6519 -0.1870 -0.0148 0.1444 0.8115
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.90194 0.16403 11.60 < 2e-16 ***
## FirstAuthorFemale1 0.01807 0.11788 0.15 0.87868
## LastAuthorFemale1 0.11434 0.14517 0.79 0.43384
## UniqueAuthors2 -0.14391 0.23347 -0.62 0.53983
## UniqueAuthors3 -0.07381 0.23600 -0.31 0.75550
## UniqueAuthors4 -0.18817 0.28759 -0.65 0.51526
## UniqueAuthors5 -0.00471 0.20979 -0.02 0.98215
## Year1997 -0.97210 0.27111 -3.59 0.00065 ***
## Year1998 -1.08207 0.24807 -4.36 4.8e-05 ***
## Year1999 -0.87643 0.99271 -0.88 0.38061
```

```

## Year2001      -0.96410      0.27200      -3.54      0.00074 ***
## Year2002      -0.96220      0.22821      -4.22      8.0e-05 ***
## Year2003      -0.61220      0.26876      -2.28      0.02608 *
## Year2005      -0.67714      0.26968      -2.51      0.01458 *
## Year2006      -0.79786      0.26000      -3.07      0.00315 **
## Year2007      -0.88948      0.25512      -3.49      0.00089 ***
## Year2008      -0.72259      0.32582      -2.22      0.03013 *
## Year2009      -0.61308      0.26005      -2.36      0.02147 *
## Year2010      -0.71946      0.26140      -2.75      0.00769 **
## Year2011      -0.87831      0.35137      -2.50      0.01501 *
## Year2012      -0.40132      0.22502      -1.78      0.07925 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.373
## Multiple R-squared:  0.283, Adjusted R-squared:  0.0594
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 70 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0115 0.8720 0.9600 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      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 7.964 1 2.822
## LastAuthorFemale 2.937 1 1.714
## Year 21.320 14 1.115

```

Residuals from first and last author



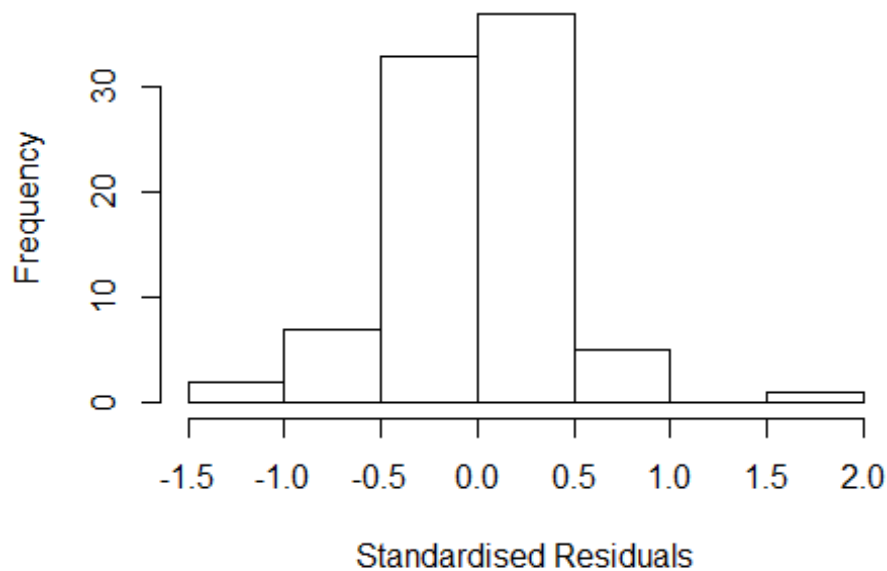
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.06723 -0.19048 0.00539 0.17442 1.65789
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.2631 0.1102 2.39 0.01970 *
## FirstAuthorFemale1 0.0287 0.1138 0.25 0.80190
## LastAuthorFemale1 0.1493 0.0925 1.61 0.11116
## Year1997 0.5122 0.1745 2.93 0.00455 **
## Year1998 0.4076 0.1407 2.90 0.00507 **
## Year1999 0.5911 1.1069 0.53 0.59510
## Year2001 0.5202 0.1759 2.96 0.00427 **
## Year2002 0.5897 0.1639 3.60 0.00061 ***
## Year2003 0.9422 0.1745 5.40 9.2e-07 ***
## Year2005 0.8864 0.2300 3.85 0.00026 ***
## Year2006 0.7568 0.2266 3.34 0.00137 **
## Year2007 0.6669 0.1635 4.08 0.00012 ***
```

```

## Year2008          0.8041      0.2233      3.60  0.00060 ***
## Year2009          0.9205      0.1782      5.17  2.3e-06 ***
## Year2010          0.7840      0.1917      4.09  0.00012 ***
## Year2011          0.7034      0.2621      2.68  0.00915 **
## Year2012          1.1723      0.1424      8.23  8.2e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.371
## Multiple R-squared:  0.265, Adjusted R-squared:  0.0922
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 72 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0079 0.8440 0.9620 0.8870 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      1.18e-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 9.341 1      3.056
## Year              9.341 14      1.083

```

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
## -1.12489 -0.18955 0.00497 0.22670 1.62306
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.2979 0.3027 0.98 0.32843
## FirstAuthorFemale1 0.0656 0.1219 0.54 0.59223
## Year1997 0.4404 0.3646 1.21 0.23116
## Year1998 0.3543 0.3307 1.07 0.28783
## Year1999 0.5378 1.0034 0.54 0.59372
## Year2001 0.4484 0.3653 1.23 0.22371
## Year2002 0.5788 0.3093 1.87 0.06554 .
## Year2003 0.8704 0.3646 2.39 0.01971 *
## Year2005 0.8734 0.3813 2.29 0.02505 *
## Year2006 0.7218 0.3885 1.86 0.06748 .
## Year2007 0.6024 0.3537 1.70 0.09303 .
## Year2008 0.8269 0.3950 2.09 0.03999 *
```

```

## Year2009          0.9020      0.3477      2.59  0.01156 *
## Year2010          0.7647      0.3666      2.09  0.04067 *
## Year2011          0.6648      0.3961      1.68  0.09777 .
## Year2012          1.1631      0.3180      3.66  0.00049 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.381
## Multiple R-squared:  0.245, Adjusted R-squared:  0.0814
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 73 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0295 0.8860 0.9560 0.8920 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      1.18e-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"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 3.124 1          1.768
## Year              3.124 14          1.042

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q    Median      3Q      Max
## -1.086730 -0.193627  0.000168  0.181154  1.668720
##
## Coefficients:

```



```

##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.2523    0.0416   6.06 6.4e-08 ***
## LastAuthorFemale1 0.1552    0.0955   1.63 0.1087
## Year1997          0.5517    0.0416  13.25 < 2e-16 ***
## Year1998          0.4327    0.0666   6.50 1.1e-08 ***
## Year1999          0.6162    1.2825   0.48 0.6324
## Year2001          0.5597    0.0471  11.89 < 2e-16 ***
## Year2002          0.6040    0.1392   4.34 4.8e-05 ***
## Year2003          0.9817    0.0416  23.59 < 2e-16 ***
## Year2005          0.9230    0.1908   4.84 7.7e-06 ***
## Year2006          0.7856    0.1470   5.34 1.1e-06 ***
## Year2007          0.7007    0.0555  12.63 < 2e-16 ***
## Year2008          0.8344    0.1506   5.54 5.1e-07 ***
## Year2009          0.9446    0.1254   7.53 1.4e-10 ***
## Year2010          0.8128    0.1297   6.27 2.8e-08 ***
## Year2011          0.7173    0.2342   3.06 0.0031 **
## Year2012          1.1856    0.0982  12.07 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.363
## Multiple R-squared:  0.269, Adjusted R-squared:  0.11
## Convergence in 30 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 71 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0014 0.8420 0.9510 0.8820 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      1.18e-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 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
##    11    3    8   12    9    9    6    9   17   11   10    7   16   34   30
## 2011 2012
##    38   34
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     2    2    1    1    3    3    3    4   10    3    5    4    7   15   12
## 2011 2012
##    18   16
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     1    1    1    0    2    3    2    3    7    1    5    4    5   13   10
## 2011 2012
##    14   12
## [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: 3.91486764116886"

## Warning in wilcox.test.default(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] "Female last author team size 2018 geometric mean: 3.91486764116886"
## [1] "Male last author team size 2018 geometric mean: 3.91486764116886"

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

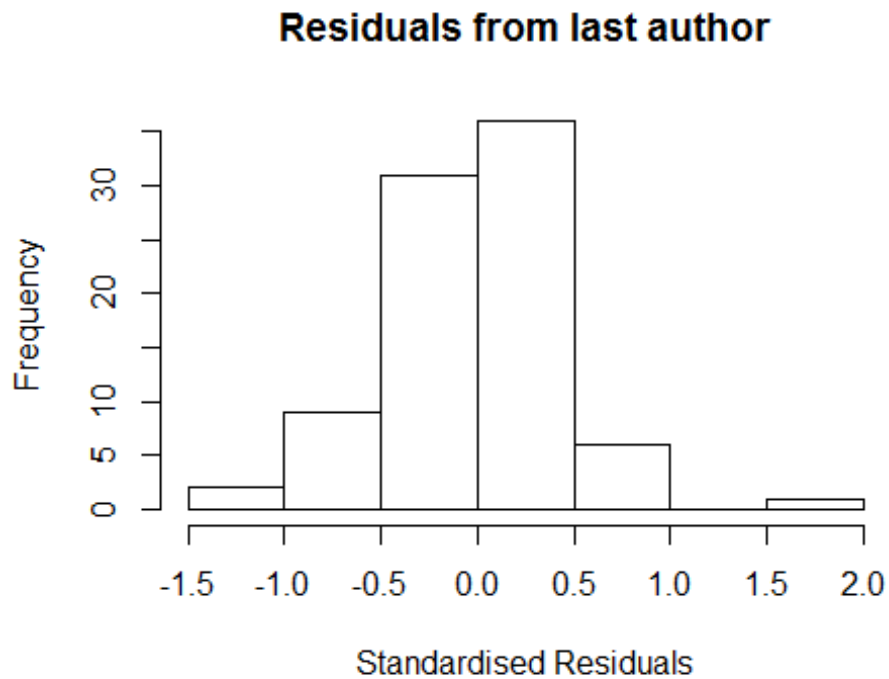
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 9, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"

## Warning in lmrob.fit(x, y, control, init = init, mf = mf): M-step did NOT
## converge. Returning unconverged SM-estimate

```

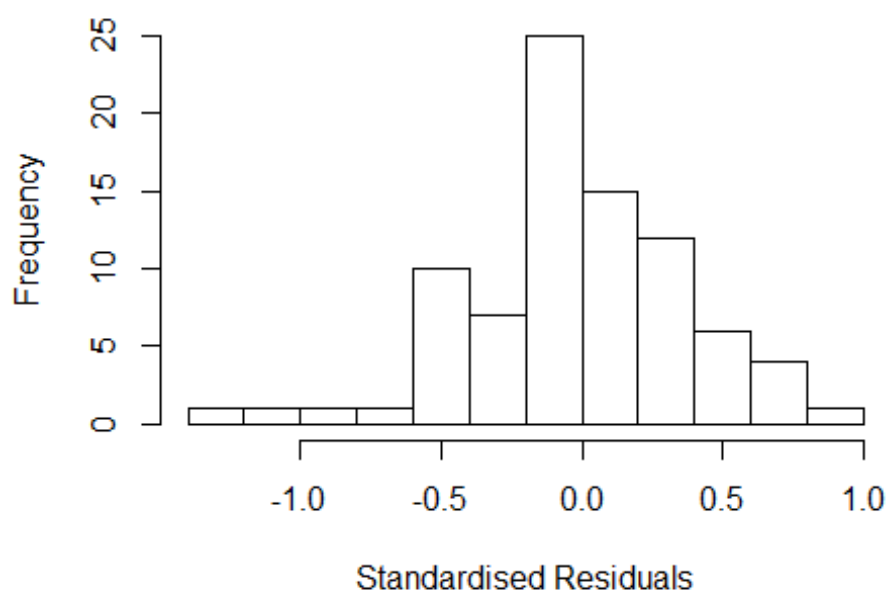
```
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"

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



```
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -3.194e+15  1      NaN
## LastAuthorFemale  1.659e+00  1      1.288
## Year              -5.298e+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
## -1.211564 -0.201552 -0.000453 0.232660 0.986940
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.66e+00 0.00e+00 Inf < 2e-16 ***
## FirstAuthorFemale1 9.00e-02 1.03e-01 8.80e-01 0.3833
## LastAuthorFemale1 -2.39e-01 1.28e-01 -1.86e+00 0.0669 .
## Year1997 -6.00e-01 4.66e-08 -1.29e+07 < 2e-16 ***
## Year1998 -1.16e+00 1.03e-01 -1.13e+01 < 2e-16 ***
## Year2000 -7.47e-01 5.14e-02 -1.45e+01 < 2e-16 ***
## Year2001 -4.58e-01 8.31e-02 -5.51e+00 6.5e-07 ***
## Year2002 -7.68e-01 3.50e-01 -2.20e+00 0.0315 *
## Year2003 -7.76e-01 2.45e-01 -3.17e+00 0.0023 **
## Year2004 -4.29e-01 1.65e-01 -2.59e+00 0.0117 *
## Year2005 -3.23e-01 2.11e-08 -1.53e+07 < 2e-16 ***
## Year2006 -5.36e-01 1.71e-01 -3.14e+00 0.0025 **
```

```

## Year2007          -3.32e-01    2.65e-01 -1.25e+00    0.2160
## Year2008          -2.97e-01    1.08e-01 -2.75e+00    0.0078 **
## Year2009          -4.51e-01    1.60e-01 -2.83e+00    0.0062 **
## Year2010          -5.41e-01    3.11e-01 -1.74e+00    0.0871 .
## Year2011          -5.41e-01    7.98e-02 -6.78e+00    3.9e-09 ***
## Year2012          -5.15e-01    1.07e-01 -4.81e+00    9.3e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.384
## Multiple R-squared:  0.139, Adjusted R-squared:  -0.0822
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 72 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.299  0.875   0.958   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      1.19e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -1.354e+15  1      NaN
## Year              -1.354e+15 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
## -1.15708 -0.21544 -0.00143  0.23358  0.98845

```

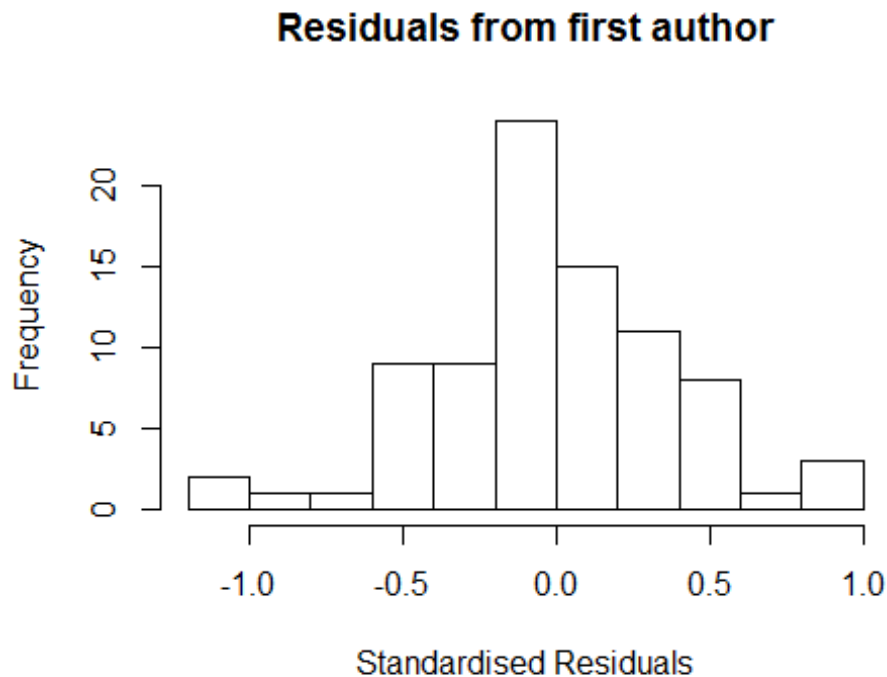
```

##
## Coefficients:
##               Estimate Std. Error   t value Pr(>|t|)
## (Intercept)    1.66e+00   1.84e-08  9.04e+07 < 2e-16 ***
## FirstAuthorFemale1 8.37e-02   1.00e-01  8.30e-01  0.40774
## Year1997       -6.00e-01   2.19e-08 -2.74e+07 < 2e-16 ***
## Year1998       -1.16e+00   1.00e-01 -1.15e+01 < 2e-16 ***
## Year2000       -7.44e-01   5.03e-02 -1.48e+01 < 2e-16 ***
## Year2001       -4.55e-01   8.23e-02 -5.53e+00 5.7e-07 ***
## Year2002       -7.65e-01   3.50e-01 -2.19e+00 0.03212 *
## Year2003       -7.74e-01   2.46e-01 -3.15e+00 0.00242 **
## Year2004       -5.04e-01   1.74e-01 -2.89e+00 0.00516 **
## Year2005       -3.23e-01   3.13e-08 -1.03e+07 < 2e-16 ***
## Year2006       -5.93e-01   1.76e-01 -3.36e+00 0.00129 **
## Year2007       -3.35e-01   2.64e-01 -1.27e+00 0.20942
## Year2008       -3.81e-01   1.53e-01 -2.49e+00 0.01520 *
## Year2009       -5.06e-01   1.47e-01 -3.45e+00 0.00098 ***
## Year2010       -5.42e-01   3.00e-01 -1.81e+00 0.07491 .
## Year2011       -5.40e-01   7.94e-02 -6.80e+00 3.4e-09 ***
## Year2012       -5.58e-01   1.08e-01 -5.18e+00 2.3e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.393
## Multiple R-squared:  0.107, Adjusted R-squared:  -0.107
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 74 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.367  0.875   0.953   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        1.19e-03        1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01        5.00e-01
##   nResample    max.it    best.r.s    k.fast.s    k.max maxit.scale
##      500         50         2         1        1000         200
##   trace.lev    mts    compute.rd
##      0         1000         0
##      psi          subsampling          cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"

```

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

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



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

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.26e+00 -2.15e-01 -7.22e-16  2.19e-01  9.72e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.6630      0.0000      Inf < 2e-16 ***
## LastAuthorFemale1 -0.2332      0.1331     -1.75  0.0844 .
##
```

```

## Year1997      -0.6000      0.0000      -Inf < 2e-16 ***
## Year1998      -1.0740      0.0000      -Inf < 2e-16 ***
## Year2000      -0.7020      0.0276    -25.41 < 2e-16 ***
## Year2001      -0.4268      0.0709     -6.02 8.2e-08 ***
## Year2002      -0.7235      0.4016     -1.80 0.0761 .
## Year2003      -0.7534      0.2794     -2.70 0.0088 **
## Year2004      -0.4133      0.1640     -2.52 0.0141 *
## Year2005      -0.3230      0.0000      -Inf < 2e-16 ***
## Year2006      -0.5164      0.1693     -3.05 0.0033 **
## Year2007      -0.3150      0.2713     -1.16 0.2497
## Year2008      -0.2811      0.1144     -2.46 0.0166 *
## Year2009      -0.4075      0.1401     -2.91 0.0049 **
## Year2010      -0.5261      0.2715     -1.94 0.0569 .
## Year2011      -0.5139      0.0708     -7.26 5.2e-10 ***
## Year2012      -0.4932      0.1016     -4.85 7.6e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.401
## Multiple R-squared:  0.125, Adjusted R-squared:  -0.0837
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 71 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.306  0.881  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           1.19e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01           5.00e-01
##   nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500           50           2           1           1000           200
##   trace.lev      mts      compute.rd
##           0           1000           0
##           psi           subsampling           cov
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 84"
## [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]"

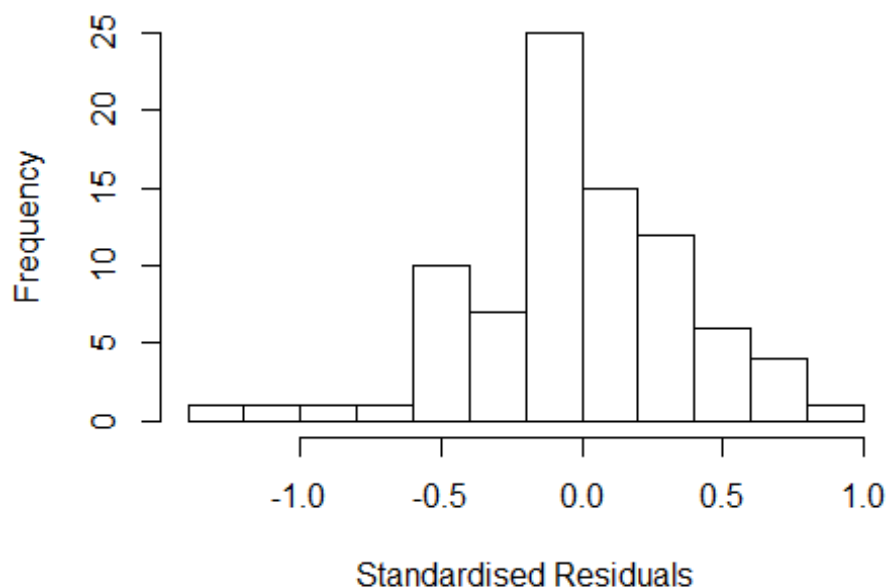
```



```
##
## 1996 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    5    4   11    5    6    7    8   12    4   13   12   16   27   25   28
## 2012
##    29
##
## 1996 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    2    1    5    4    1    3    3    7    2    7    6    8   14    8   18
## 2012
##    12
##
## 1996 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    1    2    4    1    2    3    3    2    7    6    7   11    7   14
## 2012
##     7
## [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: 3.91486764116886"

## 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, 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: 3.51948202893552"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -5.604e+14  1      NaN
## LastAuthorFemale  1.706e+01  1      4.13
## UniqueAuthors    -5.765e+31  4      NaN
## Year              -1.166e+33 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
## -1.02e+00 -1.93e-01  6.25e-16  1.22e-01  7.36e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.522949    0.286700     5.31  1.9e-06 ***
## FirstAuthorFemale1  0.023971    0.105580     0.23  0.82122
## LastAuthorFemale1  0.000955    0.153177     0.01  0.99505
## UniqueAuthors2    0.140051    0.286700     0.49  0.62711
## UniqueAuthors3    0.125287    0.308985     0.41  0.68667
## UniqueAuthors4    0.024920    0.286510     0.09  0.93100
## UniqueAuthors5    0.255251    0.320506     0.80  0.42916
## Year1998          -1.083208    0.182990    -5.92  2.1e-07 ***
## Year1999          -1.079449    0.289035    -3.73  0.00044 ***
## Year2000          -0.688003    0.093326    -7.37  8.4e-10 ***
## Year2001          -0.170869    0.116513    -1.47  0.14810
## Year2002          -0.372039    0.191581    -1.94  0.05718 .
## Year2003          -0.453032    0.334610    -1.35  0.18120
## Year2004          -0.538258    0.195554    -2.75  0.00796 **
## Year2005          -0.290949    0.295547    -0.98  0.32913
## Year2006          -0.425133    0.292998    -1.45  0.15237
## Year2007          -0.279001    0.208089    -1.34  0.18541
## Year2008          -0.702536    0.170728    -4.11  0.00013 ***
## Year2009          -0.417821    0.246943    -1.69  0.09621 .
## Year2010          -0.850610    0.172451    -4.93  7.6e-06 ***

```

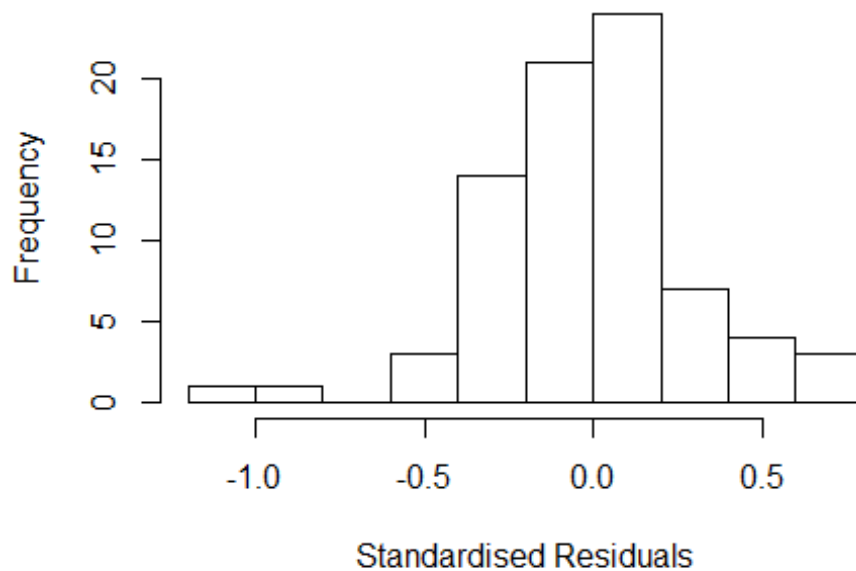
```

## Year2011          -0.385881    0.146531    -2.63    0.01091 *
## Year2012          -0.684573    0.173484    -3.95    0.00022 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.27
## Multiple R-squared:  0.442, Adjusted R-squared:  0.233
## Convergence in 35 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 69 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.122  0.859   0.957   0.889   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      1.28e-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"

## 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 9.386e+13  1      9.688e+06
## LastAuthorFemale  4.137e+00  1      2.034e+00
## Year              3.316e+14 15      3.048e+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
## -1.02890 -0.13738  0.00191  0.10586  0.80249
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.66e+00   0.00e+00      Inf < 2e-16 ***
## FirstAuthorFemale1  2.36e-04   1.29e-01   0.00e+00  0.99855
## LastAuthorFemale1  1.30e-02   1.23e-01   1.10e-01  0.91618
## Year1998        -1.07e+00   1.29e-01  -8.32e+00  1.4e-11 ***
## Year1999        -1.22e+00   3.67e-02 -3.32e+01 < 2e-16 ***
## Year2000        -7.46e-01   4.03e-02 -1.85e+01 < 2e-16 ***
## Year2001        -2.86e-01   1.18e-08 -2.42e+07 < 2e-16 ***
```

```

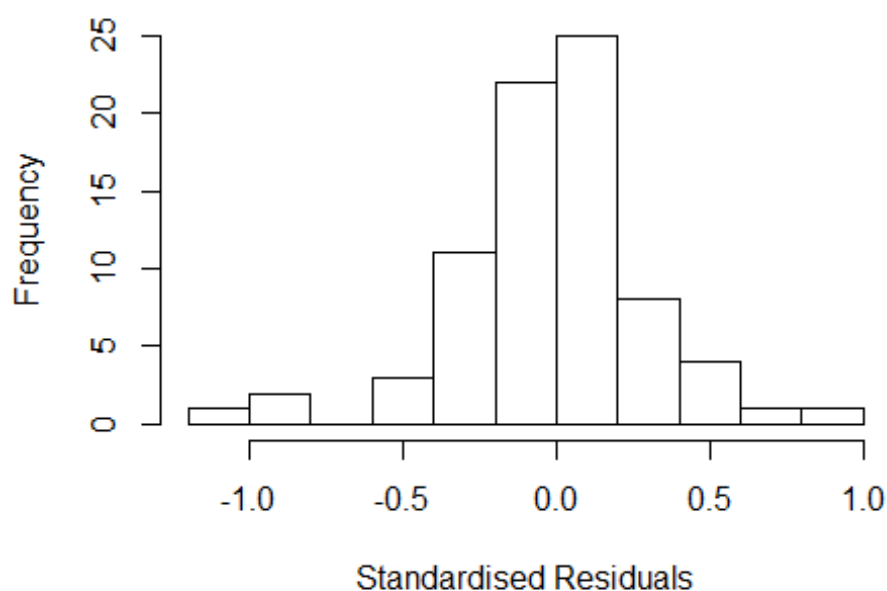
## Year2002      -4.25e-01   1.22e-01 -3.48e+00   0.00093 ***
## Year2003      -2.82e-01   4.18e-01 -6.80e-01   0.50164
## Year2004      -4.81e-01   1.82e-01 -2.64e+00   0.01059 *
## Year2005      -4.31e-01   7.19e-02 -5.99e+00   1.3e-07 ***
## Year2006      -4.26e-01   3.01e-01 -1.41e+00   0.16237
## Year2007      -3.07e-01   2.07e-01 -1.48e+00   0.14300
## Year2008      -7.20e-01   1.21e-01 -5.93e+00   1.6e-07 ***
## Year2009      -4.01e-01   2.52e-01 -1.59e+00   0.11712
## Year2010      -9.17e-01   1.05e-01 -8.73e+00   2.8e-12 ***
## Year2011      -3.72e-01   8.87e-02 -4.19e+00   9.3e-05 ***
## Year2012      -6.50e-01   1.55e-01 -4.20e+00   9.1e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.261
## Multiple R-squared:  0.455, Adjusted R-squared:  0.3
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 69 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0861 0.8310 0.9670 0.8750 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      1.28e-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"

## 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
## -1.03253 -0.13750  0.00169  0.10246  0.80143
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.6630     0.0000      Inf < 2e-16 ***
## FirstAuthorFemale1  0.0012     0.1262    0.01  0.99243
## Year1998         -1.0752     0.1262   -8.52  5.7e-12 ***
## Year1999         -1.2195     0.0367  -33.25 < 2e-16 ***
## Year2000          -0.7464     0.0396  -18.86 < 2e-16 ***
## Year2001          -0.2860     0.0000   -Inf < 2e-16 ***
## Year2002          -0.4256     0.1210   -3.52  0.00083 ***
## Year2003          -0.2791     0.4166   -0.67  0.50547
```

```

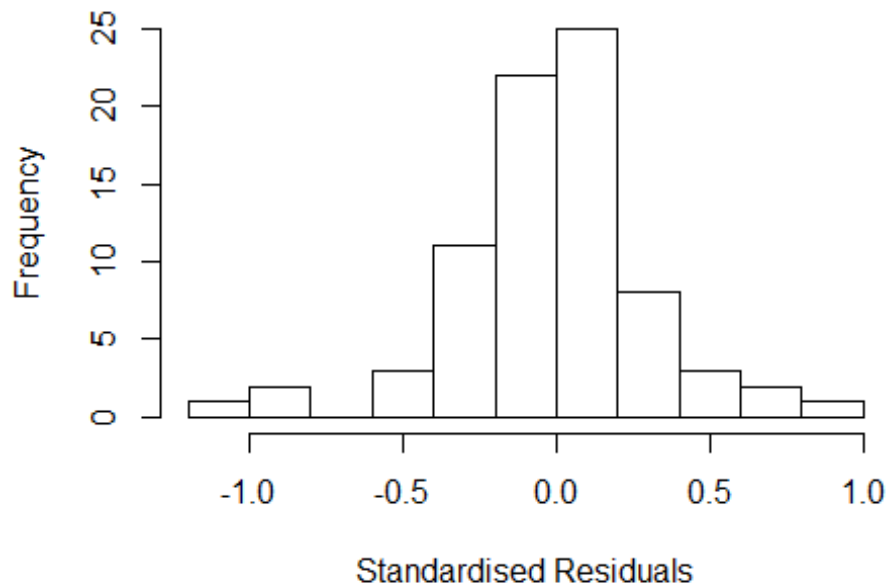
## Year2004          -0.4760      0.1695    -2.81  0.00669 **
## Year2005          -0.4310      0.0719    -5.99  1.2e-07 ***
## Year2006          -0.4243      0.3071    -1.38  0.17210
## Year2007          -0.3012      0.1895    -1.59  0.11713
## Year2008          -0.7108      0.0958    -7.42  4.4e-10 ***
## Year2009          -0.3987      0.2585    -1.54  0.12812
## Year2010          -0.9164      0.1070    -8.56  4.8e-12 ***
## Year2011          -0.3714      0.0899    -4.13  0.00011 ***
## Year2012          -0.6482      0.1567    -4.14  0.00011 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.26
## Multiple R-squared:  0.458, Adjusted R-squared:  0.316
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~ 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0779 0.8250 0.9650 0.8700 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      1.28e-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"

## 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

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.03225 -0.13790  0.00158  0.10673  0.80705
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.66e+00   7.94e-09  2.09e+08 < 2e-16 ***
## LastAuthorFemale1 1.03e-02   1.16e-01  9.00e-02  0.92969
## Year1998        -1.07e+00   3.74e-09 -2.87e+08 < 2e-16 ***
## Year1999        -1.22e+00   3.67e-02 -3.32e+01 < 2e-16 ***
## Year2000         -7.46e-01   2.62e-02 -2.85e+01 < 2e-16 ***
## Year2001         -2.86e-01   0.00e+00    -Inf < 2e-16 ***
## Year2002         -4.25e-01   1.04e-01 -4.08e+00  0.00013 ***
## Year2003         -2.56e-01   2.48e-01 -1.03e+00  0.30649
```



```

## Year2004          -4.82e-01    1.83e-01 -2.64e+00    0.01058 *
## Year2005          -4.31e-01    7.20e-02 -5.98e+00    1.2e-07 ***
## Year2006          -4.17e-01    2.21e-01 -1.89e+00    0.06402 .
## Year2007          -3.07e-01    2.19e-01 -1.40e+00    0.16538
## Year2008          -7.17e-01    1.20e-01 -5.97e+00    1.3e-07 ***
## Year2009          -3.98e-01    2.15e-01 -1.85e+00    0.06912 .
## Year2010          -9.22e-01    8.36e-02 -1.10e+01    3.7e-16 ***
## Year2011          -3.71e-01    4.06e-02 -9.13e+00    5.1e-13 ***
## Year2012          -6.48e-01    1.34e-01 -4.84e+00    9.1e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.252
## Multiple R-squared:  0.471, Adjusted R-squared:  0.332
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
##  9 weights are ~= 1. The remaining 69 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0569 0.8230 0.9650 0.8680 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      1.28e-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 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
##   10   11    7    5    7   10   15    7   16   15   14   18   13   24   19
## 2011 2012
##   33   31
##

```

```

## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      1      6      2      1      1      3      5      5      10      4      10      5      8      11      10
## 2011 2012
##      19      18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      1      3      1      0      1      3      5      5      7      3      9      4      4      9      9
## 2011 2012
##      15      13
## [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: 2.635961258426"

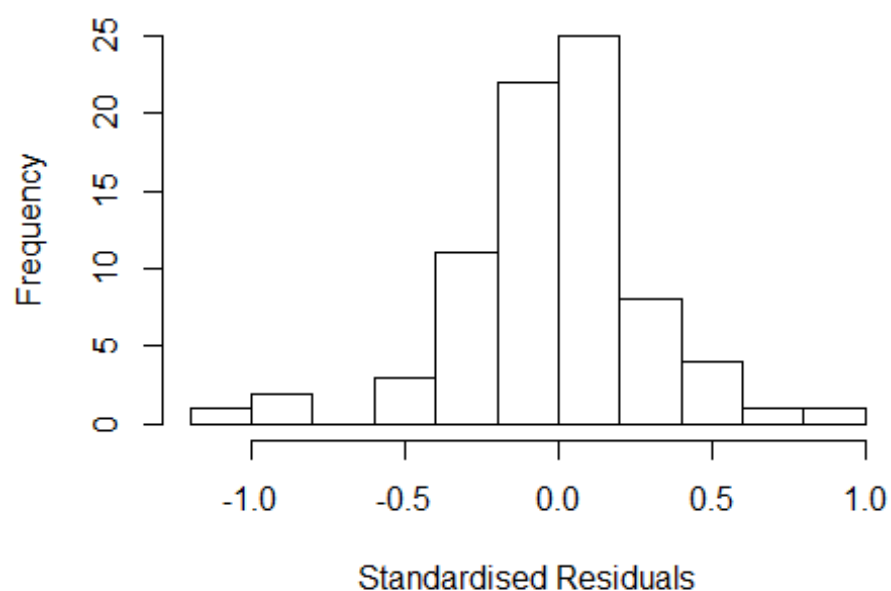
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 9, p-value = 0.1
## 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: 2.50206680971815"

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

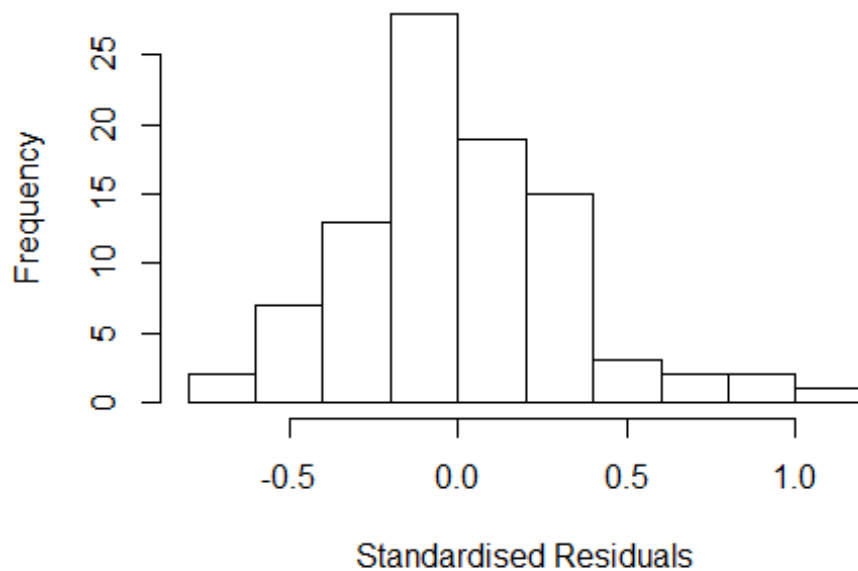
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 14, 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 -3.888e+13  1      NaN
## LastAuthorFemale  -6.445e+13  1      NaN
## UniqueAuthors    -7.968e+16  4      NaN
## Year              -2.778e+30 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
## -0.7560 -0.1898 -0.0032 0.2026 1.0026
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.665116 0.117547 14.17 < 2e-16 ***
## FirstAuthorFemale1 0.094701 0.080865 1.17 0.2455
## LastAuthorFemale1 -0.166191 0.104115 -1.60 0.1149
## UniqueAuthors2 0.017884 0.117547 0.15 0.8795
## UniqueAuthors3 0.000832 0.138387 0.01 0.9952
## UniqueAuthors4 0.112579 0.123337 0.91 0.3645
## UniqueAuthors5 0.235015 0.138394 1.70 0.0939 .
## Year1997 -0.606138 0.421718 -1.44 0.1551
## Year1998 -0.670627 0.217581 -3.08 0.0029 **
## Year2000 -1.665116 0.117547 -14.17 < 2e-16 ***
```

```

## Year2001      -0.933968    0.181696   -5.14  2.4e-06 ***
## Year2002      -0.281128    0.204330   -1.38  0.1733
## Year2003      -1.092421    0.394065   -2.77  0.0071 **
## Year2004      -0.623476    0.120320   -5.18  2.0e-06 ***
## Year2005      -0.259806    0.127252   -2.04  0.0450 *
## Year2006      -0.592122    0.135771   -4.36  4.4e-05 ***
## Year2007      -0.241639    0.177039   -1.36  0.1767
## Year2008      -0.327017    0.183303   -1.78  0.0788 .
## Year2009      -0.577611    0.176873   -3.27  0.0017 **
## Year2010      -0.628230    0.083395   -7.53  1.3e-10 ***
## Year2011      -0.581514    0.114584   -5.08  3.1e-06 ***
## Year2012      -0.944444    0.150275   -6.28  2.4e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.33
## Multiple R-squared:  0.438, Adjusted R-squared:  0.269
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 81 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.335  0.897  0.959  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           1.09e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01             5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500           50           2           1           1000           200
## trace.lev      mts      compute.rd
##           0           1000           0
##           psi             subsampling             cov
##           "bisquare"             "nonsingular"             ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.036e+13 1 3.219e+06
## LastAuthorFemale 1.587e+13 1 3.984e+06
## Year 4.408e+13 15 2.850e+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
## -8.09e-01 -1.89e-01 -2.66e-15  2.04e-01  9.65e-01
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.68e+00   1.51e-08  1.11e+08 < 2e-16 ***
## FirstAuthorFemale1 1.00e-01   6.89e-02  1.45e+00  0.1510
## LastAuthorFemale1 -1.63e-01   8.53e-02 -1.91e+00  0.0596 .
## Year1997        -6.25e-01   4.21e-01 -1.49e+00  0.1416
## Year1998        -6.97e-01   1.22e-01 -5.73e+00  2.0e-07 ***
## Year2000        -1.68e+00   3.48e-08 -4.84e+07 < 2e-16 ***
## Year2001        -9.47e-01   1.57e-01 -6.03e+00  6.0e-08 ***
## Year2002        -2.69e-01   1.53e-01 -1.76e+00  0.0832 .
## Year2003        -1.10e+00   3.88e-01 -2.83e+00  0.0059 **
## Year2004        -6.07e-01   1.32e-01 -4.59e+00  1.8e-05 ***
## Year2005        -1.99e-01   7.81e-02 -2.55e+00  0.0130 *
## Year2006        -6.03e-01   1.21e-01 -4.99e+00  3.9e-06 ***
## Year2007        -2.62e-01   1.44e-01 -1.82e+00  0.0730 .
## Year2008        -3.50e-01   1.51e-01 -2.31e+00  0.0235 *
## Year2009        -5.33e-01   1.16e-01 -4.61e+00  1.7e-05 ***
## Year2010        -5.91e-01   7.23e-02 -8.17e+00  6.1e-12 ***
## Year2011        -5.61e-01   8.97e-02 -6.25e+00  2.4e-08 ***
## Year2012        -8.74e-01   1.25e-01 -6.97e+00  1.1e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.339
## Multiple R-squared:  0.406, Adjusted R-squared:  0.27
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 83 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.397  0.883  0.965  0.911  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      1.09e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500            50          2            1            1000      200
##      trace.lev      mts      compute.rd

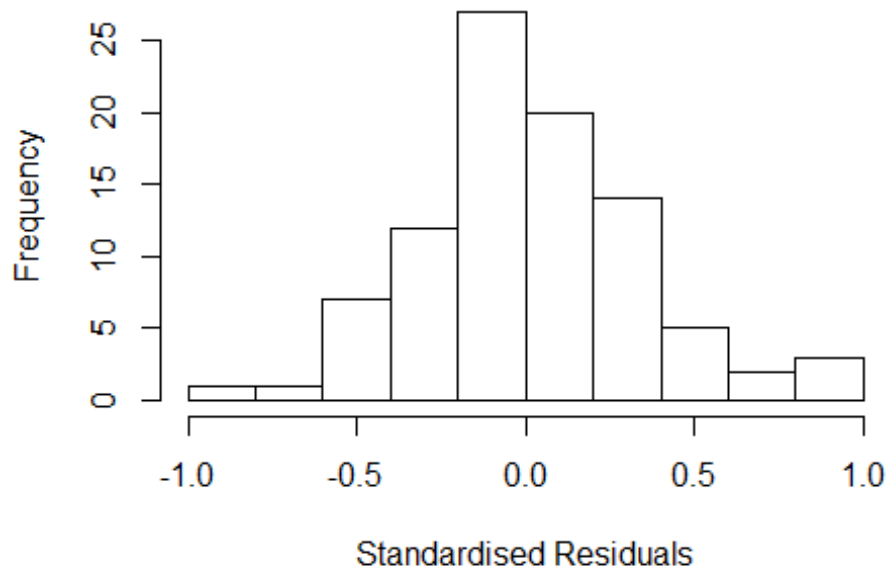
```

```
##           0           1000           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
## -0.75271 -0.18774 -0.00992  0.20029  1.00161
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.6830     0.0000      Inf < 2e-16 ***
## FirstAuthorFemale1  0.0962     0.0697       1.38  0.1714
## Year1997          -0.6273     0.4178      -1.50  0.1374
## Year1998          -0.8562     0.0697     -12.29 < 2e-16 ***
## Year2000          -1.6830     0.0000     -Inf < 2e-16 ***
## Year2001          -0.9456     0.1563      -6.05 5.3e-08 ***
## Year2002          -0.2671     0.1522      -1.76  0.0833 .
## Year2003          -1.1155     0.4477      -2.49  0.0149 *
## Year2004          -0.6309     0.1441      -4.38 3.8e-05 ***
## Year2005          -0.2487     0.1205      -2.06  0.0424 *
## Year2006          -0.6620     0.1014      -6.53 7.0e-09 ***
## Year2007          -0.2853     0.1848      -1.54  0.1269
## Year2008          -0.3940     0.1381      -2.85  0.0056 **
## Year2009          -0.5778     0.1075      -5.38 8.3e-07 ***
## Year2010          -0.5897     0.0725      -8.14 6.5e-12 ***
## Year2011          -0.5976     0.0824      -7.25 3.2e-10 ***
## Year2012          -0.9360     0.1109      -8.44 1.8e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.342
## Multiple R-squared:  0.389, Adjusted R-squared:  0.258
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
##  13 weights are ~ = 1. The remaining 79 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.371  0.901  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      1.09e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500           50           2           1           1000      200
##      trace.lev      mts      compute.rd
##      0             1000          0
##      psi          subsampling          cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"

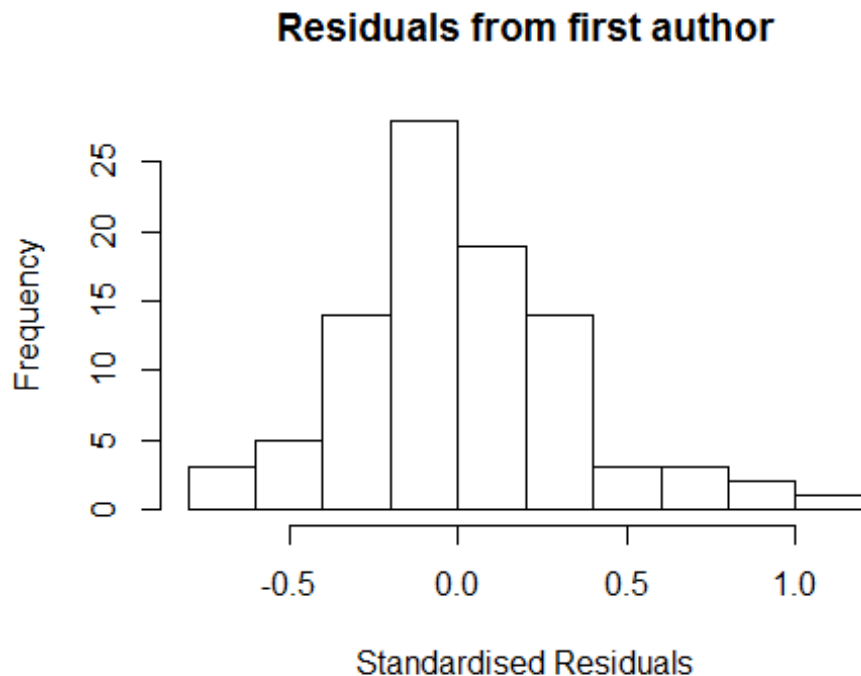
```



```
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"

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

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



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

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q   Median      3Q      Max
## -8.39e-01 -1.78e-01  1.67e-16  2.09e-01  9.31e-01
##
## Coefficients:
```

```

##               Estimate Std. Error   t value Pr(>|t|)
## (Intercept)    1.68e+00   3.03e-08  5.56e+07 < 2e-16 ***
## LastAuthorFemale1 -1.62e-01  8.73e-02 -1.85e+00  0.06778 .
## Year1997        -6.31e-01  4.12e-01 -1.53e+00  0.12951
## Year1998        -5.98e-01  8.73e-02 -6.86e+00  1.7e-09 ***
## Year2000        -1.68e+00  0.00e+00   -Inf < 2e-16 ***
## Year2001        -9.12e-01  1.34e-01 -6.80e+00  2.2e-09 ***
## Year2002        -2.28e-01  1.29e-01 -1.76e+00  0.08190 .
## Year2003        -1.03e+00  3.59e-01 -2.87e+00  0.00529 **
## Year2004        -6.07e-01  1.31e-01 -4.62e+00  1.6e-05 ***
## Year2005        -1.65e-01  8.47e-02 -1.95e+00  0.05449 .
## Year2006        -5.64e-01  1.11e-01 -5.08e+00  2.7e-06 ***
## Year2007        -2.39e-01  1.62e-01 -1.48e+00  0.14435
## Year2008        -2.95e-01  1.57e-01 -1.88e+00  0.06424 .
## Year2009        -5.05e-01  1.25e-01 -4.05e+00  0.00012 ***
## Year2010        -5.56e-01  6.86e-02 -8.10e+00  7.6e-12 ***
## Year2011        -5.27e-01  8.66e-02 -6.09e+00  4.5e-08 ***
## Year2012        -8.44e-01  1.15e-01 -7.36e+00  1.9e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.347
## Multiple R-squared:  0.391, Adjusted R-squared:  0.261
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 79 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.452  0.881  0.962  0.910  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.09e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 92"
## [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]"
##
## 1997 1999 2000 2001 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    1    2    1    2    2    1    2    5    2    6    6    5    9
##
## 1997 1999 2000 2001 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    0    1    1    2    0    1    1    4    2    2    4    5    5
##
## 1997 1999 2000 2001 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    0    1    1    2    0    1    1    4    2    2    4    5    5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 5.3835632709553"
## [1] "Male first author team size 2018 geometric mean: 3.93597934253086"

## Warning in wilcox.test.default(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 = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.29150262212918"
## [1] "Male last 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

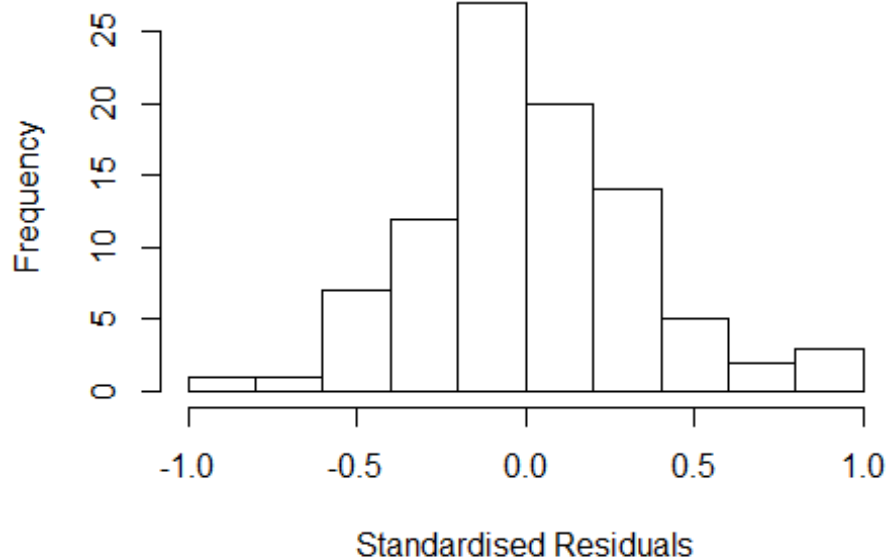
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7.5, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
## [1] "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 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
## -1.64e+00 -1.49e-01  2.39e-16  1.32e-01  5.25e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.50672    0.16709     9.02 1.9e-07 ***
## FirstAuthorFemale1  0.19028    0.16709     1.14  0.2727
## LastAuthorFemale1 -0.18468    0.19180    -0.96  0.3509
## Year2000        -0.04672    0.16709    -0.28  0.7836
## Year2001        -0.16536    0.15868    -1.04  0.3139
## Year2005        -0.41800    0.00000   -Inf < 2e-16 ***
## Year2006        -0.04872    0.16709    -0.29  0.7746
```

```

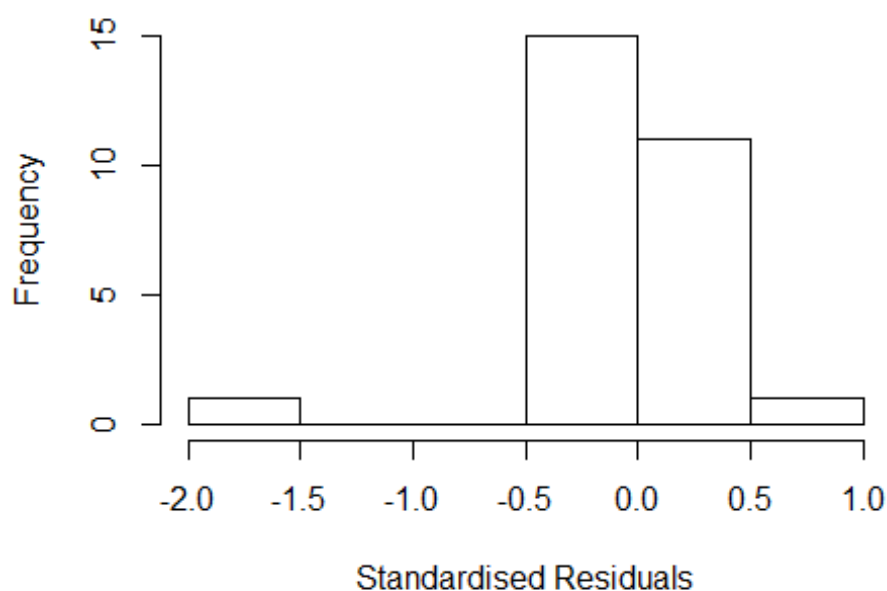
## Year2007          0.35876    0.24555    1.46    0.1646
## Year2008          -0.46922    0.35107   -1.34    0.2013
## Year2009          -0.00472    0.18299   -0.03    0.9798
## Year2010          -0.73743    0.24367   -3.03    0.0085 **
## Year2011          -0.14139    0.16933   -0.83    0.4168
## Year2012          -0.35151    0.13742   -2.56    0.0219 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.373
## Multiple R-squared:  0.621, Adjusted R-squared:  0.318
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 23 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.014  0.928  0.978  0.917  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] "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
## -1.41e+00 -1.39e-01 -1.50e-15  1.34e-01  5.23e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.6145     0.1569   10.29 1.8e-08 ***
## FirstAuthorFemale1  0.0825     0.1569    0.53 0.6065
## Year2000          -0.1545     0.1569   -0.98 0.3394
## Year2001          -0.2193     0.1947   -1.13 0.2766
## Year2005          -0.4180     0.0000    -Inf < 2e-16 ***
## Year2006          -0.1565     0.1569   -1.00 0.3334
## Year2007           0.1280     0.3275    0.39 0.7010
## Year2008          -0.5770     0.4098   -1.41 0.1782
```

```

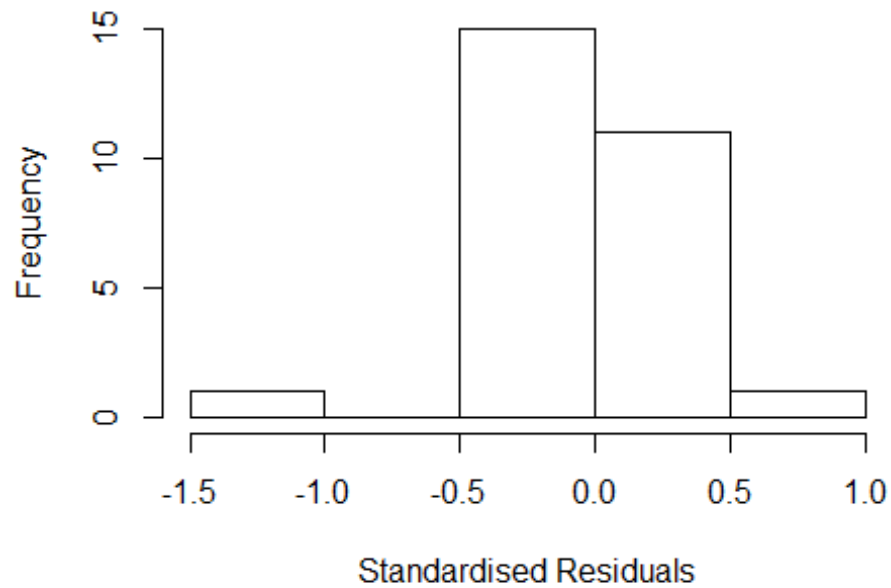
## Year2009          -0.1125      0.1737   -0.65   0.5262
## Year2010          -0.8437      0.2302   -3.66   0.0021 **
## Year2011          -0.2589      0.0778   -3.33   0.0043 **
## Year2012          -0.4554      0.1744   -2.61   0.0189 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.407
## Multiple R-squared:  0.509, Adjusted R-squared:  0.171
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 23 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.208  0.918  0.973  0.929  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      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

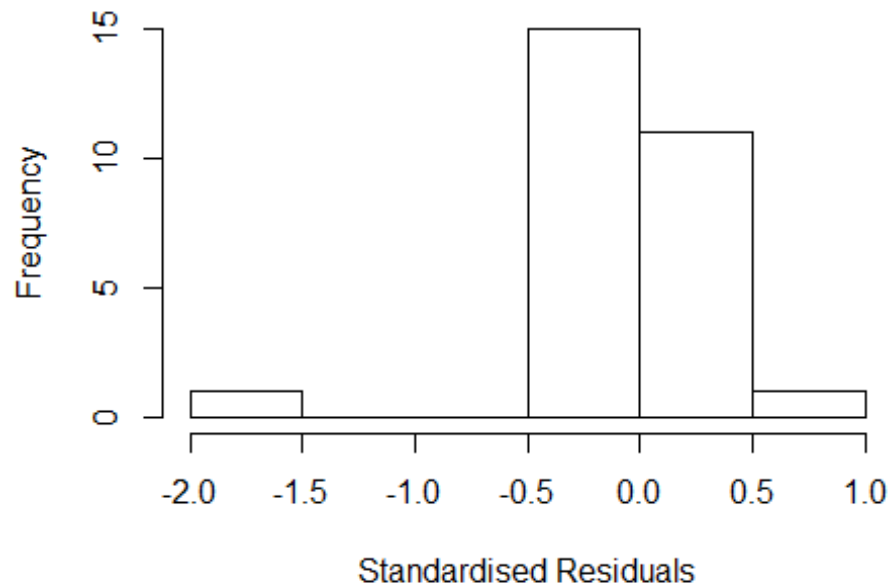
```

Residuals from first author



```
##          GVIF Df GVIF^(1/(2*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
## -1.42e+00 -1.12e-01 -2.50e-16 1.17e-01 5.24e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6970 0.0000 Inf < 2e-16 ***
## LastAuthorFemale1 -0.0868 0.1293 -0.67 0.51162
## Year2000 -0.2370 0.0000 -Inf < 2e-16 ***
## Year2001 -0.2605 0.2180 -1.19 0.24958
## Year2005 -0.4180 0.0000 -Inf < 2e-16 ***
## Year2006 -0.2390 0.0000 -Inf < 2e-16 ***
## Year2007 0.1438 0.2248 0.64 0.53148
## Year2008 -0.6595 0.3492 -1.89 0.07723 .
## Year2009 -0.1950 0.0746 -2.62 0.01875 *
## Year2010 -0.9274 0.1985 -4.67 0.00026 ***
## Year2011 -0.2566 0.0694 -3.70 0.00195 **
## Year2012 -0.4824 0.1779 -2.71 0.01539 *
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.379
## Multiple R-squared: 0.542, Adjusted R-squared: 0.227
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 23 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.128 0.925 0.979 0.919 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 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

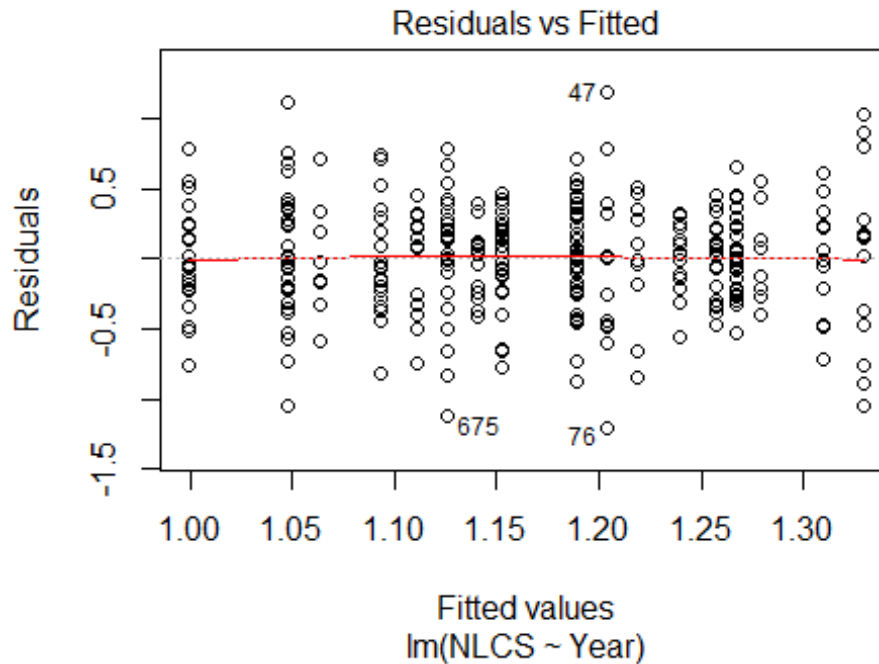
```

```

##          "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 2401"
## [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 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
## 31 31 25 39 34 32 30 32 27 47 42 55 34 42 44
## 2011 2012
## 70 48
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 13 14 12 15 9 8 10 14 16 25 22 31 20 26 27
## 2011 2012
## 51 30
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 11 12 9 12 8 6 7 10 13 20 17 30 17 21 22
## 2011 2012
## 38 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 = 45, df = 16, p-value = 1e-04
```



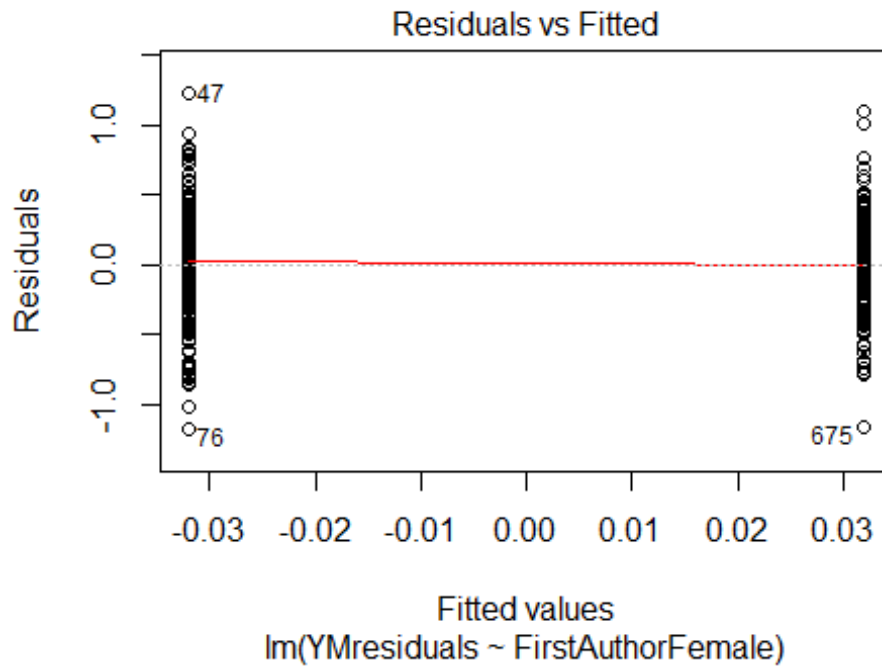
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 7.8, df = 1, p-value = 0.005

## [1] "Female first author team size 2018 geometric mean: 3.0861636884341"
## [1] "Male first author team size 2018 geometric mean: 4.06980322848949"

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

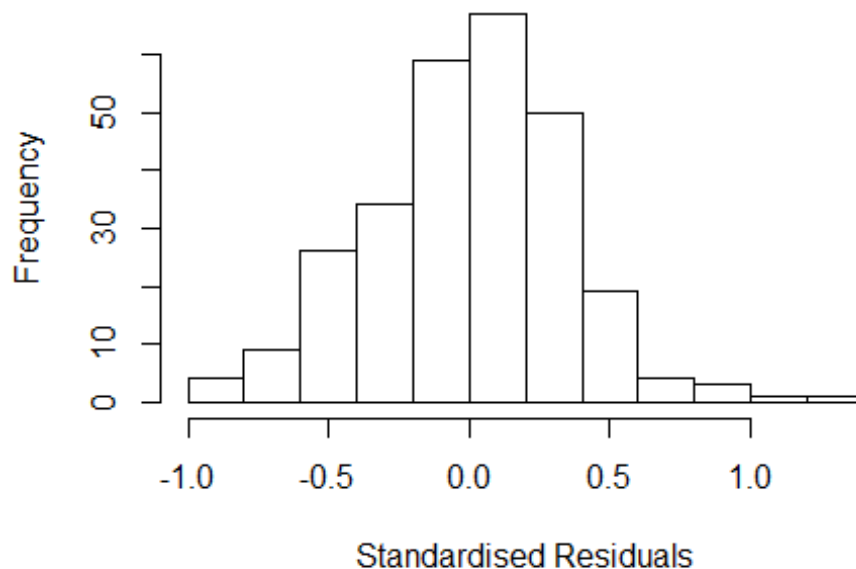
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 14, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.63424118566428"
## [1] "Male last author team size 2018 geometric mean: 3.60945639919211"

## Warning in wilcox.test.default(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 1.354 1          1.164
## LastAuthorFemale  1.405 1          1.185
## UniqueAuthors    3.367 4          1.164
## Year              4.447 16          1.048
```

Residuals from first and last author and team size



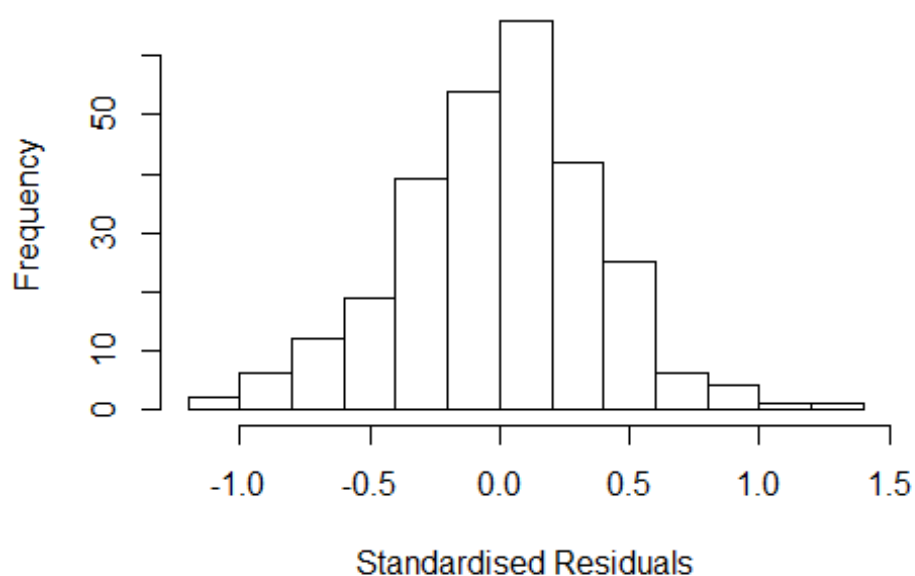
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9383 -0.2262 0.0283 0.2268 1.3570
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7761 0.2983 2.60 0.00981 **
## FirstAuthorFemale1 -0.0219 0.0453 -0.48 0.62936
## LastAuthorFemale1 -0.1307 0.0614 -2.13 0.03421 *
## UniqueAuthors2 0.4580 0.1843 2.48 0.01361 *
## UniqueAuthors3 0.5949 0.1855 3.21 0.00151 **
## UniqueAuthors4 0.6385 0.1872 3.41 0.00076 ***
## UniqueAuthors5 0.7867 0.1847 4.26 2.9e-05 ***
## Year1997 -0.1970 0.2749 -0.72 0.47429
## Year1998 0.0376 0.2751 0.14 0.89138
## Year1999 -0.2288 0.2651 -0.86 0.38900
```

```

## Year2000          -0.0238      0.2583   -0.09  0.92654
## Year2001          -0.2916      0.2662   -1.10  0.27440
## Year2002          -0.0818      0.2942   -0.28  0.78116
## Year2003          -0.1542      0.2472   -0.62  0.53321
## Year2004          -0.1210      0.2553   -0.47  0.63596
## Year2005          -0.4208      0.2566   -1.64  0.10230
## Year2006          -0.0527      0.2455   -0.21  0.83021
## Year2007          -0.2020      0.2401   -0.84  0.40106
## Year2008          -0.3226      0.2474   -1.30  0.19343
## Year2009          -0.1369      0.2451   -0.56  0.57703
## Year2010          -0.1923      0.2540   -0.76  0.44972
## Year2011          -0.2537      0.2422   -1.05  0.29590
## Year2012          -0.2957      0.2650   -1.12  0.26563
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.331
## Multiple R-squared:  0.252, Adjusted R-squared:  0.187
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 257 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0543 0.8670 0.9490 0.8990 0.9840 0.9990
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          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 1.323 1          1.150
## LastAuthorFemale 1.394 1          1.181
## Year          1.812 16          1.019

```

Residuals from first and last author



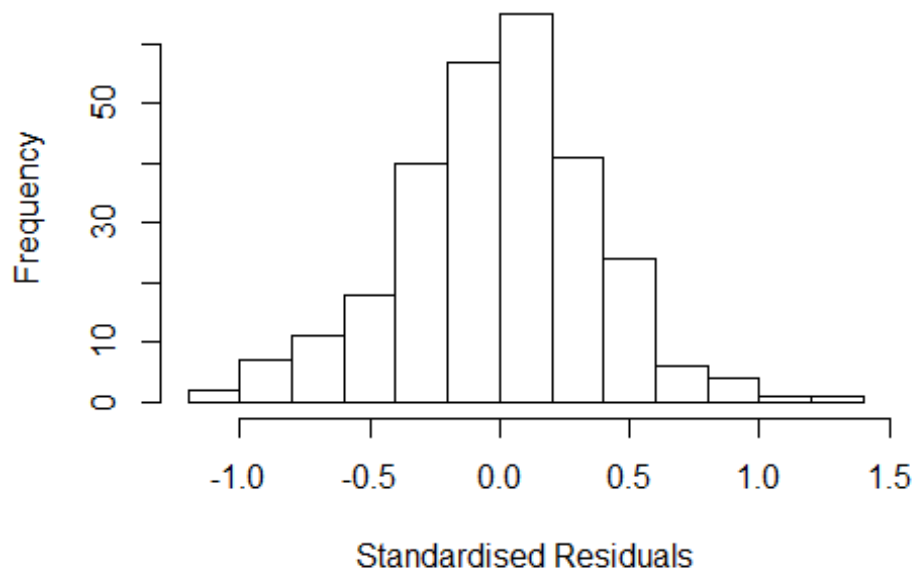
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1226 -0.2332  0.0166  0.2303  1.3605
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3318    0.2352   5.66 4e-08 ***
## FirstAuthorFemale1  0.0364    0.0496   0.73  0.46
## LastAuthorFemale1 -0.0648    0.0707  -0.92  0.36
## Year1997          -0.2983    0.2994  -1.00  0.32
## Year1998          -0.0267    0.2688  -0.10  0.92
## Year1999          -0.2933    0.2551  -1.15  0.25
## Year2000          -0.0211    0.2569  -0.08  0.93
## Year2001          -0.3463    0.2651  -1.31  0.19
## Year2002          -0.0579    0.3259  -0.18  0.86
## Year2003          -0.1569    0.2497  -0.63  0.53
## Year2004          -0.0961    0.2427  -0.40  0.69
## Year2005          -0.3415    0.2443  -1.40  0.16
```

```

## Year2006          -0.0667      0.2408   -0.28      0.78
## Year2007          -0.1705      0.2378   -0.72      0.47
## Year2008          -0.3486      0.2442   -1.43      0.15
## Year2009          -0.1190      0.2386   -0.50      0.62
## Year2010          -0.1808      0.2517   -0.72      0.47
## Year2011          -0.1939      0.2372   -0.82      0.41
## Year2012          -0.3539      0.2670   -1.33      0.19
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.354
## Multiple R-squared:  0.0869, Adjusted R-squared:  0.0232
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 252 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.106  0.868  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      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 1.319 1          1.148
## Year              1.319 16          1.009

```


Residuals from first author



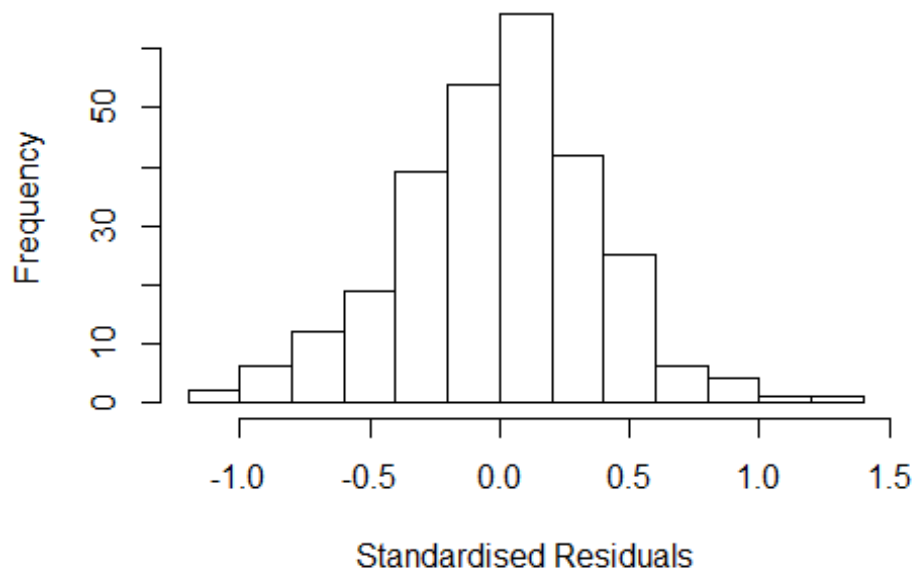
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1788 -0.2223  0.0135  0.2334  1.3592
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3210     0.2303   5.74 2.7e-08 ***
## FirstAuthorFemale1  0.0337     0.0495   0.68  0.50
## Year1997         -0.2862     0.2953  -0.97  0.33
## Year1998         -0.0244     0.2661  -0.09  0.93
## Year1999         -0.2855     0.2523  -1.13  0.26
## Year2000         -0.0187     0.2542  -0.07  0.94
## Year2001         -0.3338     0.2604  -1.28  0.20
## Year2002         -0.0527     0.3289  -0.16  0.87
## Year2003         -0.1527     0.2457  -0.62  0.53
## Year2004         -0.0840     0.2377  -0.35  0.72
## Year2005         -0.3378     0.2404  -1.41  0.16
## Year2006         -0.0615     0.2369  -0.26  0.80
```

```

## Year2007          -0.1681      0.2339   -0.72      0.47
## Year2008          -0.3457      0.2398   -1.44      0.15
## Year2009          -0.1163      0.2343   -0.50      0.62
## Year2010          -0.1758      0.2480   -0.71      0.48
## Year2011          -0.1973      0.2338   -0.84      0.40
## Year2012          -0.3525      0.2632   -1.34      0.18
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.352
## Multiple R-squared:  0.085, Adjusted R-squared:  0.025
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 248 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.103  0.865   0.950   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      3.61e-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 1.405 1          1.185
## Year            1.405 16          1.011

```

Residuals from last author



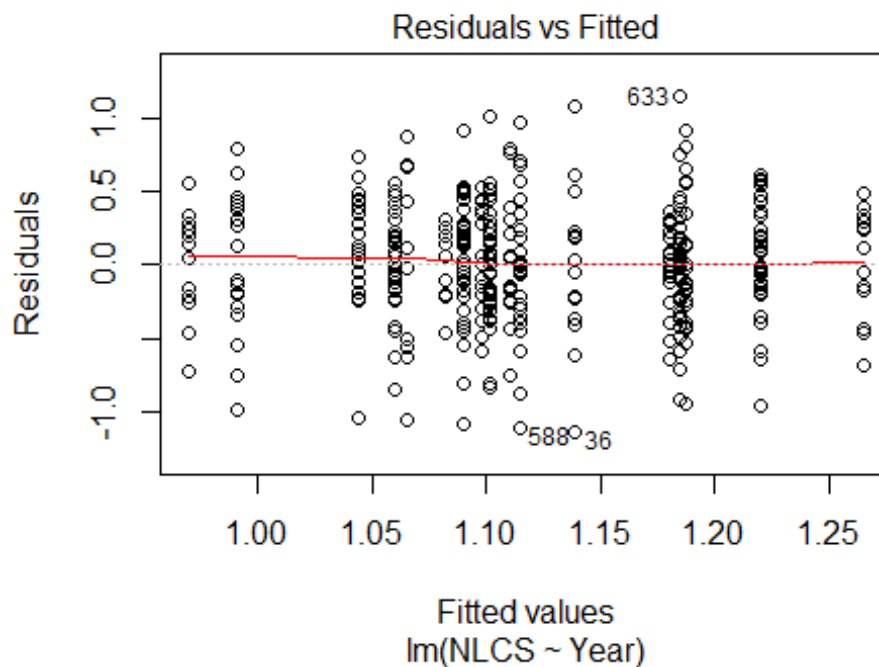
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1092 -0.2341 0.0188 0.2365 1.3378
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3469 0.2263 5.95 8.6e-09 ***
## LastAuthorFemale1 -0.0616 0.0715 -0.86 0.39
## Year1997 -0.2907 0.2954 -0.98 0.33
## Year1998 -0.0259 0.2619 -0.10 0.92
## Year1999 -0.2879 0.2500 -1.15 0.25
## Year2000 -0.0212 0.2504 -0.08 0.93
## Year2001 -0.3404 0.2594 -1.31 0.19
## Year2002 -0.0336 0.3213 -0.10 0.92
## Year2003 -0.1586 0.2428 -0.65 0.51
## Year2004 -0.0961 0.2365 -0.41 0.68
## Year2005 -0.3384 0.2386 -1.42 0.16
## Year2006 -0.0648 0.2350 -0.28 0.78
```

```

## Year2007          -0.1655      0.2317   -0.71      0.48
## Year2008          -0.3454      0.2387   -1.45      0.15
## Year2009          -0.1127      0.2332   -0.48      0.63
## Year2010          -0.1761      0.2463   -0.71      0.48
## Year2011          -0.1864      0.2315   -0.80      0.42
## Year2012          -0.3478      0.2603   -1.34      0.18
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.352
## Multiple R-squared:  0.0845, Adjusted R-squared:  0.0245
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 251 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.117  0.872  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      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 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
##   35   23   23   24   40   31   29   23   26   41   38   35   42   47   52
## 2011 2012
##   49   49
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   15   12   10   15   26    9   15   16   22   32   23   21   28   32   38
## 2011 2012

```

```
## 37 38
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 12 9 10 11 23 7 11 14 14 23 16 18 21 24 33
## 2011 2012
## 34 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 = 28, df = 16, p-value = 0.03
```

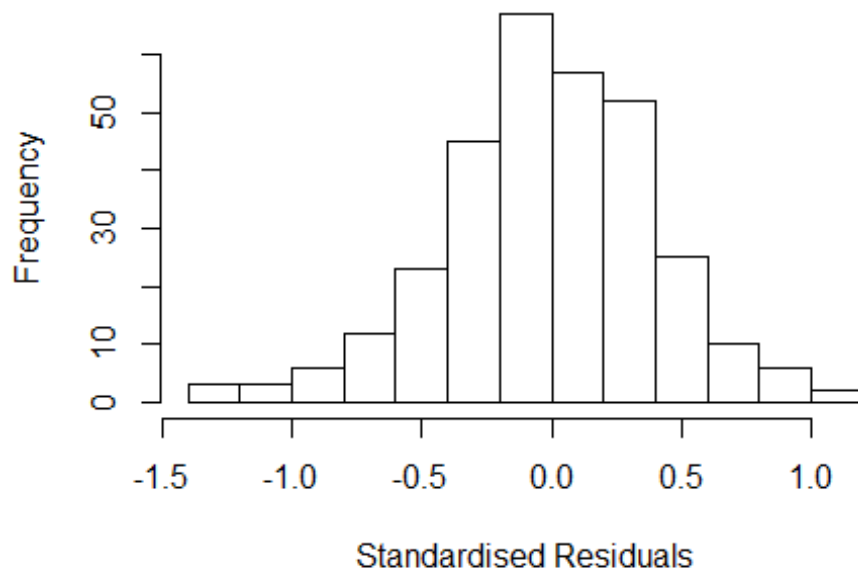


```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.0014, df = 1, p-value = 1

## [1] "Female first author team size 2018 geometric mean: 6.80303169124811"
## [1] "Male first author team size 2018 geometric mean: 5.33104539820977"

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


Residuals from first and last author and team size



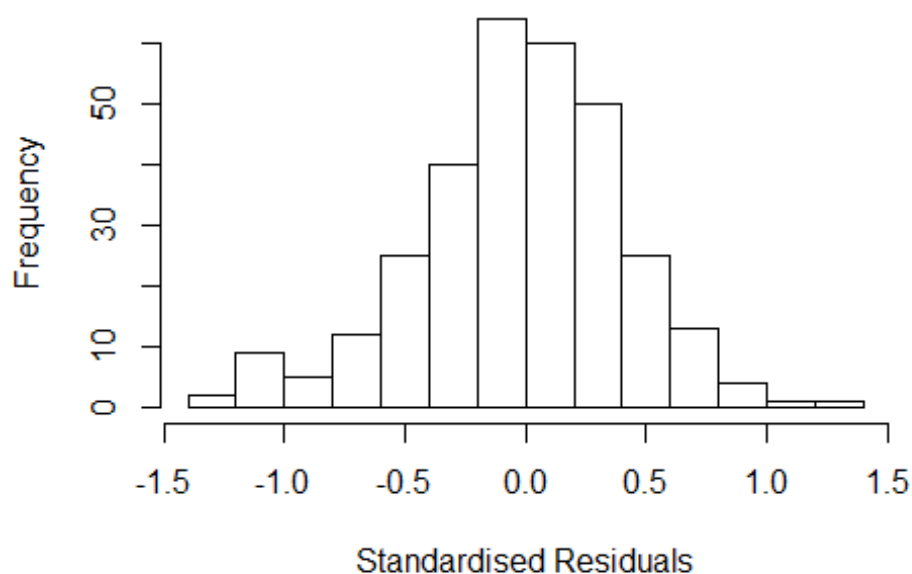
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.39047 -0.25614 -0.00454 0.24766 1.14022
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8929 0.2428 3.68 0.00028 ***
## FirstAuthorFemale1 0.0458 0.0509 0.90 0.36941
## LastAuthorFemale1 -0.0694 0.0634 -1.09 0.27465
## UniqueAuthors2 0.1251 0.1834 0.68 0.49590
## UniqueAuthors3 0.1517 0.1886 0.80 0.42179
## UniqueAuthors4 0.3214 0.1901 1.69 0.09195 .
## UniqueAuthors5 0.3840 0.1844 2.08 0.03821 *
## Year1997 -0.0767 0.2376 -0.32 0.74703
## Year1998 -0.1448 0.3749 -0.39 0.69963
## Year1999 0.0689 0.2345 0.29 0.76927
```

```

## Year2000          0.0679      0.2335      0.29  0.77148
## Year2001         -0.0317      0.2294     -0.14  0.89024
## Year2002          0.0172      0.2468      0.07  0.94452
## Year2003         -0.1328      0.2283     -0.58  0.56114
## Year2004          0.0716      0.2105      0.34  0.73394
## Year2005         -0.1623      0.2135     -0.76  0.44788
## Year2006          0.1020      0.2481      0.41  0.68116
## Year2007         -0.1042      0.2327     -0.45  0.65457
## Year2008         -0.0463      0.2222     -0.21  0.83512
## Year2009          0.0239      0.2170      0.11  0.91245
## Year2010         -0.0784      0.2133     -0.37  0.71362
## Year2011          0.0442      0.2144      0.21  0.83679
## Year2012         -0.0332      0.2172     -0.15  0.87855
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.377
## Multiple R-squared:  0.13,   Adjusted R-squared:  0.064
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 281 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.145  0.860  0.953  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      3.22e-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 1.268 1      1.126
## LastAuthorFemale  1.378 1      1.174
## Year              1.689 16      1.017

```


Residuals from first and last author



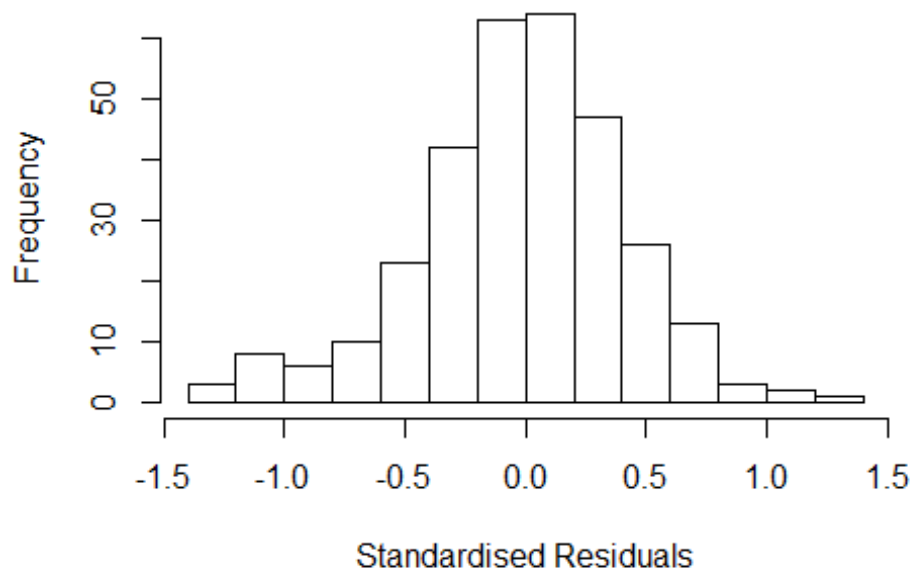
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.21769 -0.26082 -0.00189 0.27723 1.21106
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.00894 0.17469 5.78 2e-08 ***
## FirstAuthorFemale1 0.09546 0.05003 1.91 0.057 .
## LastAuthorFemale1 -0.05464 0.06455 -0.85 0.398
## Year1997 0.00360 0.20952 0.02 0.986
## Year1998 0.07017 0.44699 0.16 0.875
## Year1999 0.22223 0.19833 1.12 0.263
## Year2000 0.11328 0.20977 0.54 0.590
## Year2001 0.06634 0.19396 0.34 0.733
## Year2002 0.14522 0.20600 0.70 0.481
## Year2003 0.00694 0.20161 0.03 0.973
## Year2004 0.15035 0.18062 0.83 0.406
## Year2005 -0.00960 0.18280 -0.05 0.958
```

```

## Year2006          0.24316      0.22569      1.08      0.282
## Year2007          0.04467      0.20778      0.21      0.830
## Year2008          0.04794      0.20137      0.24      0.812
## Year2009          0.13592      0.18551      0.73      0.464
## Year2010          0.02927      0.18034      0.16      0.871
## Year2011          0.19562      0.18040      1.08      0.279
## Year2012          0.11229      0.18533      0.61      0.545
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.386
## Multiple R-squared:  0.0498, Adjusted R-squared:  -0.00879
## Convergence in 27 IRWLS iterations
##
## Robustness weights:
## 32 weights are ~= 1. The remaining 279 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.300  0.860  0.947  0.888  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      3.22e-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 1.248 1          1.117
## Year              1.248 16          1.007

```

Residuals from first author



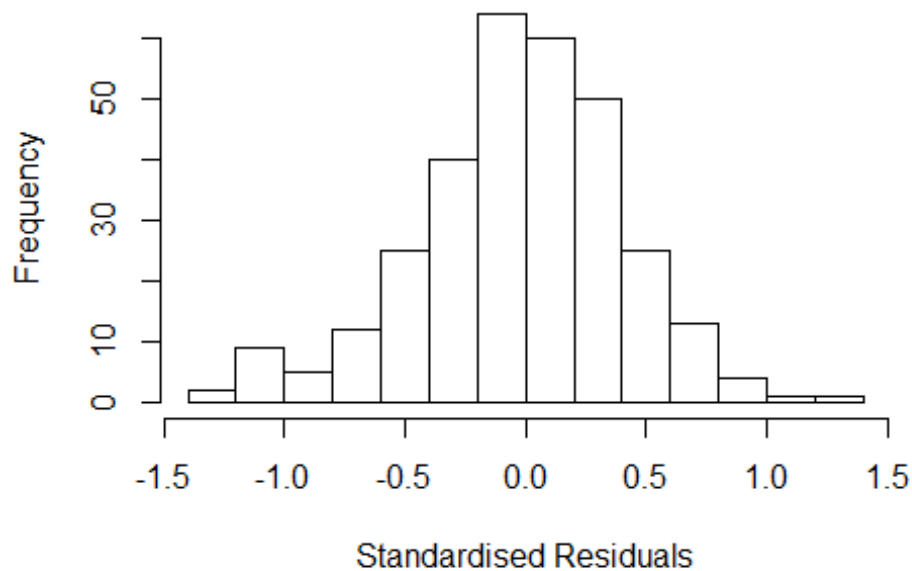
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.20204 -0.25611  0.00253  0.27748  1.22428
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.99572    0.17087   5.83 1.5e-08 ***
## FirstAuthorFemale1 0.09340    0.04987   1.87  0.062 .
## Year1997        0.00313    0.21030   0.01  0.988
## Year1998        0.08498    0.42994   0.20  0.843
## Year1999        0.23616    0.19455   1.21  0.226
## Year2000        0.11292    0.21069   0.54  0.592
## Year2001        0.07236    0.19040   0.38  0.704
## Year2002        0.14175    0.20473   0.69  0.489
## Year2003        0.01633    0.19952   0.08  0.935
## Year2004        0.15679    0.17795   0.88  0.379
## Year2005       -0.00294    0.17967  -0.02  0.987
## Year2006        0.23696    0.22519   1.05  0.294
```

```

## Year2007          0.05654    0.20289    0.28    0.781
## Year2008          0.05675    0.19805    0.29    0.775
## Year2009          0.13821    0.18333    0.75    0.452
## Year2010          0.02613    0.17989    0.15    0.885
## Year2011          0.20283    0.17762    1.14    0.254
## Year2012          0.10725    0.18418    0.58    0.561
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.389
## Multiple R-squared:  0.0466, Adjusted R-squared:  -0.00872
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 35 weights are ~= 1. The remaining 276 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.302  0.867   0.948   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      3.22e-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 1.357 1      1.165
## Year              1.357 16      1.010

```

Residuals from last author



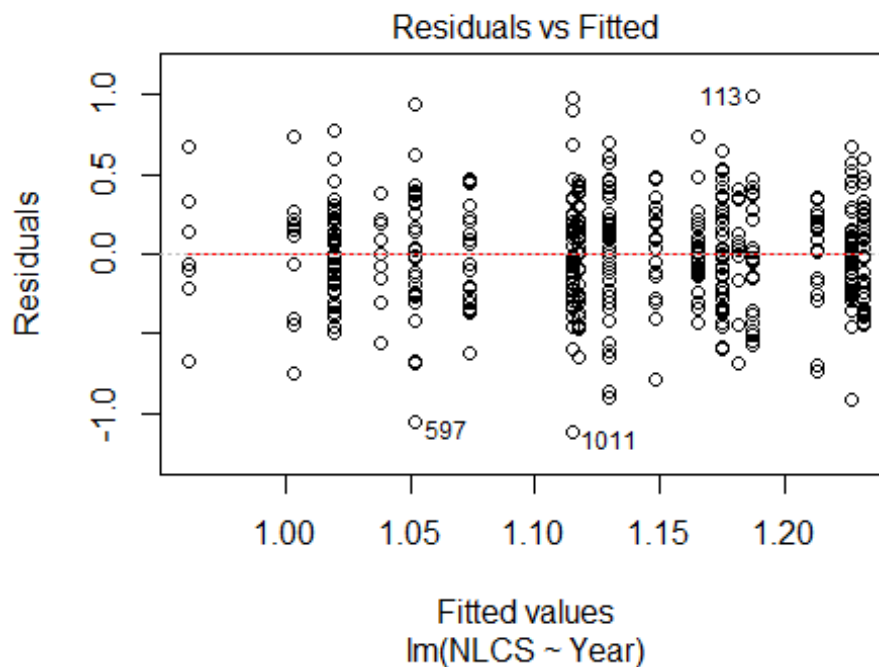
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.19091 -0.25020 0.00756 0.26662 1.15033
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.06967 0.16948 6.31 1e-09 ***
## LastAuthorFemale1 -0.04894 0.06383 -0.77 0.44
## Year1997 -0.02149 0.20612 -0.10 0.92
## Year1998 0.03282 0.37348 0.09 0.93
## Year1999 0.20086 0.19126 1.05 0.29
## Year2000 0.08534 0.20168 0.42 0.67
## Year2001 0.04660 0.20043 0.23 0.82
## Year2002 0.12506 0.20526 0.61 0.54
## Year2003 -0.00293 0.20015 -0.01 0.99
## Year2004 0.13351 0.17915 0.75 0.46
## Year2005 -0.02411 0.18095 -0.13 0.89
## Year2006 0.22964 0.22533 1.02 0.31
```

```

## Year2007      0.03990    0.20884    0.19    0.85
## Year2008      0.02500    0.20093    0.12    0.90
## Year2009      0.14176    0.18482    0.77    0.44
## Year2010      0.01455    0.17880    0.08    0.94
## Year2011      0.18543    0.17904    1.04    0.30
## Year2012      0.12124    0.18506    0.66    0.51
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.397
## Multiple R-squared:  0.037, Adjusted R-squared:  -0.0189
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 285 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.347  0.876  0.954  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      3.22e-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: 311"
## [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
##   25   23   23   30   48   35   31   34   34   57   52   50   52   62   51
## 2011 2012
##   72   49
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    8   13   10   23   25   11   19   22   18   42   32   32   39   46   36
## 2011 2012

```

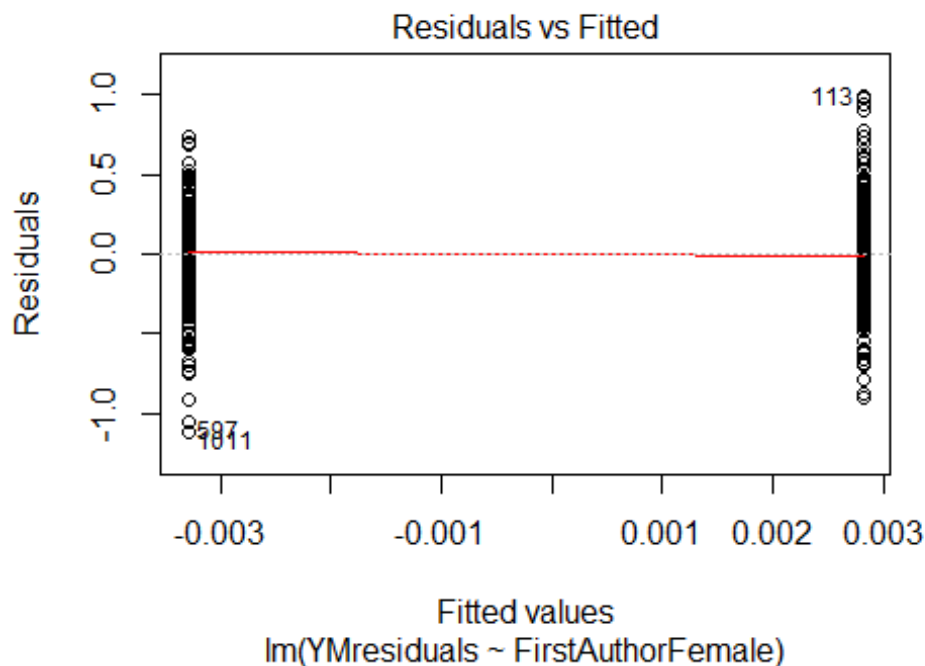
```
## 47 41
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3 9 7 19 19 8 11 20 13 31 23 27 32 37 32
## 2011 2012
## 40 36
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 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: 5.72465317064809"
## [1] "Male first author team size 2018 geometric mean: 4.70381336291021"
## Warning in wilcox.test.default(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: 5.32298202783709"
## [1] "Male last author team size 2018 geometric mean: 4.83856598138417"

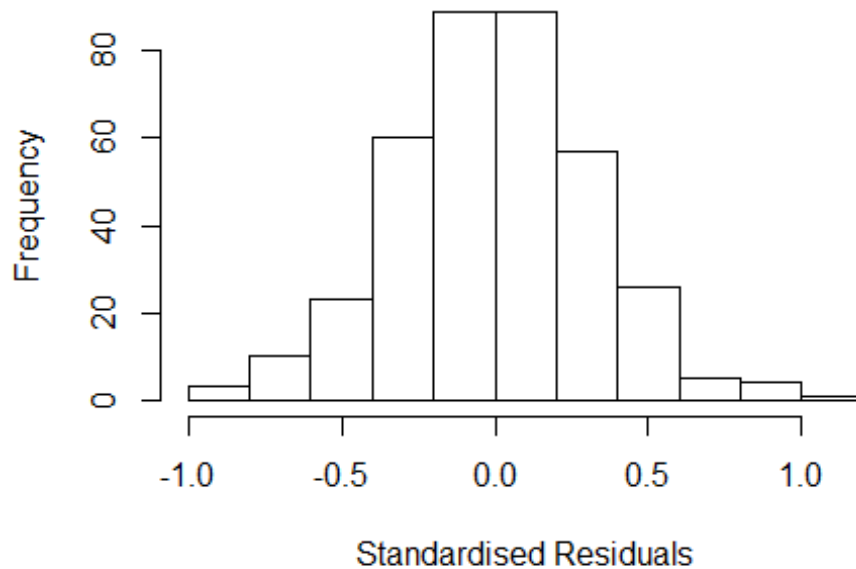
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 120, p-value = 0.5
## 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 | 1.446 | 1 | 1.203 |
| LastAuthorFemale | 1.265 | 1 | 1.125 |
| UniqueAuthors | 2.503 | 4 | 1.122 |
| Year | 2.643 | 16 | 1.031 |

Residuals from first and last author and team size



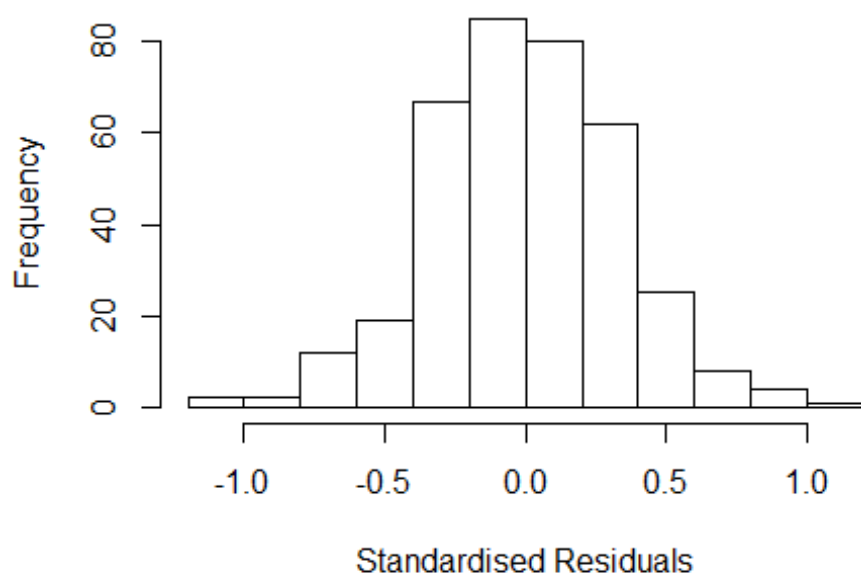
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.98280 -0.21141 -0.00365 0.20729 1.10625
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6951 0.2068 3.36 0.00086 ***
## FirstAuthorFemale1 -0.0555 0.0376 -1.48 0.14100
## LastAuthorFemale1 -0.0177 0.0563 -0.31 0.75339
## UniqueAuthors2 0.3231 0.1528 2.11 0.03521 *
## UniqueAuthors3 0.4700 0.1529 3.07 0.00229 **
## UniqueAuthors4 0.4905 0.1572 3.12 0.00197 **
## UniqueAuthors5 0.4994 0.1565 3.19 0.00154 **
## Year1997 0.1449 0.1588 0.91 0.36212
## Year1998 -0.0247 0.2519 -0.10 0.92185
## Year1999 0.1100 0.1720 0.64 0.52298
```

```

## Year2000          0.0825      0.1515      0.54  0.58639
## Year2001          0.0265      0.1709      0.16  0.87677
## Year2002          0.0217      0.1610      0.13  0.89310
## Year2003         -0.0109      0.1570     -0.07  0.94448
## Year2004          0.1239      0.1627      0.76  0.44676
## Year2005         -0.1345      0.1520     -0.88  0.37706
## Year2006          0.1063      0.1590      0.67  0.50410
## Year2007          0.0312      0.1593      0.20  0.84495
## Year2008          0.0951      0.1538      0.62  0.53685
## Year2009          0.0424      0.1547      0.27  0.78438
## Year2010          0.0114      0.1482      0.08  0.93896
## Year2011          0.1062      0.1469      0.72  0.47028
## Year2012         -0.0317      0.1550     -0.20  0.83784
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.322
## Multiple R-squared:  0.126, Adjusted R-squared:  0.0703
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 27 weights are ~= 1. The remaining 340 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.213  0.883  0.953  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      2.72e-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 1.215 1      1.102
## LastAuthorFemale  1.166 1      1.080
## Year              1.352 16      1.009

```

Residuals from first and last author



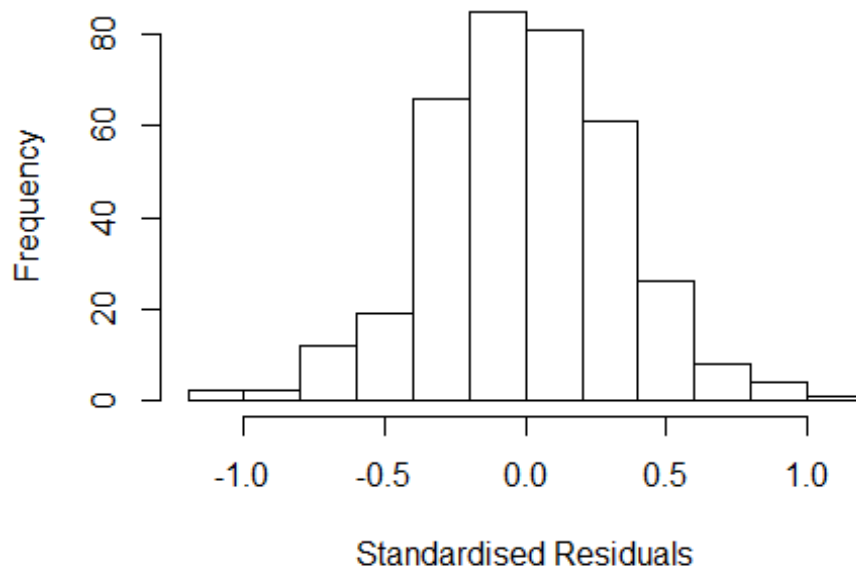
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.11605 -0.23193 -0.00662 0.21429 1.00915
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.10506 0.11727 9.42 <2e-16 ***
## FirstAuthorFemale1 -0.02820 0.03660 -0.77 0.44
## LastAuthorFemale1 0.00373 0.05410 0.07 0.95
## Year1997 0.08914 0.13557 0.66 0.51
## Year1998 -0.15203 0.25022 -0.61 0.54
## Year1999 0.14568 0.15108 0.96 0.34
## Year2000 0.04585 0.12904 0.36 0.72
## Year2001 0.00873 0.14787 0.06 0.95
## Year2002 0.02063 0.14383 0.14 0.89
## Year2003 -0.01208 0.13663 -0.09 0.93
## Year2004 0.12529 0.15087 0.83 0.41
## Year2005 -0.09599 0.12784 -0.75 0.45
```

```

## Year2006          0.12934      0.13418      0.96      0.34
## Year2007          0.00299      0.14223      0.02      0.98
## Year2008          0.10161      0.13173      0.77      0.44
## Year2009          0.08765      0.13064      0.67      0.50
## Year2010          0.04003      0.12490      0.32      0.75
## Year2011          0.14456      0.12174      1.19      0.24
## Year2012          0.01099      0.12901      0.09      0.93
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.334
## Multiple R-squared:  0.0526, Adjusted R-squared:  0.00357
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 34 weights are ~= 1. The remaining 333 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.243  0.883  0.949  0.907  0.984  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      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 1.194 1      1.093
## Year              1.194 16      1.006

```

Residuals from first author



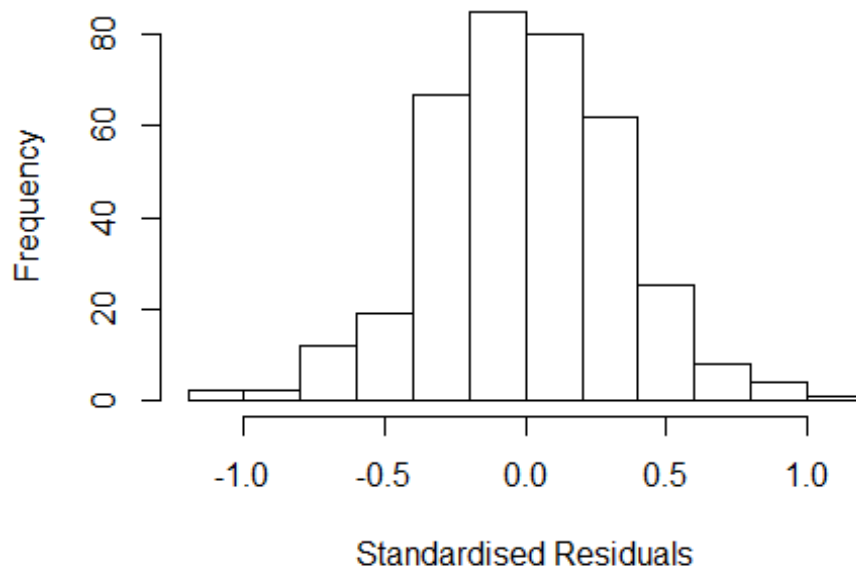
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.11677 -0.23137 -0.00725 0.21444 1.00815
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.10487 0.11714 9.43 <2e-16 ***
## FirstAuthorFemale1 -0.02792 0.03628 -0.77 0.44
## Year1997 0.08958 0.13524 0.66 0.51
## Year1998 -0.15200 0.25024 -0.61 0.54
## Year1999 0.14575 0.15102 0.97 0.34
## Year2000 0.04633 0.12875 0.36 0.72
## Year2001 0.00978 0.14707 0.07 0.95
## Year2002 0.02141 0.14320 0.15 0.88
## Year2003 -0.01151 0.13626 -0.08 0.93
## Year2004 0.12671 0.14861 0.85 0.39
## Year2005 -0.09553 0.12772 -0.75 0.45
## Year2006 0.12978 0.13424 0.97 0.33
```

```

## Year2007          0.00377    0.14115    0.03    0.98
## Year2008          0.10214    0.13146    0.78    0.44
## Year2009          0.08852    0.13033    0.68    0.50
## Year2010          0.04055    0.12514    0.32    0.75
## Year2011          0.14537    0.12137    1.20    0.23
## Year2012          0.01190    0.12878    0.09    0.93
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.334
## Multiple R-squared:  0.0525, Adjusted R-squared:  0.0064
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 34 weights are ~= 1. The remaining 333 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.242  0.880   0.949   0.907   0.984   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      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 1.135 1          1.065
## Year            1.135 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
## -1.0989 -0.2289 -0.0109 0.2193 0.9981
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.085765 0.106027 10.24 <2e-16 ***
## LastAuthorFemale1 -0.000128 0.053633 0.00 1.00
## Year1997 0.089550 0.129858 0.69 0.49
## Year1998 -0.149028 0.251689 -0.59 0.55
## Year1999 0.152116 0.145367 1.05 0.30
## Year2000 0.053282 0.121649 0.44 0.66
## Year2001 0.017758 0.140082 0.13 0.90
## Year2002 0.031941 0.137134 0.23 0.82
## Year2003 -0.005707 0.130172 -0.04 0.97
## Year2004 0.130966 0.144289 0.91 0.36
## Year2005 -0.095007 0.121005 -0.79 0.43
## Year2006 0.134910 0.127401 1.06 0.29
```

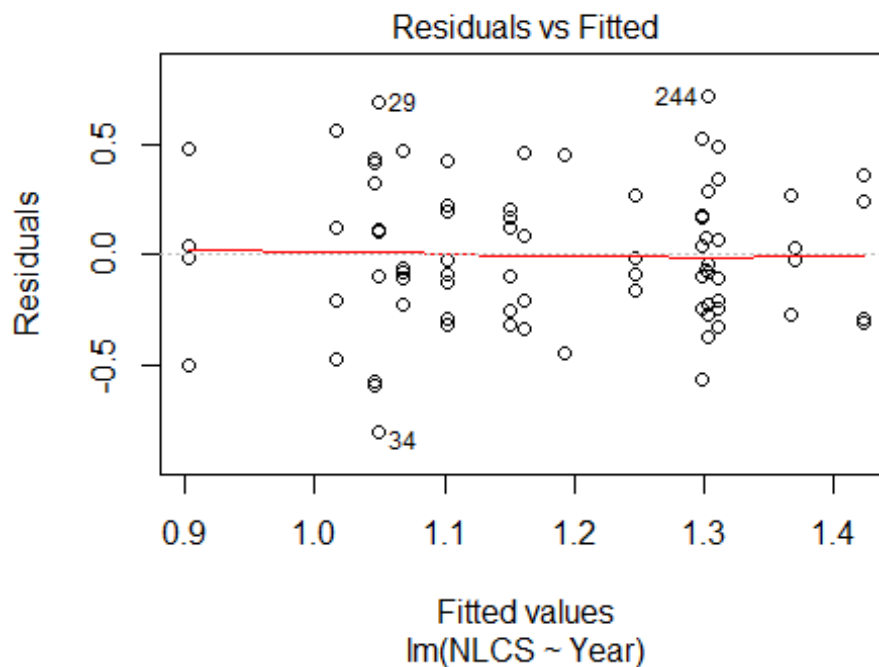
```

## Year2007      0.001955    0.135216    0.01    0.99
## Year2008      0.113816    0.123320    0.92    0.36
## Year2009      0.093041    0.122798    0.76    0.45
## Year2010      0.043144    0.117728    0.37    0.71
## Year2011      0.148809    0.114461    1.30    0.19
## Year2012      0.013141    0.122603    0.11    0.91
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.334
## Multiple R-squared:  0.0512, Adjusted R-squared:  0.005
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 31 weights are ~= 1. The remaining 336 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.257  0.884  0.948  0.908  0.986  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      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: 367"
## [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
##    3    6   12    5    6   12   10    7    7    8   10    9   11    9   12
## 2011 2012
##    8   10
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    4    5    4    2    2    5    4    5    4    7    2    7    4    8
## 2011 2012

```



```
##      7      7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      1      4      4      4      2      2      2      4      5      3      5      2      6      4      7
## 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 = 14, df = 16, p-value = 0.6
```



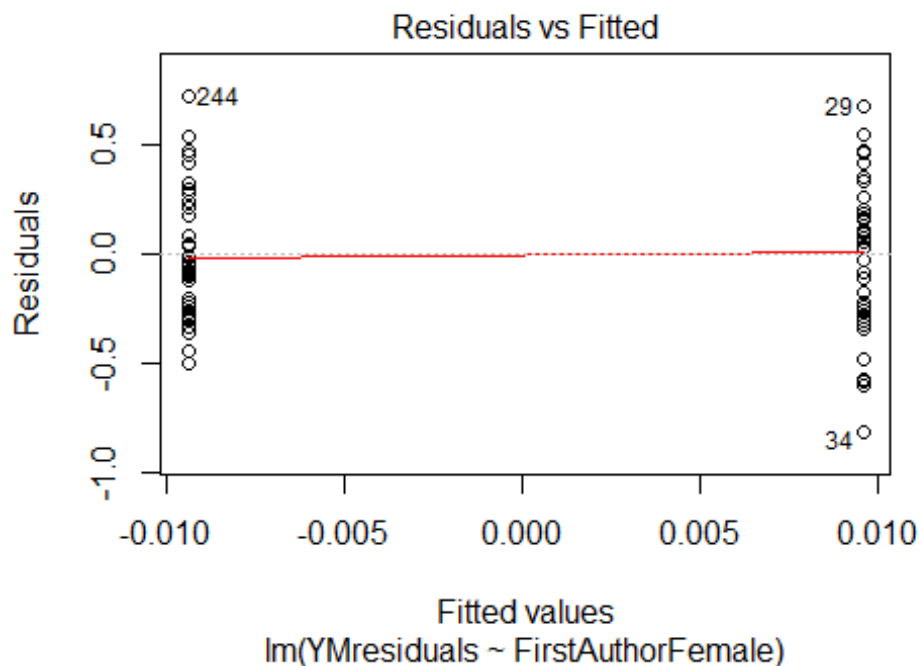
```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2, df = 1, p-value = 0.2

## [1] "Female first author team size 2018 geometric mean: 6.68740304976422"
## [1] "Male first author team size 2018 geometric mean: 5.03755003176877"

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

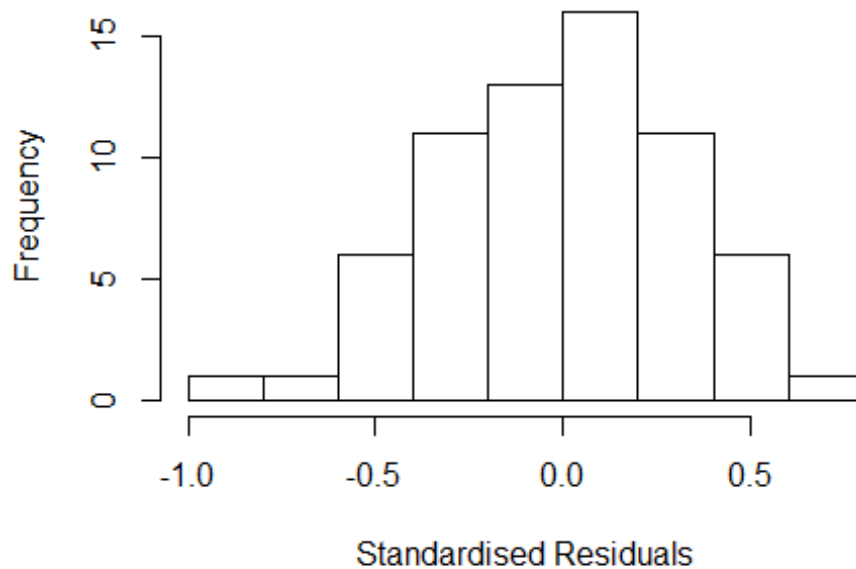
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 20, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6.40744299507361"
## [1] "Male last author team size 2018 geometric mean: 4.98779672490831"

## Warning in wilcox.test.default(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] "Regression 1: First author gender, last author gender, team size,
##      Year as factors"
##
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 9.228e+00  1      3.038
## LastAuthorFemale  1.105e+01  1      3.325
## UniqueAuthors    6.156e+04  4      3.969
## Year              3.255e+05 16      1.487
```

Residuals from first and last author and team size



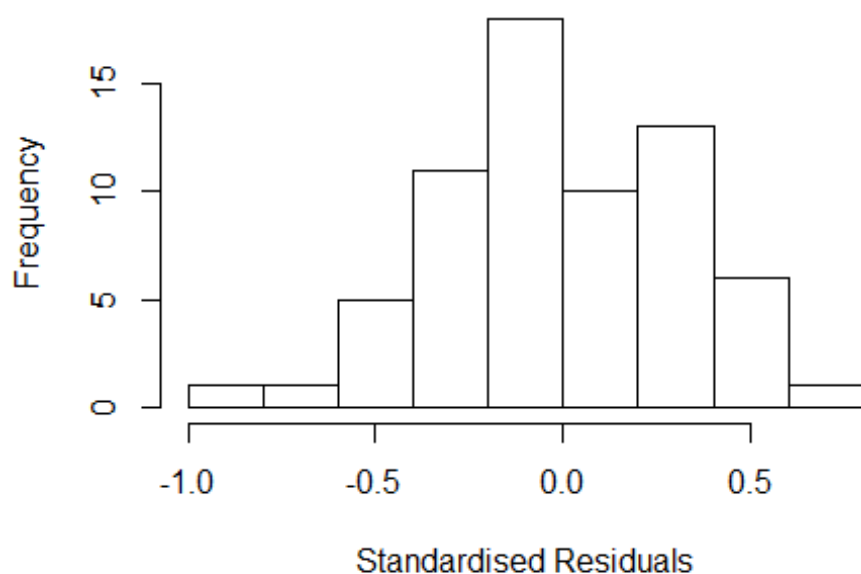
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.884 -0.213 0.025 0.231 0.746
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.33773 0.29329 4.56 4.2e-05 ***
## FirstAuthorFemale1 -0.04674 0.10687 -0.44 0.66
## LastAuthorFemale1 -0.11767 0.14486 -0.81 0.42
## UniqueAuthors2 -0.19500 0.30156 -0.65 0.52
## UniqueAuthors3 -0.14406 0.26299 -0.55 0.59
## UniqueAuthors4 0.14645 0.26685 0.55 0.59
## UniqueAuthors5 0.07012 0.31492 0.22 0.82
## Year1997 -0.45033 0.30736 -1.47 0.15
## Year1998 -0.20521 0.38154 -0.54 0.59
## Year1999 0.18495 0.16532 1.12 0.27
```

```

## Year2000      -0.00269    0.30989   -0.01    0.99
## Year2001      0.22357    0.21141    1.06    0.30
## Year2002     -0.00429    0.76270   -0.01    1.00
## Year2003     -0.29255    0.26688   -1.10    0.28
## Year2004     -0.18105    0.38219   -0.47    0.64
## Year2005     -0.05930    0.24109   -0.25    0.81
## Year2006     -0.19975    0.22786   -0.88    0.39
## Year2007     -0.10048    0.24251   -0.41    0.68
## Year2008      0.00693    0.24163    0.03    0.98
## Year2009     -0.09877    0.25945   -0.38    0.71
## Year2010     -0.24106    0.27639   -0.87    0.39
## Year2011      0.13649    0.22129    0.62    0.54
## Year2012      0.03608    0.27730    0.13    0.90
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.331
## Multiple R-squared:  0.272, Adjusted R-squared:  -0.1
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 58 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.455  0.871  0.949  0.912  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.52e-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 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 13.56 1      3.682
## LastAuthorFemale  16.26 1      4.032
## Year              46.19 16      1.127

```

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
## -0.9203 -0.2148 -0.0173 0.2365 0.6891
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.09834 0.10468 10.49 6.6e-14 ***
## FirstAuthorFemale1 -0.00234 0.10468 -0.02 0.9823
## LastAuthorFemale1 -0.10690 0.12313 -0.87 0.3897
## Year1997 -0.19305 0.22127 -0.87 0.3874
## Year1998 0.07096 0.33060 0.21 0.8310
## Year1999 0.35280 0.16214 2.18 0.0346 *
## Year2000 0.14995 0.34520 0.43 0.6660
## Year2001 0.27450 0.01946 14.11 < 2e-16 ***
## Year2002 0.20006 0.31792 0.63 0.5322
## Year2003 -0.09114 0.24147 -0.38 0.7076
## Year2004 0.01904 0.26464 0.07 0.9429
## Year2005 0.13512 0.21999 0.61 0.5420
```

```

## Year2006          0.06261    0.10758    0.58    0.5633
## Year2007          0.25778    0.07434    3.47    0.0011 **
## Year2008          0.17269    0.15188    1.14    0.2613
## Year2009          0.17601    0.08295    2.12    0.0392 *
## Year2010         -0.00547    0.09273   -0.06    0.9532
## Year2011          0.34765    0.15789    2.20    0.0326 *
## Year2012          0.23289    0.18448    1.26    0.2130
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.418
## Multiple R-squared:  0.156, Adjusted R-squared:  -0.167
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 55 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.607  0.917  0.949  0.933  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      1.52e-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 4.294 1      2.072
## Year              4.294 16      1.047
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max

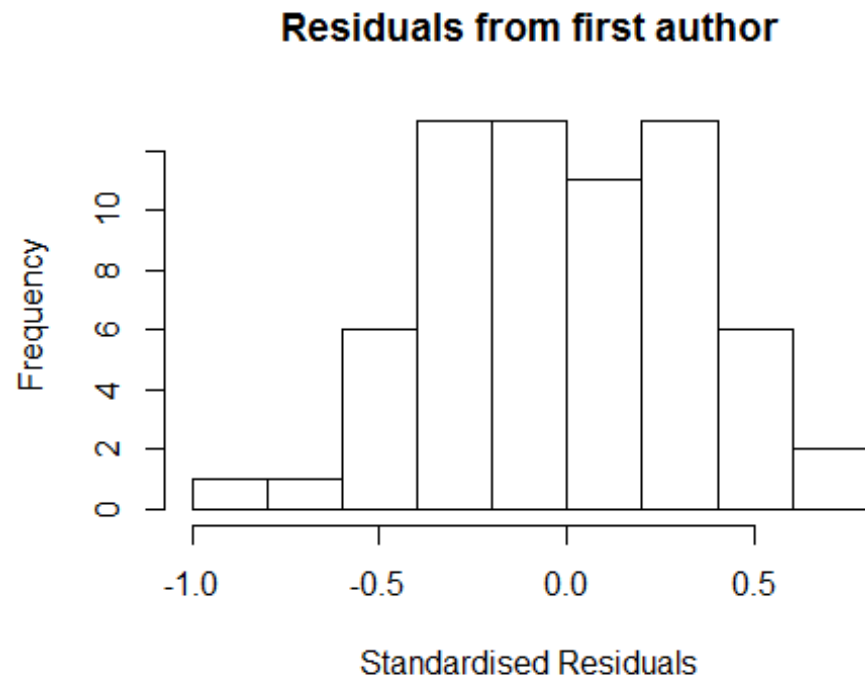
```

```

## -0.87581 -0.24724 -0.00818 0.24259 0.72488
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0777      0.1023   10.54 4.5e-14 ***
## FirstAuthorFemale1 0.0183      0.1023    0.18 0.8586
## Year1997         -0.1763      0.2247   -0.78 0.4365
## Year1998          0.0471      0.3203    0.15 0.8836
## Year1999          0.3352      0.1721    1.95 0.0573 .
## Year2000          0.0965      0.4083    0.24 0.8142
## Year2001          0.2745      0.0195   14.11 < 2e-16 ***
## Year2002          0.1138      0.3001    0.38 0.7061
## Year2003         -0.0703      0.2398   -0.29 0.7707
## Year2004         -0.0103      0.2643   -0.04 0.9692
## Year2005          0.1433      0.2190    0.65 0.5162
## Year2006          0.0496      0.1183    0.42 0.6772
## Year2007          0.2147      0.0695    3.09 0.0034 **
## Year2008          0.1829      0.1493    1.22 0.2268
## Year2009          0.1542      0.0853    1.81 0.0768 .
## Year2010         -0.0236      0.0950   -0.25 0.8049
## Year2011          0.3427      0.1498    2.29 0.0266 *
## Year2012          0.1971      0.1677    1.18 0.2456
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.421
## Multiple R-squared: 0.145, Adjusted R-squared: -0.158
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 54 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.645 0.902 0.957 0.932 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      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"

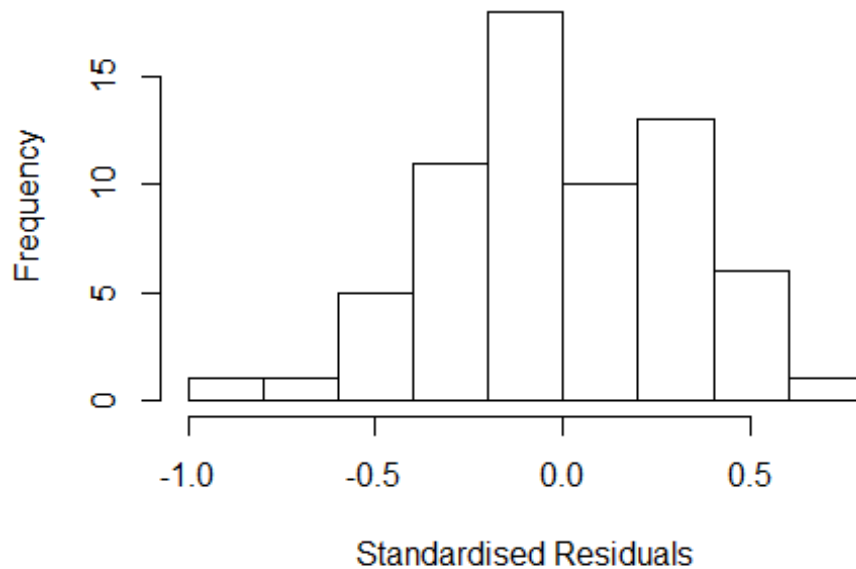
```

```
## 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 | 16.86 | 1 | 4.106 |
| ## | Year | 16.86 | 16 | 1.092 |

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
## -0.9203 -0.2137 -0.0171 0.2359 0.6889
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.09600 0.00000 Inf < 2e-16 ***
## LastAuthorFemale1 -0.10636 0.12115 -0.88 0.38435
## Year1997 -0.19115 0.18730 -1.02 0.31258
## Year1998 0.07327 0.32847 0.22 0.82443
## Year1999 0.35383 0.15660 2.26 0.02844 *
## Year2000 0.14968 0.34398 0.44 0.66541
## Year2001 0.27450 0.01946 14.11 < 2e-16 ***
## Year2002 0.20186 0.30703 0.66 0.51402
## Year2003 -0.08883 0.21622 -0.41 0.68304
## Year2004 0.02006 0.27398 0.07 0.94193
## Year2005 0.13605 0.20697 0.66 0.51411
## Year2006 0.06348 0.09794 0.65 0.51999
```

```

## Year2007      0.25868      0.06260      4.13  0.00014 ***
## Year2008      0.17383      0.14550      1.19  0.23808
## Year2009      0.17647      0.07952      2.22  0.03124 *
## Year2010     -0.00474      0.09615     -0.05  0.96089
## Year2011      0.34939      0.13893      2.51  0.01531 *
## Year2012      0.23306      0.18390      1.27  0.21116
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.418
## Multiple R-squared:  0.156, Adjusted R-squared:  -0.143
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 55 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.606  0.917  0.949  0.933  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      1.52e-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 2406"
## [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    2    2    6    3    1    4    3   12    7   14   23   17   10   16
## 2012
##    18
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    1    1    1    2    1    1    1    7    4    8   14   14    6    9
## 2012

```

```
## 12
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
## 1 0 1 1 2 1 1 1 6 4 7 13 13 5 7
## 2012
## 11
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 7.07106781186548"
## [1] "Male first author team size 2018 geometric mean: 8"
##
## 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: 7.36806299728077"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
## GVI F Df GVI F^(1/(2*Df))
## FirstAuthorFemale 1.375e+15 1 3.708e+07
## LastAuthorFemale 1.332e+00 1 1.154e+00
## UniqueAuthors -8.502e+31 4 NaN
## Year -2.131e+32 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 +
UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -5.20e-01 -1.38e-01 -2.48e-16 1.70e-01 6.70e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.22e+00 1.08e-01 1.13e+01 8.6e-16 ***
## FirstAuthorFemale1 4.15e-02 9.05e-02 4.60e-01 0.6484
## LastAuthorFemale1 1.49e-01 1.89e-01 7.90e-01 0.4338
## UniqueAuthors2 -4.28e-01 1.31e-01 -3.26e+00 0.0019 **
## UniqueAuthors3 -3.47e-01 1.54e-01 -2.25e+00 0.0287 *
## UniqueAuthors4 -1.05e-01 1.29e-01 -8.20e-01 0.4171
```

```

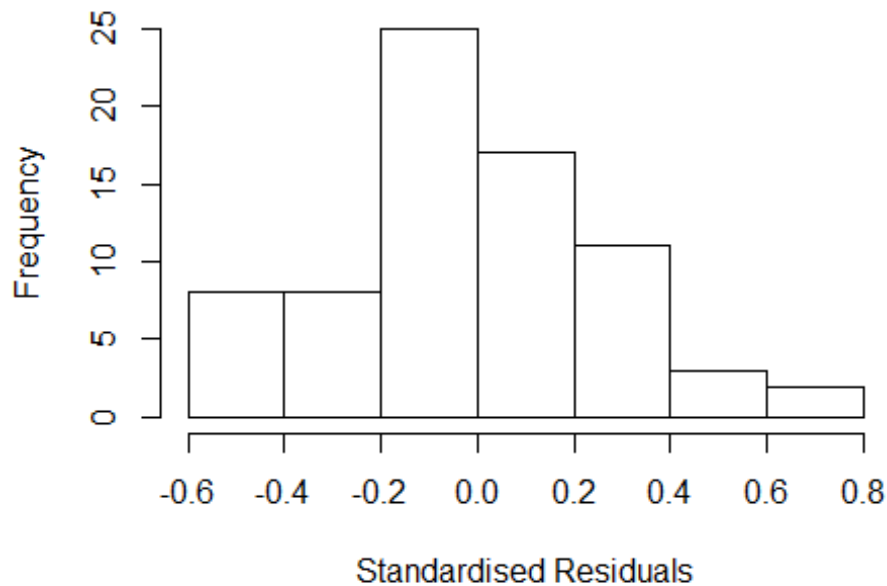
## UniqueAuthors5      -2.69e-01   1.22e-01 -2.20e+00   0.0322 *
## Year1999             2.32e-02   9.69e-02  2.40e-01   0.8115
## Year2000             1.30e-01   3.63e-08  3.58e+06 < 2e-16 ***
## Year2001            -1.48e-01   3.17e-01 -4.70e-01   0.6437
## Year2002             1.33e-01   9.69e-02  1.38e+00   0.1748
## Year2003            -1.43e-01   3.27e-08 -4.37e+06 < 2e-16 ***
## Year2004             1.40e-01   1.46e-01  9.60e-01   0.3425
## Year2005            -3.39e-01   1.80e-01 -1.88e+00   0.0654 .
## Year2006             1.52e-01   9.13e-02  1.66e+00   0.1028
## Year2007             1.04e-01   1.52e-01  6.80e-01   0.4977
## Year2008             1.62e-01   1.11e-01  1.46e+00   0.1488
## Year2009             1.45e-01   8.71e-02  1.66e+00   0.1029
## Year2010             1.25e-02   1.74e-01  7.00e-02   0.9429
## Year2011             4.49e-02   1.68e-01  2.70e-01   0.7907
## Year2012            -9.00e-02   9.45e-02 -9.50e-01   0.3454
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.335
## Multiple R-squared:  0.266, Adjusted R-squared:  -0.0112
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 60 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.668  0.898  0.960  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      1.35e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##   nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##   trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          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

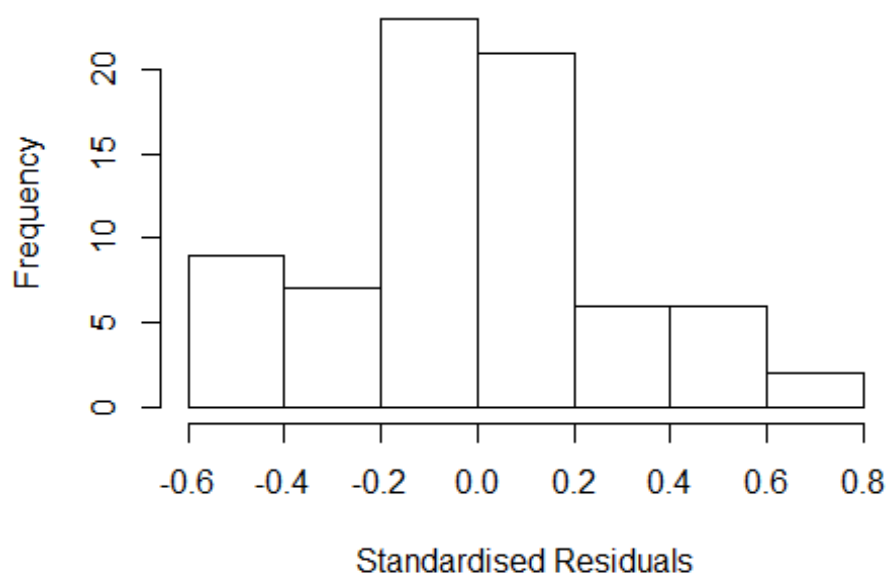
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q      Median        3Q       Max
## -5.27e-01 -1.48e-01  7.13e-16  1.69e-01  7.14e-01
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    9.71e-01   9.91e-02  9.80e+00  7.9e-14 ***
## FirstAuthorFemale1 2.42e-02   9.91e-02  2.40e-01   0.8079
## LastAuthorFemale1 1.55e-01   1.70e-01  9.20e-01   0.3633
## Year1999        -1.36e-01   0.00e+00   -Inf < 2e-16 ***
## Year2000         1.30e-01   0.00e+00    Inf < 2e-16 ***
## Year2001        -7.44e-02   5.10e-01 -1.50e-01   0.8844
## Year2002        -2.60e-02   0.00e+00   -Inf < 2e-16 ***
```

```

## Year2003      -1.43e-01   1.11e-08 -1.29e+07 < 2e-16 ***
## Year2004      4.42e-02   9.91e-02  4.50e-01  0.6572
## Year2005     -3.56e-01   2.01e-01 -1.77e+00  0.0818 .
## Year2006      2.60e-01   8.14e-02  3.19e+00  0.0023 **
## Year2007      8.39e-02   1.29e-01  6.50e-01  0.5184
## Year2008      1.42e-01   1.17e-01  1.22e+00  0.2288
## Year2009      8.25e-02   9.41e-02  8.80e-01  0.3841
## Year2010     -9.19e-02   1.49e-01 -6.20e-01  0.5393
## Year2011      6.12e-02   1.42e-01  4.30e-01  0.6679
## Year2012     -4.36e-02   8.58e-02 -5.10e-01  0.6129
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.314
## Multiple R-squared:  0.192, Adjusted R-squared:  -0.0348
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~ 1. The remaining 61 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.585  0.843   0.966   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      1.35e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          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

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.55476 -0.18289  0.00161  0.15391  0.70565
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    9.60e-01   9.13e-02  1.05e+01  4.7e-15 ***
## FirstAuthorFemale1 3.52e-02   9.13e-02  3.90e-01   0.701
## Year1999        -1.36e-01   9.25e-09 -1.47e+07 < 2e-16 ***
## Year2000         1.30e-01   1.24e-08  1.05e+07 < 2e-16 ***
## Year2001        -6.89e-02   5.07e-01 -1.40e-01   0.892
## Year2002        -2.60e-02   2.34e-08 -1.11e+06 < 2e-16 ***
## Year2003        -1.43e-01   0.00e+00    -Inf < 2e-16 ***
## Year2004         5.52e-02   9.13e-02  6.00e-01   0.548
```

```

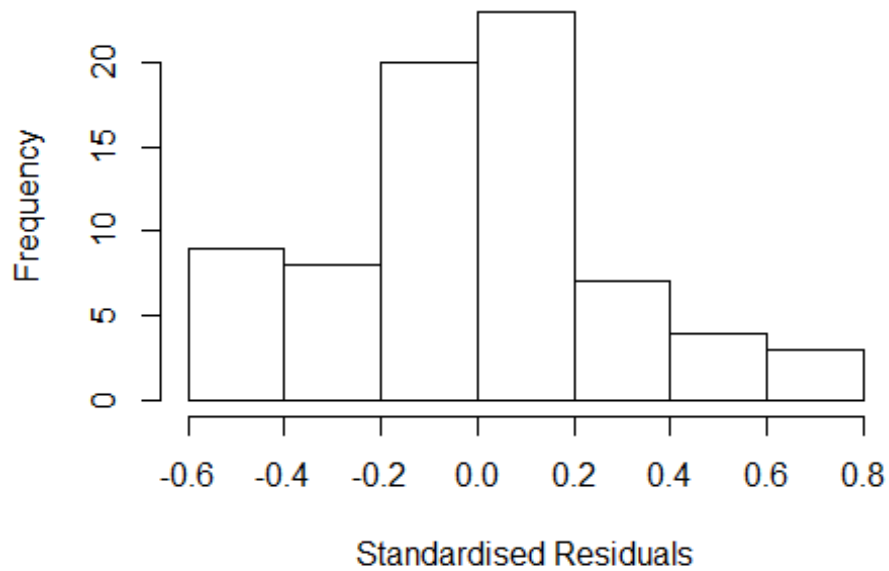
## Year2005          -3.36e-01   1.78e-01 -1.89e+00   0.063 .
## Year2006          2.63e-01   8.12e-02  3.24e+00   0.002 **
## Year2007          1.11e-01   1.35e-01  8.20e-01   0.414
## Year2008          1.49e-01   1.15e-01  1.30e+00   0.200
## Year2009          8.88e-02   9.08e-02  9.80e-01   0.332
## Year2010         -5.88e-02   1.28e-01 -4.60e-01   0.648
## Year2011          1.34e-01   1.22e-01  1.10e+00   0.277
## Year2012         -1.82e-02   9.08e-02 -2.00e-01   0.842
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.308
## Multiple R-squared:  0.183, Adjusted R-squared:  -0.0286
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 62 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.580  0.851  0.960  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      1.35e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          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

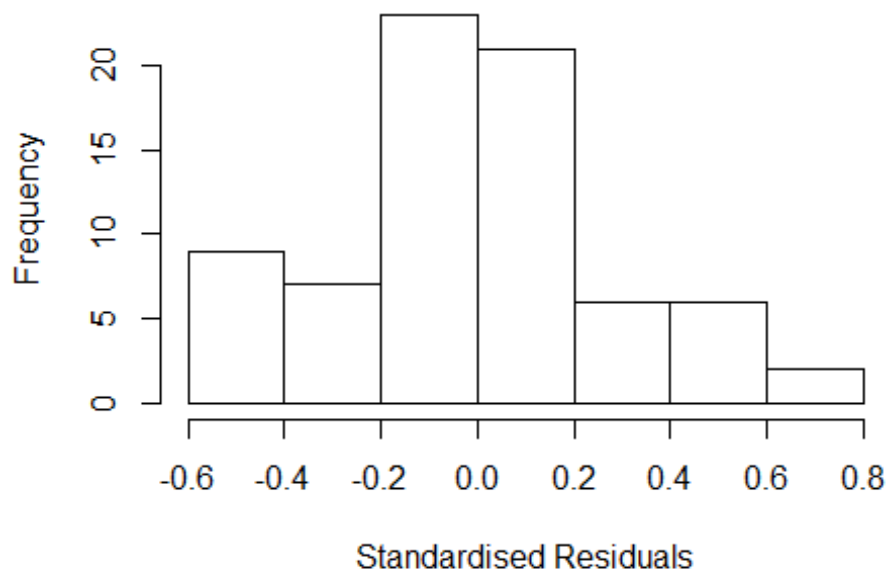
```


Residuals from first author



| | | | | |
|----|------------------|------|----|-------------------------|
| ## | | GVIF | Df | $GVIF^{1/(2 \cdot Df)}$ |
| ## | LastAuthorFemale | NaN | 1 | NaN |
| ## | Year | NaN | 14 | NaN |

Residuals from last author

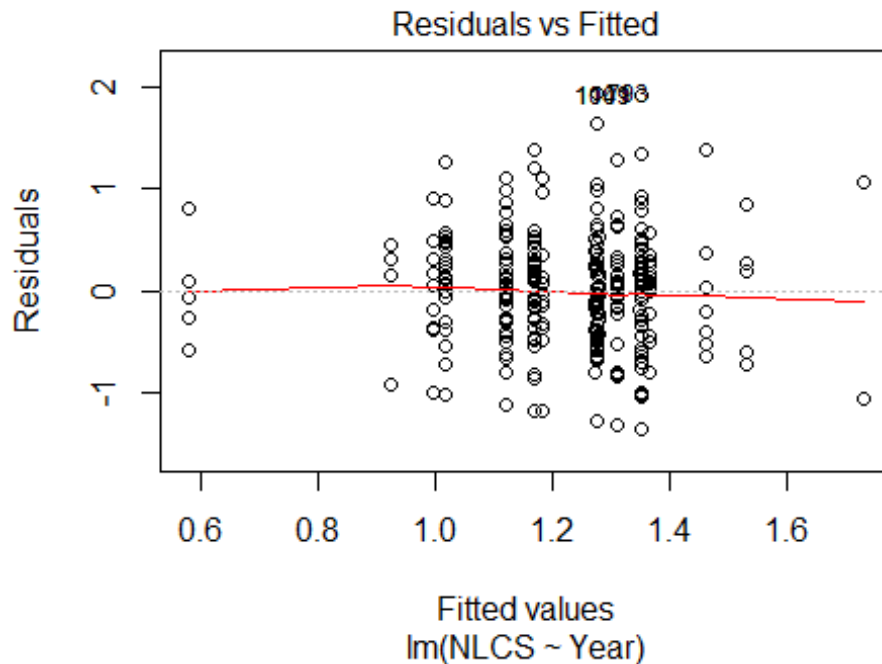


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -5.57e-01 -1.38e-01 -4.01e-16 1.58e-01 7.04e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 9.95e-01 1.60e-08 6.23e+07 <2e-16 ***
## LastAuthorFemale1 1.61e-01 1.65e-01 9.80e-01 0.3336
## Year1999 -1.36e-01 0.00e+00 -Inf <2e-16 ***
## Year2000 1.30e-01 1.42e-08 9.18e+06 <2e-16 ***
## Year2001 -8.65e-02 5.10e-01 -1.70e-01 0.8660
## Year2002 -2.60e-02 1.84e-08 -1.42e+06 <2e-16 ***
## Year2003 -1.43e-01 0.00e+00 -Inf <2e-16 ***
## Year2004 2.00e-02 0.00e+00 Inf <2e-16 ***
## Year2005 -3.56e-01 2.02e-01 -1.76e+00 0.0837 .
## Year2006 2.53e-01 7.50e-02 3.38e+00 0.0013 **
## Year2007 7.54e-02 1.27e-01 5.90e-01 0.5546
## Year2008 1.28e-01 9.55e-02 1.34e+00 0.1864
## Year2009 6.85e-02 7.99e-02 8.60e-01 0.3949
## Year2010 -1.08e-01 1.24e-01 -8.70e-01 0.3894
## Year2011 4.25e-02 1.25e-01 3.40e-01 0.7344
## Year2012 -5.22e-02 8.21e-02 -6.40e-01 0.5275
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.328
## Multiple R-squared: 0.187, Adjusted R-squared: -0.0233
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 59 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.624 0.853 0.955 0.916 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 1.35e-03 1.82e-12
## warn.limit.reject warn.limit.meanrw
## 5.00e-01 5.00e-01
## nResample max.it best.r.s k.fast.s k.max maxit.scale
```

```

##           500           50           2           1           1000           200
## trace.lev      mts compute.rd
##           0           1000           0
##           psi           subsampling           cov
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 74"
## [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
##   11   17   20   33   28   24   24   35   34   39   43   70   47   69   73
## 2011 2012
##   86   76
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    4    2    7    5    5    9   10   10   15   12   28   24   36   41
## 2011 2012
##   50   50
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    4    2    7    5    5    9    8    9   10   11   25   19   31   35
## 2011 2012
##   43   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 = 24, df = 16, p-value = 0.1

```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.8, df = 1, p-value = 0.1

## [1] "Female first author team size 2018 geometric mean: 3.55181745772802"
## [1] "Male first author team size 2018 geometric mean: 3.44320910574379"

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

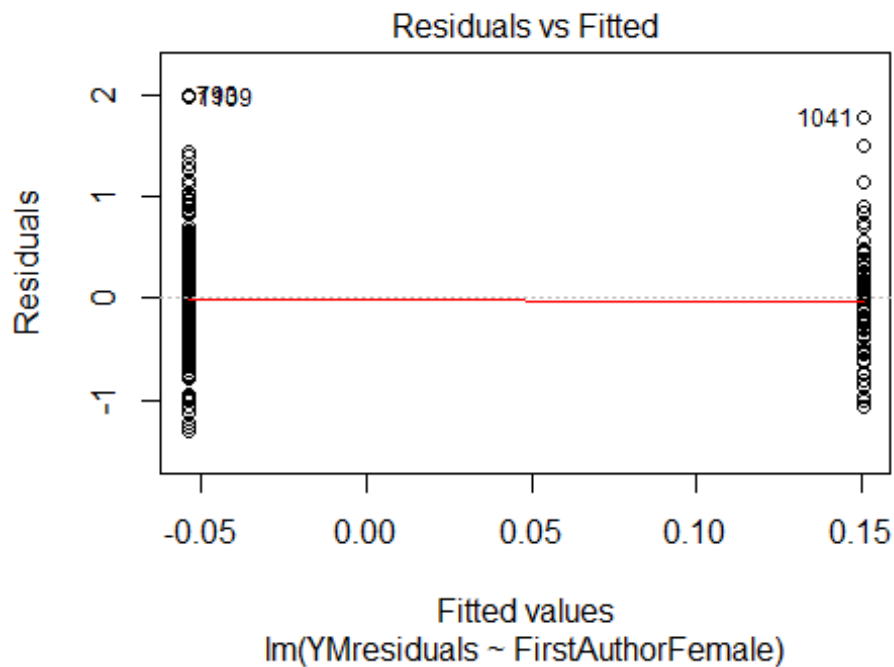
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 240, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.13887172043383"
## [1] "Male last author team size 2018 geometric mean: 3.33975150247096"

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

##
## Wilcoxon rank sum test with continuity correction
##
```

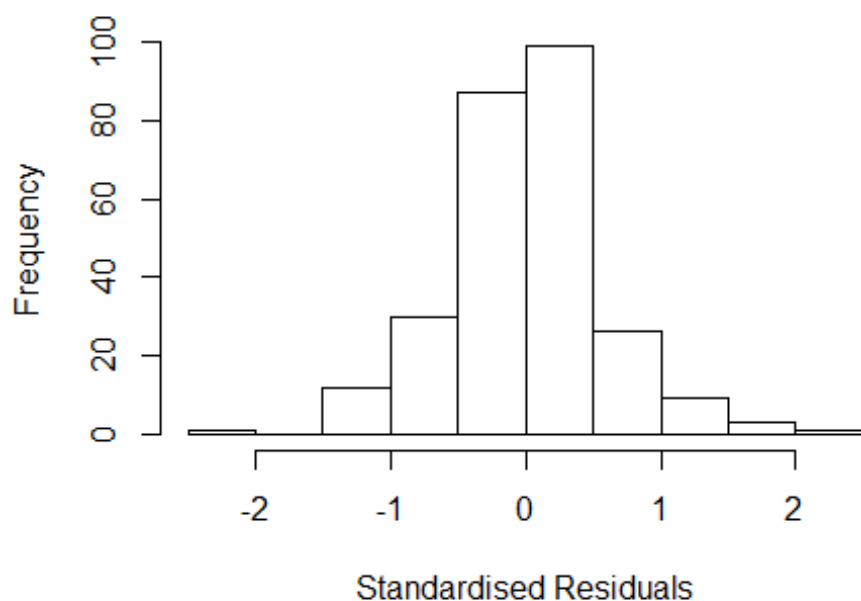
```
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 210, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"

## 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 1.267 1          1.125
## LastAuthorFemale  1.241 1          1.114
## Year              1.524 16          1.013
```

Residuals from first and last author



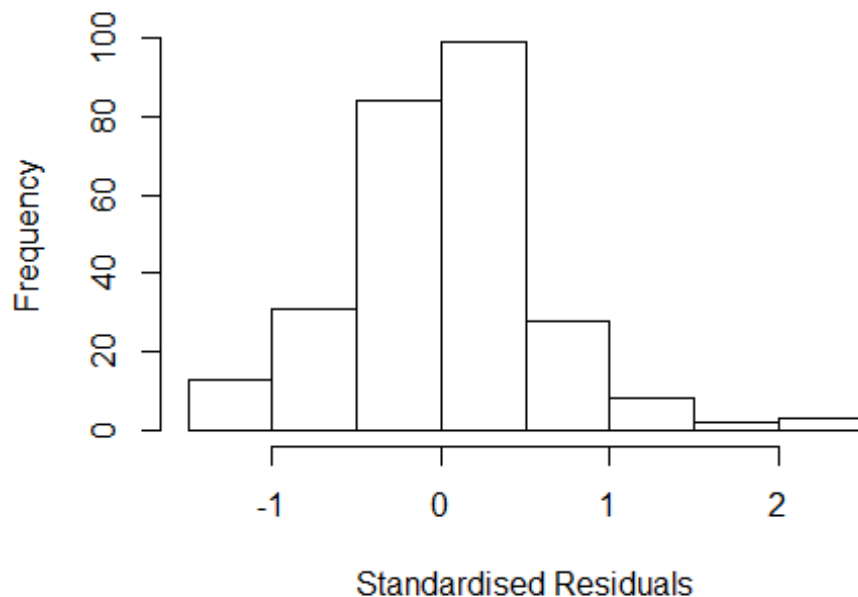
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -2.07456 -0.32444 0.00524 0.31378 2.00383
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3485 0.1279 10.54 < 2e-16 ***
## FirstAuthorFemale1 0.2504 0.0699 3.58 0.00041 ***
## LastAuthorFemale1 -0.2146 0.1064 -2.02 0.04474 *
## Year1997 -0.4444 0.3261 -1.36 0.17413
## Year1998 1.3931 0.2811 4.96 1.3e-06 ***
## Year1999 -0.0847 0.2163 -0.39 0.69557
## Year2000 0.1684 0.3646 0.46 0.64452
## Year2001 -0.8633 0.2579 -3.35 0.00094 ***
## Year2002 -0.4122 0.2396 -1.72 0.08668 .
## Year2003 -0.1377 0.1603 -0.86 0.39113
## Year2004 -0.0714 0.1622 -0.44 0.66008
## Year2005 -0.1142 0.1865 -0.61 0.54072
```

```

## Year2006          -0.3099      0.2331   -1.33   0.18499
## Year2007          -0.3459      0.1756   -1.97   0.04998 *
## Year2008          -0.0994      0.1571   -0.63   0.52749
## Year2009          -0.2288      0.1580   -1.45   0.14871
## Year2010          -0.1046      0.1647   -0.64   0.52587
## Year2011          -0.2808      0.1602   -1.75   0.08097 .
## Year2012          -0.0753      0.1661   -0.45   0.65050
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.484
## Multiple R-squared:  0.152, Adjusted R-squared:  0.0908
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 242 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0262 0.8490 0.9550 0.8770 0.9860 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      3.73e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## 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
## -1.2534 -0.3134 0.0291 0.3040 2.0822
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3485 0.1280 10.54 < 2e-16 ***
## FirstAuthorFemale1 0.2453 0.0705 3.48 0.00060 ***
## Year1997 -0.5076 0.3375 -1.50 0.13381
## Year1998 -0.6347 0.2502 -2.54 0.01178 *
## Year1999 -0.0849 0.2165 -0.39 0.69532
## Year2000 0.1055 0.3405 0.31 0.75694
## Year2001 -0.8623 0.2577 -3.35 0.00094 ***
## Year2002 -0.4108 0.2395 -1.72 0.08747 .
## Year2003 -0.1602 0.1536 -1.04 0.29803
## Year2004 -0.1203 0.1559 -0.77 0.44118
## Year2005 -0.1131 0.1864 -0.61 0.54471
## Year2006 -0.3392 0.2327 -1.46 0.14612
```



```

## Year2007          -0.3689      0.1753   -2.10  0.03630 *
## Year2008          -0.1290      0.1595   -0.81  0.41937
## Year2009          -0.2368      0.1582   -1.50  0.13560
## Year2010          -0.1551      0.1628   -0.95  0.34139
## Year2011          -0.3196      0.1621   -1.97  0.04982 *
## Year2012          -0.0951      0.1683   -0.57  0.57253
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.482
## Multiple R-squared:  0.11,   Adjusted R-squared:  0.0494
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 27 weights are ~= 1. The remaining 241 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0224 0.8330 0.9570 0.8750 0.9860 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      3.73e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.275 1          1.129
## Year              1.275 16          1.008

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.3306 -0.3170  0.0271  0.3041  2.1541

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.34850    0.12786   10.55 <2e-16 ***
## LastAuthorFemale1 -0.20974    0.11906   -1.76  0.0793 .
## Year1997       -0.27821    0.29891   -0.93  0.3529
## Year1998       -0.59568    0.45927   -1.30  0.1958
## Year1999       -0.03774    0.24899   -0.15  0.8796
## Year2000        0.23688    0.39545    0.60  0.5497
## Year2001       -0.79892    0.24954   -3.20  0.0015 **
## Year2002       -0.34322    0.22518   -1.52  0.1287
## Year2003       -0.01470    0.16757   -0.09  0.9302
## Year2004       -0.04244    0.16021   -0.26  0.7913
## Year2005       -0.05666    0.18088   -0.31  0.7543
## Year2006       -0.20957    0.21710   -0.97  0.3353
## Year2007       -0.30378    0.17766   -1.71  0.0885 .
## Year2008        0.00495    0.15457    0.03  0.9745
## Year2009       -0.14857    0.15601   -0.95  0.3419
## Year2010       -0.07674    0.16887   -0.45  0.6499
## Year2011       -0.17563    0.15756   -1.11  0.2661
## Year2012       -0.01787    0.16488   -0.11  0.9138
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.489
## Multiple R-squared:  0.088, Adjusted R-squared:  0.0259
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 242 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0134 0.8450 0.9500 0.8760 0.9870 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      3.73e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 268"
## [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]"
##
## 1999 2003 2005 2006 2007 2008 2009 2010
##    1    1    2    1    3    4    1    1
##
## 1999 2003 2005 2006 2007 2008 2009 2010
##    1    0    2    0    1    3    1    1
##
## 1999 2003 2005 2006 2007 2008 2009 2010
##    1    0    2    0    1    3    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: 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: 9"
## [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
##    4    7    4    1    8    5    6    4    5    5    9    13    20    41    42
## 2011 2012
##   32   52
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    1    1    2    2    1    0    2    2    3    5    14    27    30
## 2011 2012
##   20   31
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    1    1    1    2    2    1    0    2    2    3    4    11    21    22
## 2011 2012
##   16   25
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"

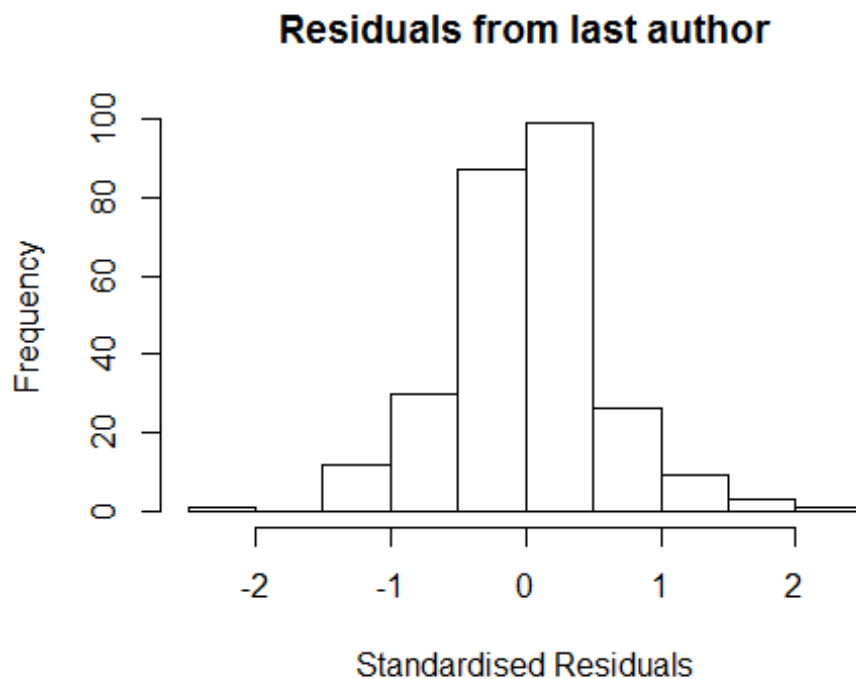
```

```
## [1] "Female first author team size 2018 geometric mean: 4.55805095233363"
## [1] "Male first author team size 2018 geometric mean: 3.45988211384833"

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

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 71, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6.32455532033676"
## [1] "Male last author team size 2018 geometric mean: 3.75400681762488"

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



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 31, 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 2.493e+13  1      4.993e+06
## LastAuthorFemale  1.018e+00  1      1.009e+00
## UniqueAuthors    1.130e+41  4      1.354e+05
## Year              3.157e+41 14      3.035e+01

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q   Median      3Q      Max
## -1.31e+00 -2.38e-01  2.72e-15  2.33e-01  1.25e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -0.6492    0.1559   -4.16  7.0e-05 ***
## FirstAuthorFemale1 -0.3182    0.1006   -3.16  0.0021 **
## LastAuthorFemale1  0.1245    0.1004    1.24  0.2179
## UniqueAuthors2    0.9902    0.1668    5.94  5.0e-08 ***
## UniqueAuthors3    1.3212    0.1559    8.48  3.4e-13 ***
## UniqueAuthors4    1.3134    0.1718    7.64  1.9e-11 ***
## UniqueAuthors5    1.3367    0.1503    8.89  4.5e-14 ***
## Year1998          0.7915    0.1098    7.21  1.5e-10 ***
## Year1999          0.6410    0.1306    4.91  3.9e-06 ***
## Year2000          0.6016    0.0949    6.34  8.1e-09 ***
## Year2001          0.2659    0.2795    0.95  0.3440
## Year2002          1.8752    0.1559   12.03 < 2e-16 ***
## Year2004          0.3261    0.2511    1.30  0.1973
## Year2005          0.5856    0.2732    2.14  0.0347 *
## Year2006          0.6019    0.1603    3.76  0.0003 ***
## Year2007          0.6492    0.1559    4.16  7.0e-05 ***
## Year2008          0.8409    0.1807    4.65  1.1e-05 ***
## Year2009          0.6080    0.1091    5.57  2.4e-07 ***
## Year2010          0.6424    0.1264    5.08  1.9e-06 ***
## Year2011          0.6837    0.1206    5.67  1.6e-07 ***
## Year2012          0.6125    0.1206    5.08  1.9e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.408
## Multiple R-squared:  0.265, Adjusted R-squared:  0.107

```

```

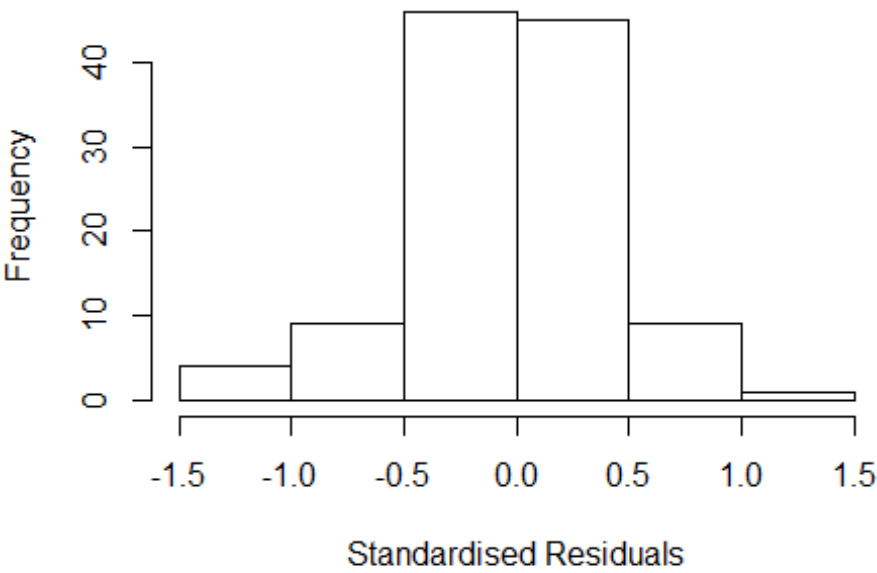
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 97 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.283 0.875 0.954 0.895 0.981 0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      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      8.77e-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"

## 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

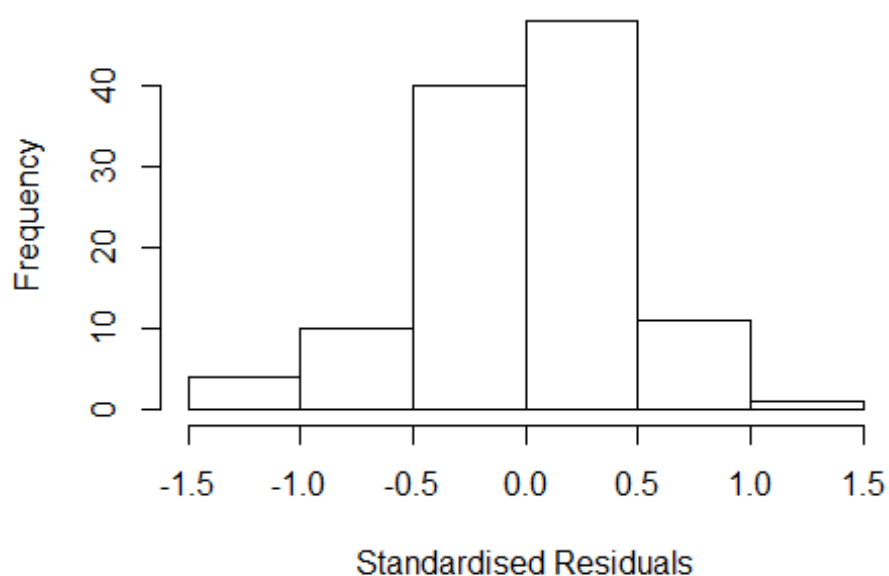
```

Residuals from first and last author and team size



| ## | | GVIF | Df | $GVIF^{1/(2 \cdot Df)}$ |
|----|-------------------|------|----|-------------------------|
| ## | FirstAuthorFemale | NaN | 1 | NaN |
| ## | LastAuthorFemale | NaN | 1 | NaN |
| ## | Year | NaN | 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
## -1.1988 -0.2460  0.0262  0.2419  1.3792
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.6720    0.0000      Inf < 2e-16 ***
## FirstAuthorFemale1 -0.2483    0.1014   -2.45  0.01618 *
## LastAuthorFemale1  0.1324    0.1379    0.96  0.33953
## Year1998         0.8070    0.0000      Inf < 2e-16 ***
## Year1999         0.5633    0.1014    5.55  2.4e-07 ***
## Year2000         0.4011    0.1609    2.49  0.01436 *
## Year2001         0.0965    0.1186    0.81  0.41791
## Year2002         0.5540    0.0000      Inf < 2e-16 ***
## Year2004         0.3066    0.2619    1.17  0.24448
## Year2005         0.5506    0.2347    2.35  0.02100 *
## Year2006         0.5583    0.1456    3.83  0.00022 ***
## Year2007         0.4200    0.3336    1.26  0.21104
```

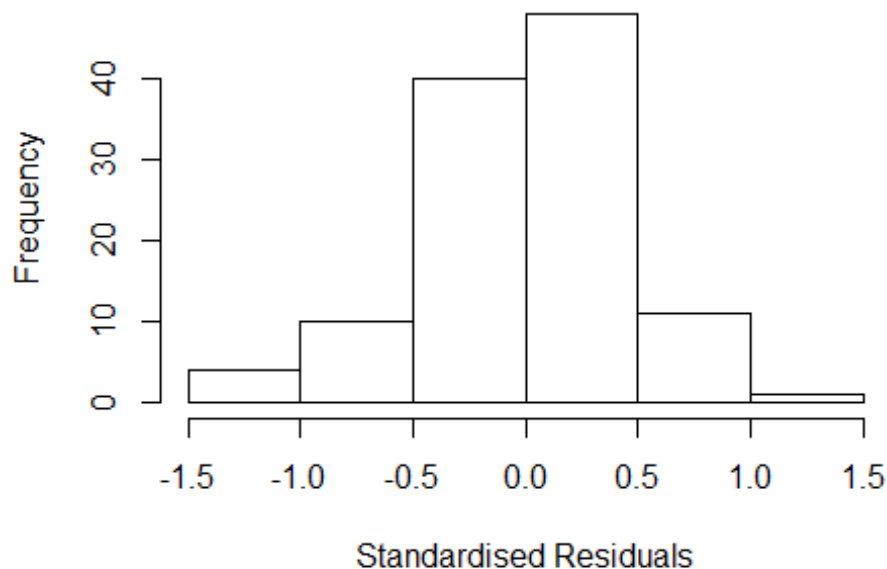


```

## Year2008          0.7780      0.1426      5.46  3.7e-07 ***
## Year2009          0.5676      0.0787      7.21  1.2e-10 ***
## Year2010          0.5268      0.1255      4.20  6.0e-05 ***
## Year2011          0.6386      0.1374      4.65  1.1e-05 ***
## Year2012          0.5238      0.0942      5.56  2.4e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.44
## Multiple R-squared:  0.133, Adjusted R-squared:  -0.00982
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 102 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.305  0.881  0.964  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      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 4.228e+14 1      2.056e+07
## Year              4.228e+14 14      3.329e+00

```

Residuals from first author



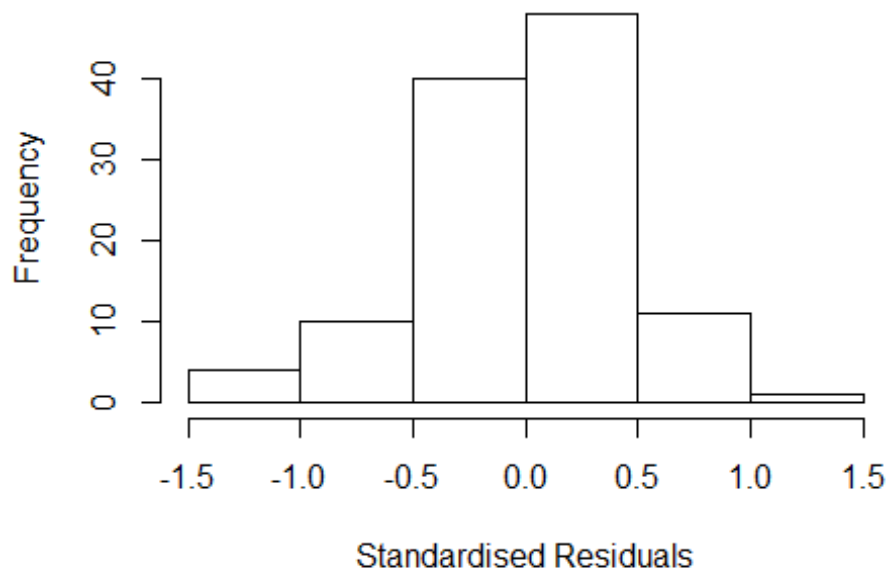
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2159 -0.2557 0.0255 0.2396 1.3621
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 6.72e-01 5.86e-08 1.15e+07 < 2e-16 ***
## FirstAuthorFemale1 -2.42e-01 1.02e-01 -2.38e+00 0.01909 *
## Year1998 8.07e-01 5.77e-08 1.40e+07 < 2e-16 ***
## Year1999 5.57e-01 1.02e-01 5.48e+00 3.3e-07 ***
## Year2000 3.98e-01 1.59e-01 2.51e+00 0.01374 *
## Year2001 9.65e-02 1.19e-01 8.10e-01 0.41795
## Year2002 5.54e-01 6.27e-08 8.83e+06 < 2e-16 ***
## Year2004 3.04e-01 2.66e-01 1.14e+00 0.25582
## Year2005 5.48e-01 2.32e-01 2.36e+00 0.02029 *
## Year2006 5.54e-01 1.45e-01 3.83e+00 0.00022 ***
## Year2007 4.18e-01 3.33e-01 1.25e+00 0.21267
## Year2008 7.76e-01 1.42e-01 5.45e+00 3.8e-07 ***
```

```

## Year2009          5.74e-01  7.89e-02  7.28e+00  8.6e-11 ***
## Year2010          5.44e-01  1.22e-01  4.45e+00  2.2e-05 ***
## Year2011          6.56e-01  1.31e-01  5.01e+00  2.4e-06 ***
## Year2012          5.28e-01  9.60e-02  5.50e+00  3.0e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.439
## Multiple R-squared:  0.129, Adjusted R-squared:  -0.00473
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 104 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.314  0.889  0.963  0.903  0.988  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      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 62.01  1          7.875
## Year            62.01 14          1.159

```

Residuals from last author



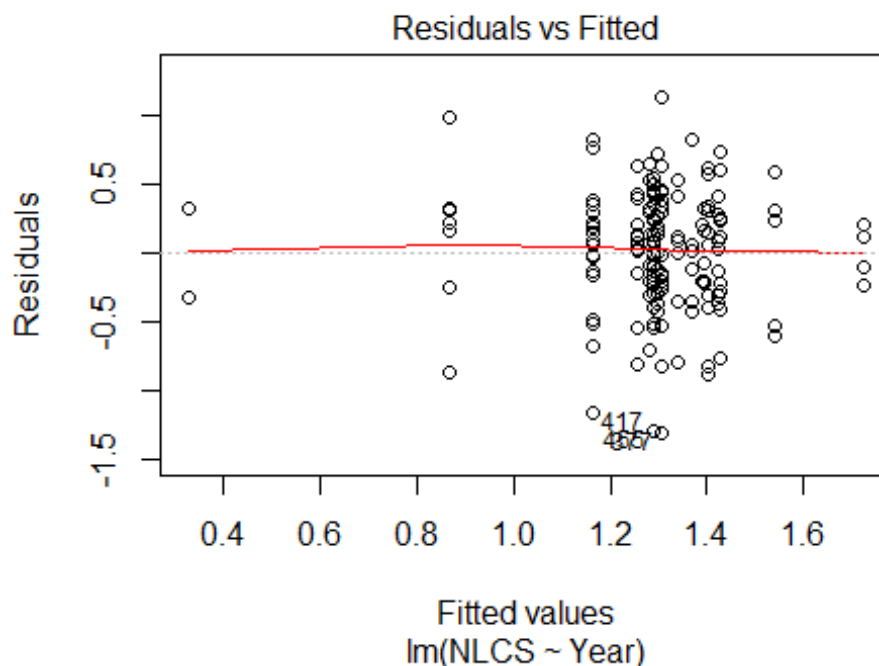
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1387 -0.2736 0.0178 0.2632 1.4393
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 6.72e-01 3.43e-08 1.96e+07 < 2e-16 ***
## LastAuthorFemale1 9.26e-02 1.44e-01 6.40e-01 0.52279
## Year1998 8.07e-01 2.28e-08 3.53e+07 < 2e-16 ***
## Year1999 3.15e-01 2.82e-08 1.12e+07 < 2e-16 ***
## Year2000 2.77e-01 5.91e-02 4.69e+00 8.9e-06 ***
## Year2001 9.65e-02 1.19e-01 8.10e-01 0.41773
## Year2002 5.54e-01 5.07e-08 1.09e+07 < 2e-16 ***
## Year2004 1.82e-01 3.97e-01 4.60e-01 0.64698
## Year2005 4.26e-01 1.25e-01 3.40e+00 0.00098 ***
## Year2006 3.96e-01 1.06e-01 3.75e+00 0.00030 ***
## Year2007 3.14e-01 3.33e-01 9.40e-01 0.34760
## Year2008 6.92e-01 1.26e-01 5.47e+00 3.4e-07 ***
```

```

## Year2009          5.40e-01    8.23e-02 6.57e+00  2.5e-09 ***
## Year2010          4.67e-01    1.36e-01 3.43e+00  0.00089 ***
## Year2011          5.88e-01    1.52e-01 3.88e+00  0.00019 ***
## Year2012          4.55e-01    9.86e-02 4.61e+00  1.2e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.443
## Multiple R-squared:  0.0813, Adjusted R-squared:  -0.0594
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 101 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.268  0.886  0.954  0.898  0.988  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      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 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
##   19   12   26   15   12   19   29   18   37   26   15   15   22   16   22
## 2011 2012
##   27   36
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    2    6    5    4   10    5    8   12   11    7    7   11   13   16
## 2011 2012
##   19   27
##

```

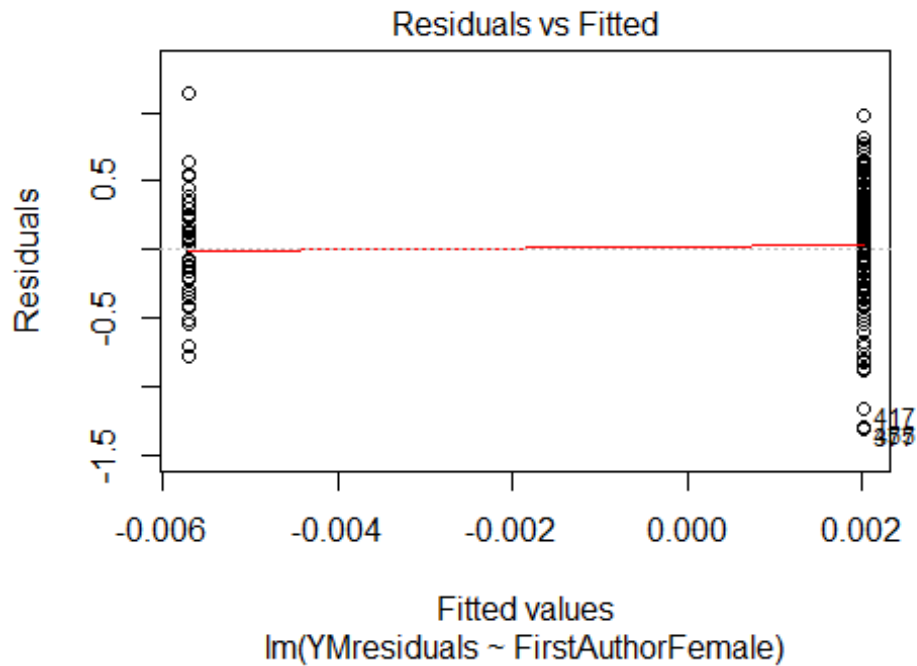
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    2    5    5    4    8    5    7    8    8    5    6    9    8    14
## 2011 2012
##   15   22
## [1] "Heteroscedasticity checks, confirming that there are problems with
##      these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 14, df = 16, p-value = 0.6
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.3, df = 1, p-value = 0.3

## [1] "Female first author team size 2018 geometric mean: 2.44948974278318"
## [1] "Male first author team size 2018 geometric mean: 3.72791927319135"

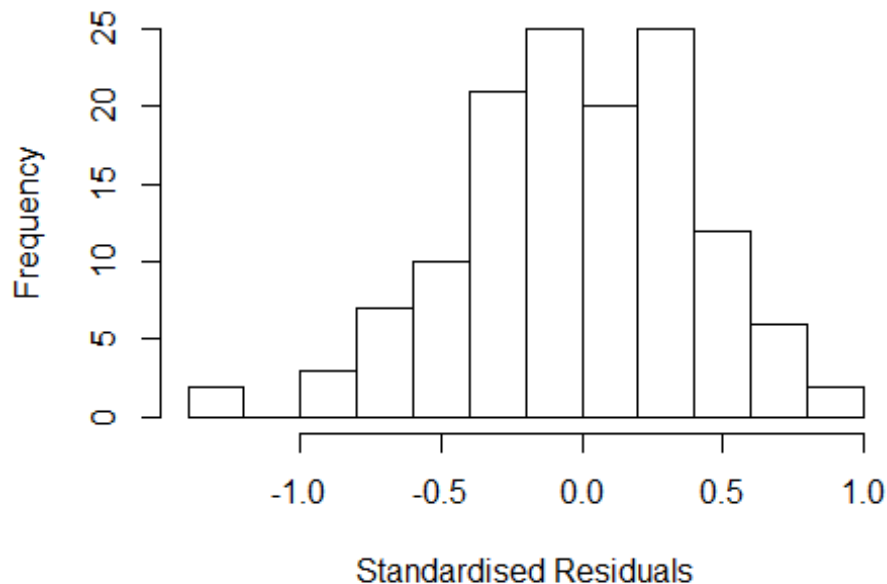
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1, p-value = 0.2
## 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: 3.30639565054252"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
```

| | GVIF | Df | GVIF^(1/(2*Df)) |
|-------------------|--------|----|-----------------|
| FirstAuthorFemale | 1.326 | 1 | 1.151 |
| LastAuthorFemale | 1.980 | 1 | 1.407 |
| UniqueAuthors | 7.368 | 4 | 1.284 |
| Year | 14.702 | 16 | 1.088 |

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
## -1.328 -0.253 -0.012 0.259 0.991
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.22686 0.33692 3.64 0.00041 ***
## FirstAuthorFemale1 -0.03314 0.07793 -0.43 0.67147
## LastAuthorFemale1 0.03310 0.12226 0.27 0.78712
## UniqueAuthors2 0.14609 0.28552 0.51 0.60992
## UniqueAuthors3 0.09630 0.28130 0.34 0.73274
## UniqueAuthors4 0.08137 0.28130 0.29 0.77293
## UniqueAuthors5 -0.00392 0.29893 -0.01 0.98957
## Year1997 -0.96354 0.32067 -3.00 0.00329 **
## Year1998 -0.13046 0.20393 -0.64 0.52368
## Year1999 0.05711 0.20912 0.27 0.78527
```

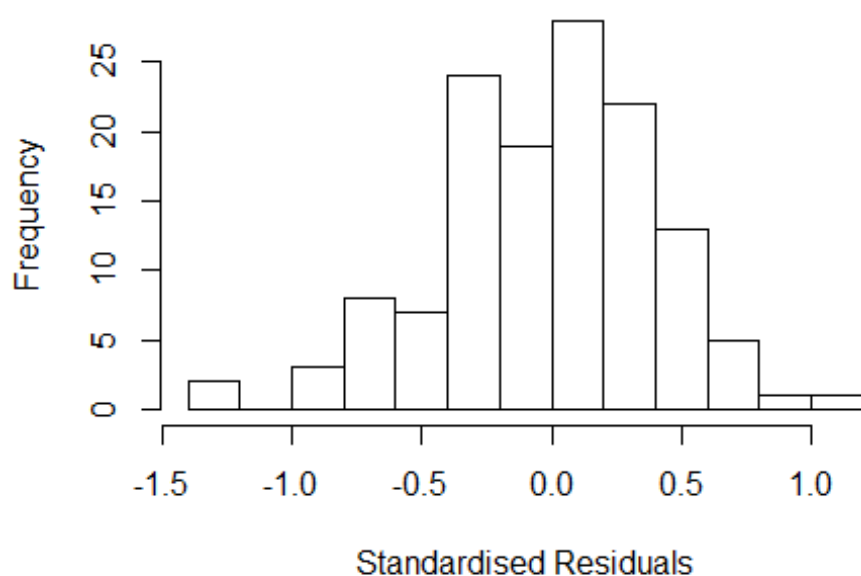


```

## Year2000          0.39652      0.21002      1.89  0.06165 .
## Year2001          0.10422      0.25028      0.42  0.67790
## Year2002          0.28038      0.39672      0.71  0.48121
## Year2003         -0.51226      0.33018     -1.55  0.12366
## Year2004         -0.08292      0.23241     -0.36  0.72195
## Year2005         -0.04925      0.29308     -0.17  0.86685
## Year2006          0.02195      0.21667      0.10  0.91948
## Year2007          0.02381      0.25005      0.10  0.92431
## Year2008          0.09584      0.25608      0.37  0.70894
## Year2009         -0.24311      0.24082     -1.01  0.31493
## Year2010         -0.04526      0.24106     -0.19  0.85142
## Year2011         -0.13281      0.21069     -0.63  0.52976
## Year2012          0.00819      0.19517      0.04  0.96660
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.439
## Multiple R-squared:  0.205, Adjusted R-squared:  0.0456
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 121 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.339  0.888  0.967  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      7.52e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.231 1      1.110
## LastAuthorFemale  1.674 1      1.294
## Year              1.931 16      1.021

```

Residuals from first and last author



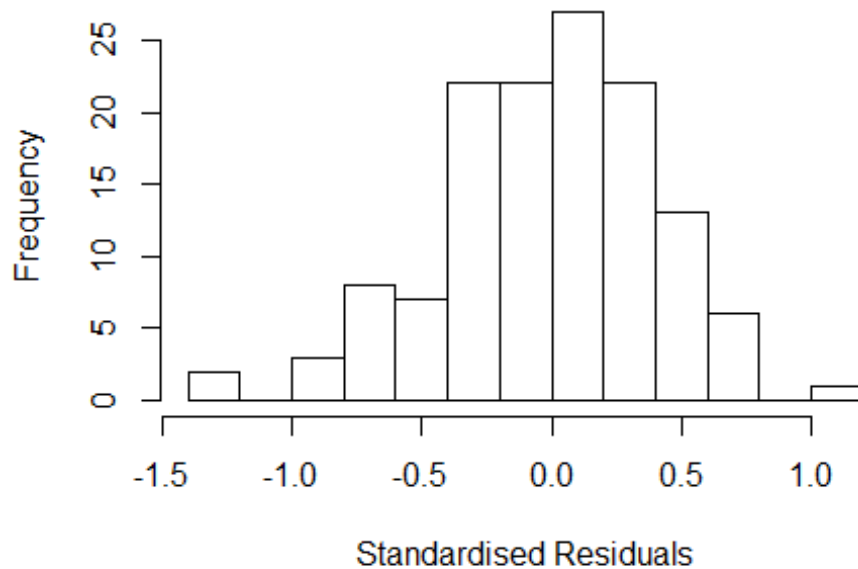
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3222 -0.2505 0.0212 0.2608 1.0287
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3630 0.1869 7.29 4.4e-11 ***
## FirstAuthorFemale1 -0.0420 0.0768 -0.55 0.5852
## LastAuthorFemale1 0.0529 0.1169 0.45 0.6517
## Year1997 -1.0669 0.3063 -3.48 0.0007 ***
## Year1998 -0.1721 0.2027 -0.85 0.3976
## Year1999 0.0392 0.2097 0.19 0.8519
## Year2000 0.3815 0.2105 1.81 0.0725 .
## Year2001 0.0898 0.2644 0.34 0.7347
## Year2002 0.2127 0.3228 0.66 0.5113
## Year2003 -0.5397 0.3463 -1.56 0.1219
## Year2004 -0.1055 0.2238 -0.47 0.6384
## Year2005 -0.1063 0.2940 -0.36 0.7184
```

```

## Year2006          -0.0312      0.2179   -0.14   0.8865
## Year2007          -0.0492      0.2590   -0.19   0.8495
## Year2008           0.0336      0.2337    0.14   0.8861
## Year2009          -0.2800      0.2354   -1.19   0.2367
## Year2010          -0.1176      0.2284   -0.51   0.6078
## Year2011          -0.1808      0.2143   -0.84   0.4007
## Year2012          -0.0408      0.1962   -0.21   0.8356
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.446
## Multiple R-squared:  0.19,   Adjusted R-squared:  0.0619
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 118 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.361  0.888  0.961  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      7.52e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## 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      1.094
## Year              1.198 16      1.006

```

Residuals from first author



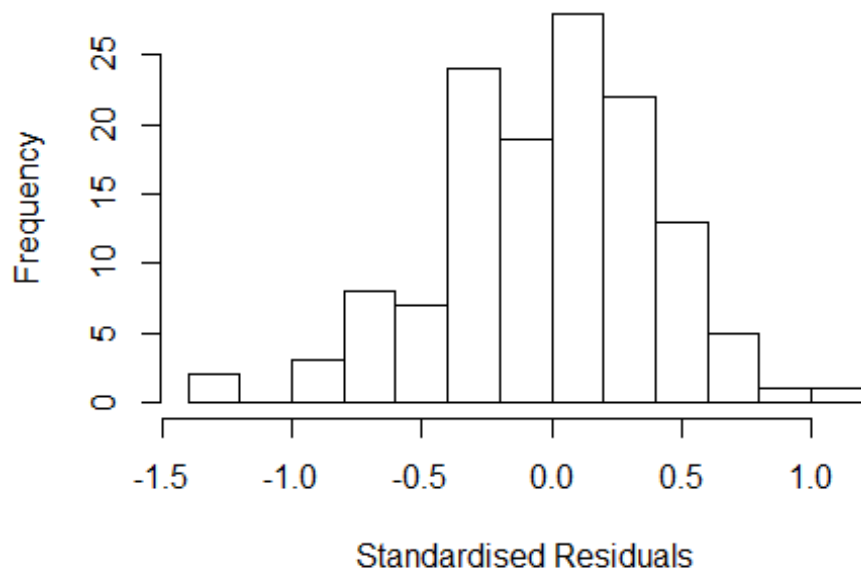
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3294 -0.2523 0.0232 0.2535 1.0154
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3895 0.1562 8.90 9.5e-15 ***
## FirstAuthorFemale1 -0.0447 0.0768 -0.58 0.56177
## Year1997 -1.0391 0.2867 -3.62 0.00043 ***
## Year1998 -0.1763 0.1822 -0.97 0.33525
## Year1999 0.0133 0.1835 0.07 0.94249
## Year2000 0.3564 0.1859 1.92 0.05765 .
## Year2001 0.0638 0.2445 0.26 0.79475
## Year2002 0.1879 0.3085 0.61 0.54363
## Year2003 -0.5529 0.3467 -1.59 0.11355
## Year2004 -0.1244 0.2016 -0.62 0.53850
## Year2005 -0.1292 0.2661 -0.49 0.62809
## Year2006 -0.0465 0.1945 -0.24 0.81145
```

```

## Year2007          -0.0735      0.2398   -0.31  0.75970
## Year2008           0.0141      0.2154    0.07  0.94774
## Year2009          -0.3009      0.2146   -1.40  0.16363
## Year2010          -0.1326      0.2119   -0.63  0.53273
## Year2011          -0.1937      0.1945   -1.00  0.32136
## Year2012          -0.0601      0.1741   -0.34  0.73073
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.441
## Multiple R-squared:  0.189, Adjusted R-squared:  0.0691
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 122 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.344  0.891   0.962   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      7.52e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.598 1          1.264
## Year              1.598 16          1.015

```

Residuals from last author



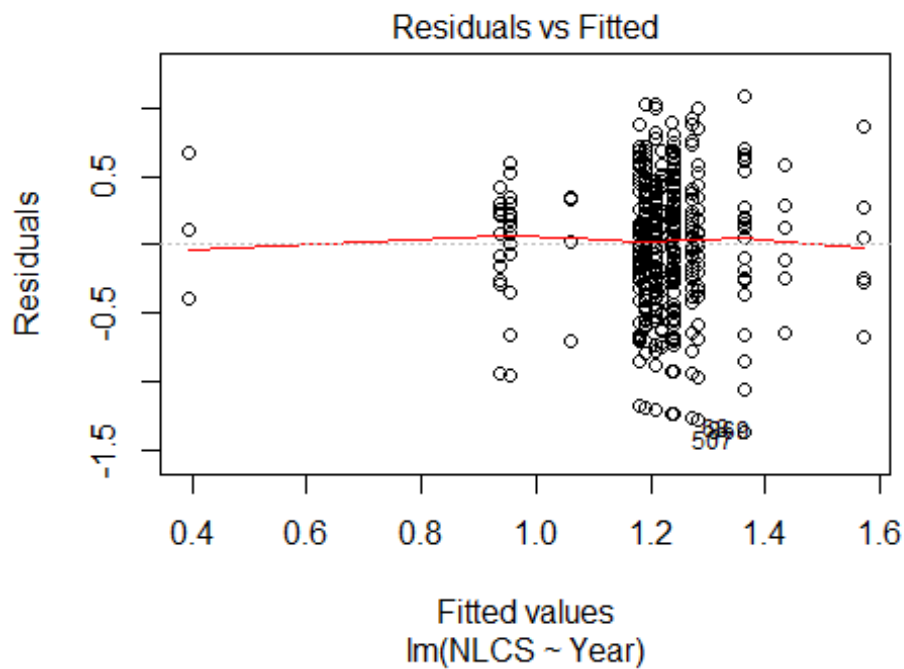
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3121 -0.2481 0.0305 0.2394 1.0394
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3616 0.1876 7.26 5e-11 ***
## LastAuthorFemale1 0.0557 0.1137 0.49 0.62499
## Year1997 -1.0894 0.3192 -3.41 0.00089 ***
## Year1998 -0.1800 0.2014 -0.89 0.37342
## Year1999 0.0321 0.2120 0.15 0.87993
## Year2000 0.3613 0.2075 1.74 0.08426 .
## Year2001 0.0844 0.2661 0.32 0.75157
## Year2002 0.1956 0.3237 0.60 0.54679
## Year2003 -0.5490 0.3408 -1.61 0.10991
## Year2004 -0.1208 0.2247 -0.54 0.59186
## Year2005 -0.1116 0.2921 -0.38 0.70322
## Year2006 -0.0395 0.2180 -0.18 0.85643
```

```

## Year2007          -0.0758      0.2594   -0.29   0.77075
## Year2008           0.0195      0.2321    0.08   0.93304
## Year2009          -0.2842      0.2378   -1.19   0.23459
## Year2010          -0.1299      0.2260   -0.58   0.56641
## Year2011          -0.1899      0.2138   -0.89   0.37629
## Year2012          -0.0495      0.1969   -0.25   0.80193
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.445
## Multiple R-squared:  0.188, Adjusted R-squared:  0.0684
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 122 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.365  0.891  0.961   0.918  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      7.52e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 133"
## [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
##   22   15   26   31   36   61   44   34   58   46   65   82   74  105  113
## 2011 2012
##  118  102
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    4   12    6    6   12   16   16   18   22   21   32   28   57   55
## 2011 2012

```

```
## 68 73
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3 4 11 6 6 11 14 15 15 20 14 24 24 49 47
## 2011 2012
## 61 60
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 16, df = 16, p-value = 0.4
```



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

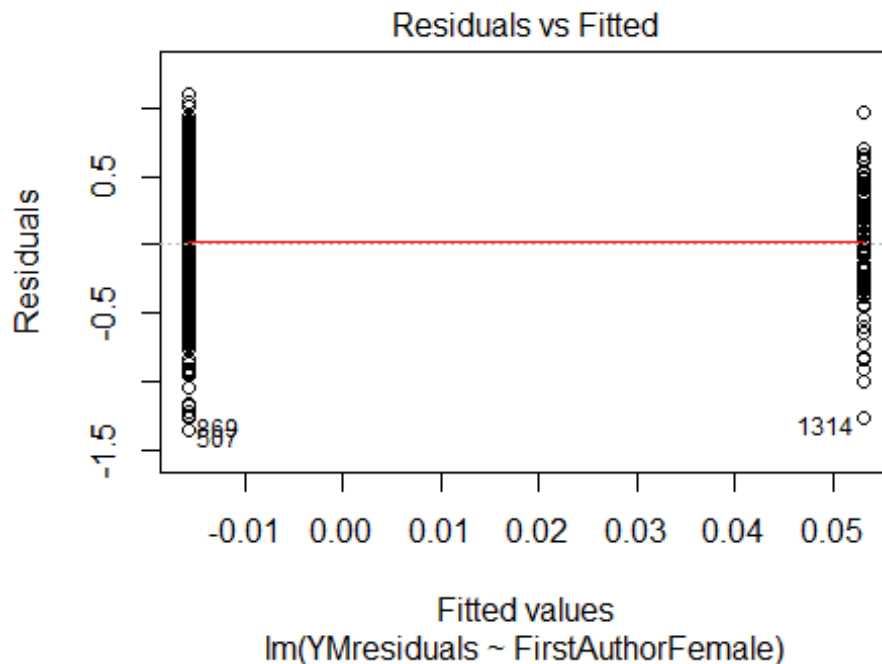
## [1] "Female first author team size 2018 geometric mean: 5.3835632709553"
## [1] "Male first author team size 2018 geometric mean: 3.67805321931899"

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



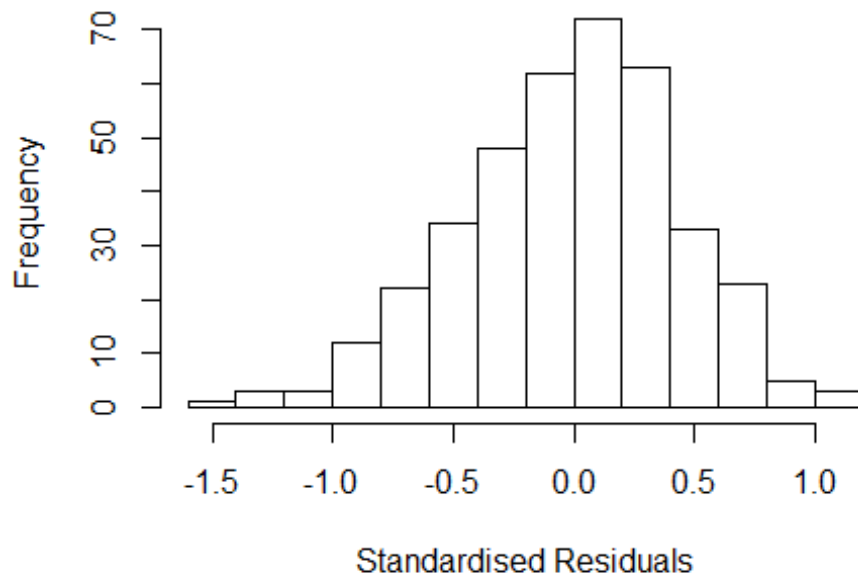
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 52, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5"
## [1] "Male last author team size 2018 geometric mean: 3.78514355229227"

## Warning in wilcox.test.default(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] "Regression 1: First author gender, last author gender, team size,
## Year as factors"
##
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.178 1      1.085
## LastAuthorFemale  1.198 1      1.094
## UniqueAuthors    2.282 4      1.109
## Year              2.969 16     1.035
```

Residuals from first and last author and team size



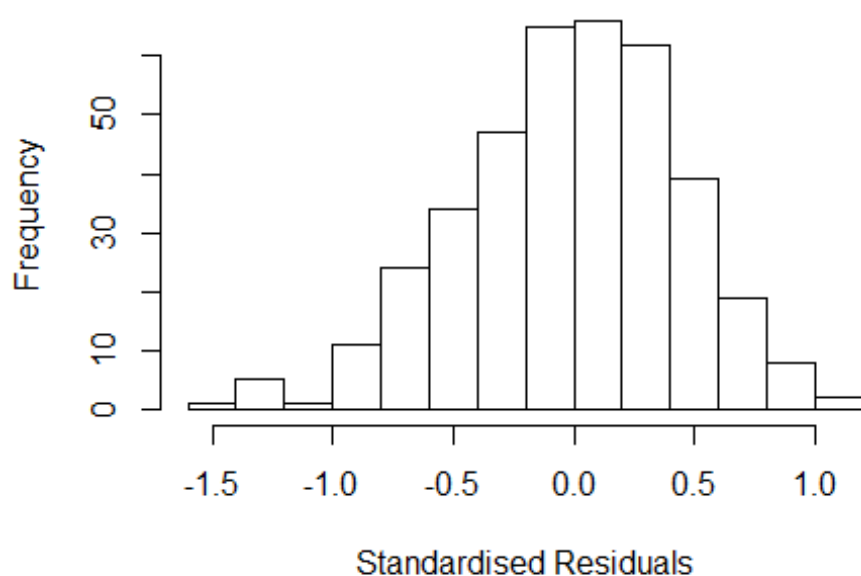
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4495 -0.2954 0.0159 0.2948 1.0916
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5444 0.3111 1.75 0.0810 .
## FirstAuthorFemale1 0.0423 0.0518 0.82 0.4137
## LastAuthorFemale1 0.0279 0.0681 0.41 0.6819
## UniqueAuthors2 -0.1224 0.1551 -0.79 0.4305
## UniqueAuthors3 -0.0939 0.1503 -0.63 0.5322
## UniqueAuthors4 -0.0741 0.1538 -0.48 0.6301
## UniqueAuthors5 -0.0489 0.1552 -0.31 0.7530
## Year1997 0.6474 0.3659 1.77 0.0777 .
## Year1998 0.7696 0.3469 2.22 0.0271 *
## Year1999 0.9856 0.3437 2.87 0.0044 **
```

```

## Year2000          1.0675      0.3562      3.00      0.0029 **
## Year2001          0.4761      0.3122      1.52      0.1282
## Year2002          0.9383      0.3092      3.03      0.0026 **
## Year2003          0.5287      0.3146      1.68      0.0937 .
## Year2004          0.7020      0.3075      2.28      0.0230 *
## Year2005          0.9792      0.3212      3.05      0.0025 **
## Year2006          0.8362      0.3198      2.61      0.0093 **
## Year2007          0.7692      0.3124      2.46      0.0143 *
## Year2008          0.8886      0.3072      2.89      0.0041 **
## Year2009          0.7084      0.2986      2.37      0.0182 *
## Year2010          0.7575      0.2981      2.54      0.0115 *
## Year2011          0.7378      0.2977      2.48      0.0137 *
## Year2012          0.7712      0.2987      2.58      0.0102 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.449
## Multiple R-squared:  0.0795, Adjusted R-squared:  0.0234
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 36 weights are ~= 1. The remaining 348 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.277  0.871  0.950  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          2.60e-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 1.142 1      1.069
## LastAuthorFemale  1.185 1      1.089
## Year              1.352 16      1.009

```

Residuals from first and last author



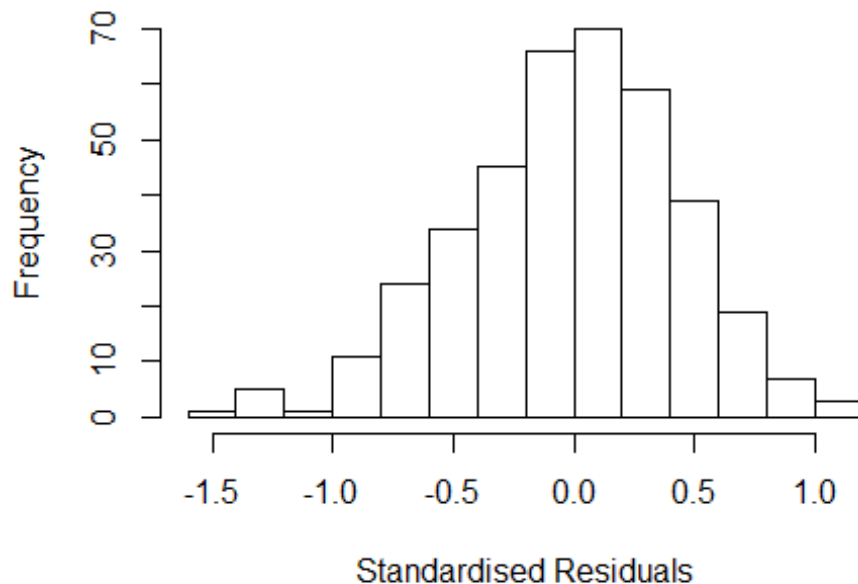
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4371 -0.3035 0.0202 0.2959 1.0456
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5052 0.2948 1.71 0.0874 .
## FirstAuthorFemale1 0.0375 0.0511 0.73 0.4640
## LastAuthorFemale1 0.0274 0.0681 0.40 0.6881
## Year1997 0.5801 0.3675 1.58 0.1153
## Year1998 0.7296 0.3473 2.10 0.0364 *
## Year1999 0.9271 0.3406 2.72 0.0068 **
## Year2000 1.0103 0.3611 2.80 0.0054 **
## Year2001 0.4420 0.3108 1.42 0.1558
## Year2002 0.8933 0.3088 2.89 0.0041 **
## Year2003 0.4732 0.3204 1.48 0.1406
## Year2004 0.6384 0.3056 2.09 0.0374 *
## Year2005 0.9319 0.3225 2.89 0.0041 **
```

```

## Year2006          0.7910      0.3153      2.51      0.0125 *
## Year2007          0.7376      0.3106      2.37      0.0181 *
## Year2008          0.8578      0.3077      2.79      0.0056 **
## Year2009          0.6713      0.3013      2.23      0.0265 *
## Year2010          0.7169      0.3008      2.38      0.0177 *
## Year2011          0.6955      0.2978      2.34      0.0200 *
## Year2012          0.7316      0.3003      2.44      0.0153 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.447
## Multiple R-squared:  0.0758, Adjusted R-squared:  0.0302
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 360 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.280  0.872  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      2.60e-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 1.144 1      1.070
## Year              1.144 16      1.004

```

Residuals from first author



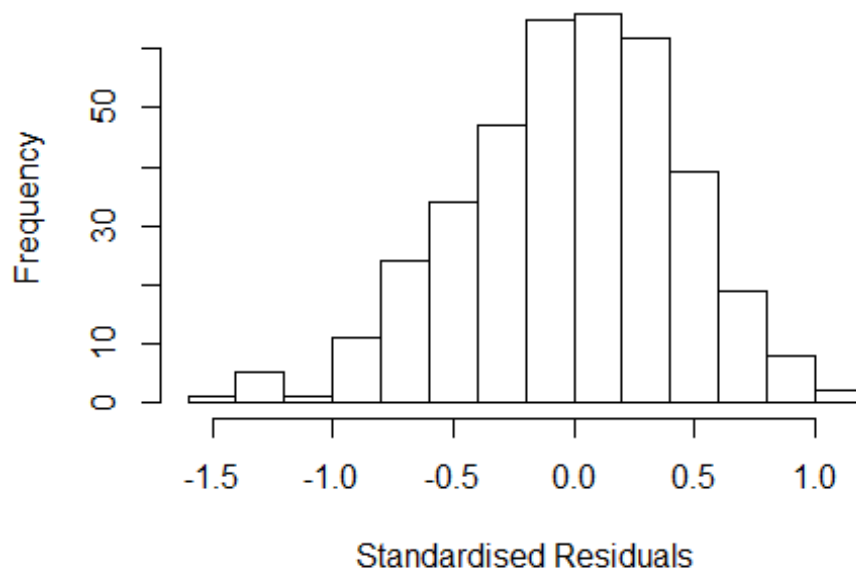
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4428 -0.3060 0.0172 0.2958 1.0439
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5045 0.2950 1.71 0.0881 .
## FirstAuthorFemale1 0.0389 0.0514 0.76 0.4492
## Year1997 0.5806 0.3671 1.58 0.1146
## Year1998 0.7381 0.3431 2.15 0.0321 *
## Year1999 0.9275 0.3409 2.72 0.0068 **
## Year2000 1.0128 0.3599 2.81 0.0052 **
## Year2001 0.4423 0.3104 1.42 0.1551
## Year2002 0.8938 0.3083 2.90 0.0040 **
## Year2003 0.4803 0.3188 1.51 0.1328
## Year2004 0.6445 0.3052 2.11 0.0354 *
## Year2005 0.9383 0.3210 2.92 0.0037 **
## Year2006 0.7958 0.3150 2.53 0.0119 *
```

```

## Year2007          0.7390      0.3102      2.38      0.0177 *
## Year2008          0.8594      0.3078      2.79      0.0055 **
## Year2009          0.6737      0.3015      2.23      0.0261 *
## Year2010          0.7208      0.3005      2.40      0.0170 *
## Year2011          0.7004      0.2975      2.35      0.0191 *
## Year2012          0.7344      0.3000      2.45      0.0148 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.446
## Multiple R-squared:  0.0755, Adjusted R-squared:  0.0325
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 359 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.274  0.871  0.954  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      2.60e-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 1.184 1      1.088
## Year              1.184 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
## -1.4466 -0.2963 0.0158 0.3049 1.0382
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5217 0.2912 1.79 0.0740 .
## LastAuthorFemale1 0.0311 0.0689 0.45 0.6518
## Year1997 0.5747 0.3697 1.55 0.1210
## Year1998 0.7193 0.3471 2.07 0.0389 *
## Year1999 0.9176 0.3375 2.72 0.0069 **
## Year2000 1.0137 0.3623 2.80 0.0054 **
## Year2001 0.4401 0.3103 1.42 0.1570
## Year2002 0.8886 0.3080 2.88 0.0041 **
## Year2003 0.4664 0.3194 1.46 0.1450
## Year2004 0.6319 0.3037 2.08 0.0382 *
## Year2005 0.9250 0.3212 2.88 0.0042 **
## Year2006 0.7832 0.3139 2.50 0.0130 *
```

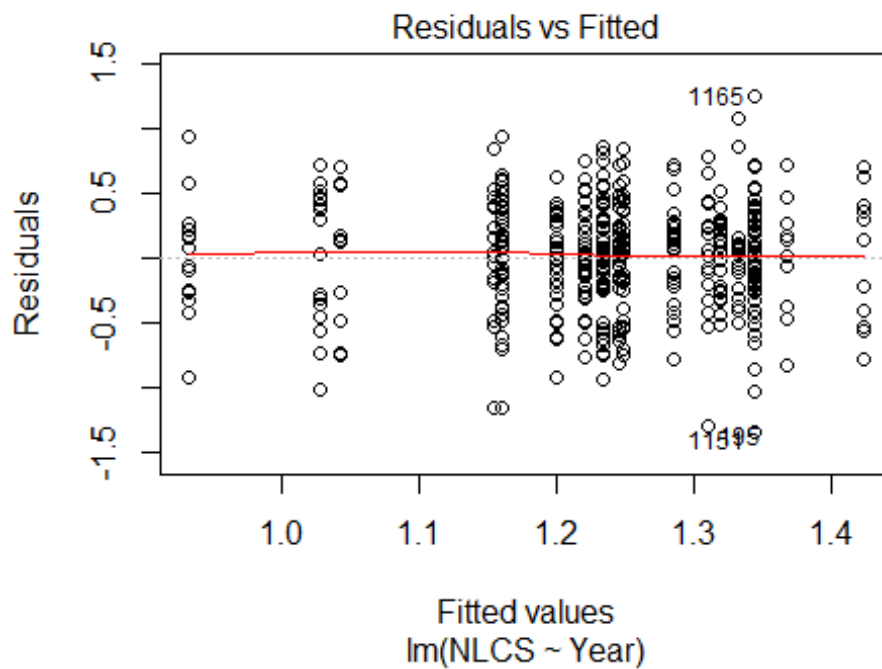


```

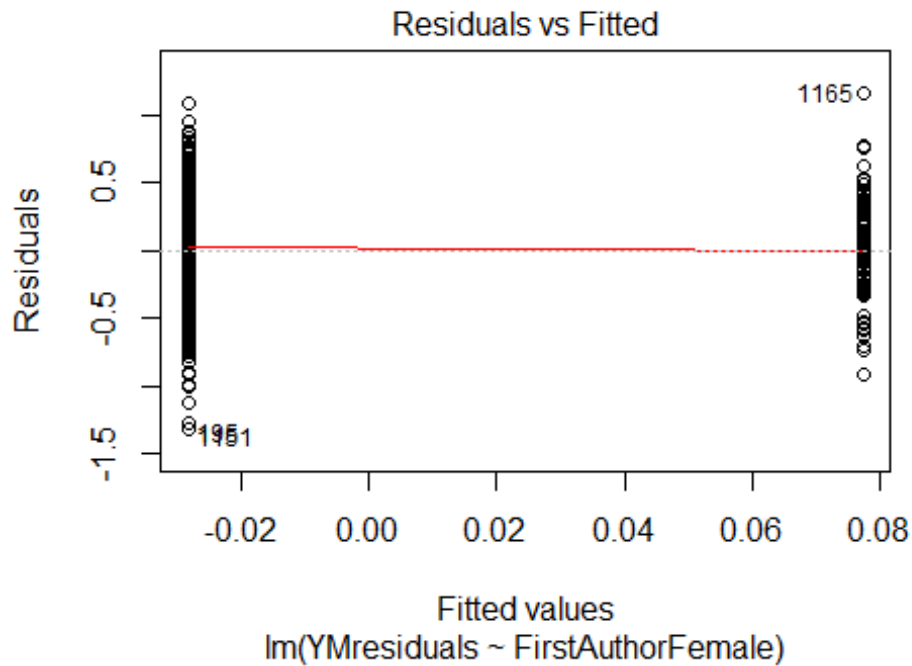
## Year2007          0.7358      0.3109      2.37      0.0185 *
## Year2008          0.8536      0.3062      2.79      0.0056 **
## Year2009          0.6621      0.2992      2.21      0.0275 *
## Year2010          0.7065      0.2990      2.36      0.0187 *
## Year2011          0.6876      0.2961      2.32      0.0208 *
## Year2012          0.7218      0.2988      2.42      0.0162 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.446
## Multiple R-squared:  0.0749, Adjusted R-squared:  0.0319
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 361 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.270  0.871  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      2.60e-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: 384"
## [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
##   35   28   34   39   34   50   24   41   44   45   50   48   52   56   60
## 2011 2012
##   65   84
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   10   16   18   17   11   14   11   21   23   30   30   32   30   36   42
## 2011 2012

```

```
## 50 61
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 10 16 17 15 11 13 9 18 21 23 26 29 27 33 39
## 2011 2012
## 43 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 = 15, df = 16, p-value = 0.5
```



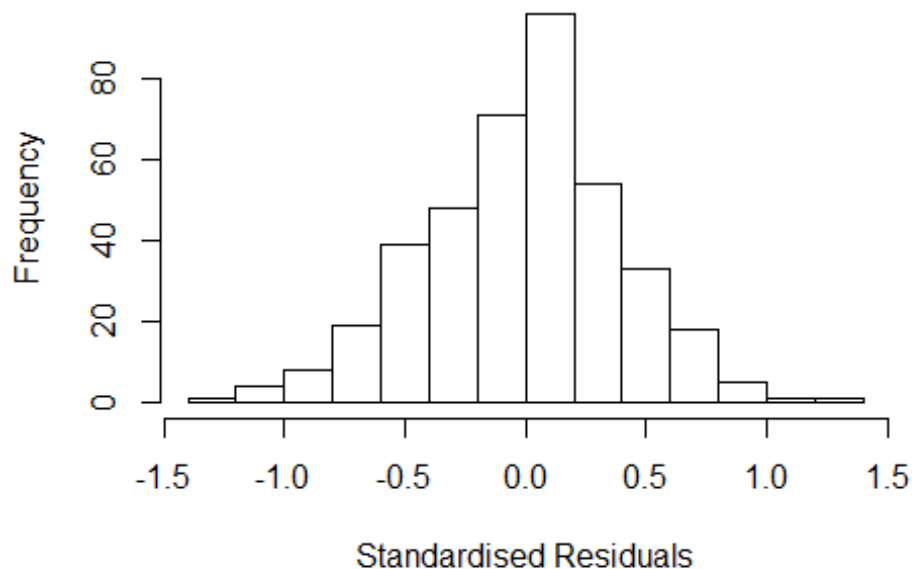
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 7.4, df = 1, p-value = 0.007
## [1] "Female first author team size 2018 geometric mean: 3.87298334620742"
## [1] "Male first author team size 2018 geometric mean: 2.21336383940064"
## 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 = 7.5, 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: 2.66716827534"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
```

| | GVIF | Df | GVIF^(1/(2*Df)) |
|-------------------|-------|----|-----------------|
| FirstAuthorFemale | 1.311 | 1 | 1.145 |
| LastAuthorFemale | 1.229 | 1 | 1.109 |
| UniqueAuthors | 2.724 | 4 | 1.133 |
| Year | 3.014 | 16 | 1.035 |

Residuals from first and last author and team size



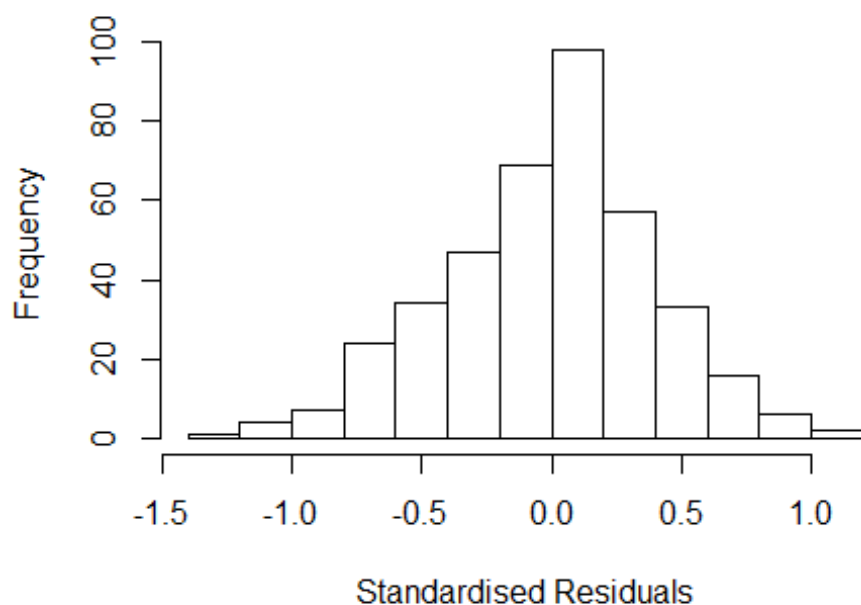
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3075 -0.2584  0.0295  0.2303  1.2173
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.33893    0.35793     3.74 0.00021 ***
## FirstAuthorFemale1  0.12015    0.04688     2.56 0.01077 *
## LastAuthorFemale1 -0.00877    0.05561    -0.16 0.87476
## UniqueAuthors2    -0.28393    0.21902    -1.30 0.19566
## UniqueAuthors3    -0.28377    0.22216    -1.28 0.20228
## UniqueAuthors4    -0.29908    0.21903    -1.37 0.17291
## UniqueAuthors5    -0.17364    0.21487    -0.81 0.41954
## Year1997          -0.08637    0.29048    -0.30 0.76639
## Year1998           0.14588    0.23531     0.62 0.53568
## Year1999           0.32710    0.25441     1.29 0.19932
```

```

## Year2000      0.27120      0.26015      1.04  0.29787
## Year2001     -0.18570      0.25989     -0.71  0.47535
## Year2002      0.28808      0.31503      0.91  0.36107
## Year2003      0.09015      0.25109      0.36  0.71977
## Year2004      0.19520      0.24175      0.81  0.41992
## Year2005      0.23624      0.24898      0.95  0.34333
## Year2006      0.10493      0.24596      0.43  0.66990
## Year2007      0.13957      0.24139      0.58  0.56349
## Year2008      0.08508      0.24139      0.35  0.72469
## Year2009      0.09065      0.24023      0.38  0.70612
## Year2010      0.07329      0.24003      0.31  0.76028
## Year2011      0.06261      0.24384      0.26  0.79751
## Year2012      0.26767      0.23571      1.14  0.25686
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.378
## Multiple R-squared:  0.105, Adjusted R-squared:  0.053
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 34 weights are ~= 1. The remaining 364 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.207  0.856  0.954  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      2.51e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500      50      2      1      1000      200
## trace.lev      mts      compute.rd
##      0      1000      0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.185 1      1.088
## LastAuthorFemale  1.127 1      1.062
## Year              1.250 16      1.007

```

Residuals from first and last author



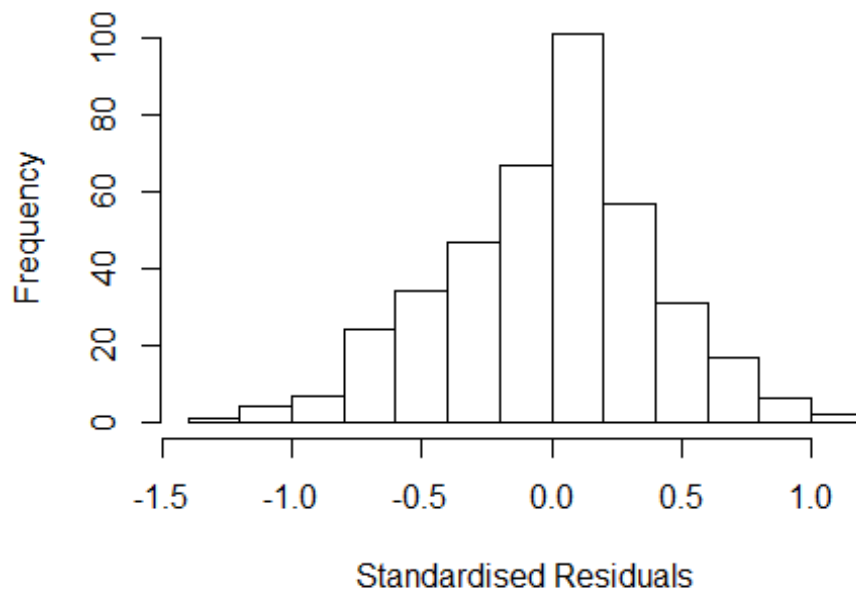
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.367 -0.259 0.021 0.237 1.184
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0435 0.2069 5.04 7.1e-07 ***
## FirstAuthorFemale1 0.1164 0.0447 2.60 0.0096 **
## LastAuthorFemale1 -0.0243 0.0548 -0.44 0.6584
## Year1997 -0.0211 0.2576 -0.08 0.9348
## Year1998 0.1750 0.2134 0.82 0.4127
## Year1999 0.3475 0.2320 1.50 0.1350
## Year2000 0.2843 0.2430 1.17 0.2428
## Year2001 -0.1462 0.2419 -0.60 0.5459
## Year2002 0.3139 0.3101 1.01 0.3120
## Year2003 0.1253 0.2302 0.54 0.5867
## Year2004 0.2335 0.2176 1.07 0.2840
## Year2005 0.2948 0.2220 1.33 0.1851
```

```

## Year2006          0.1660      0.2239      0.74      0.4589
## Year2007          0.1956      0.2143      0.91      0.3618
## Year2008          0.1518      0.2161      0.70      0.4827
## Year2009          0.1541      0.2153      0.72      0.4747
## Year2010          0.1219      0.2172      0.56      0.5751
## Year2011          0.1232      0.2166      0.57      0.5698
## Year2012          0.3236      0.2113      1.53      0.1265
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.383
## Multiple R-squared:  0.0841, Adjusted R-squared:  0.0406
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 41 weights are ~= 1. The remaining 357 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##   0.175  0.841   0.950   0.894   0.985   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      2.51e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
##   nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##   trace.lev      mts      compute.rd
##      0          1000          0
##           psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## 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
## -1.3644 -0.2693 0.0244 0.2375 1.1908
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0394 0.2033 5.11 5.1e-07 ***
## FirstAuthorFemale1 0.1164 0.0448 2.60 0.0098 **
## Year1997 -0.0193 0.2575 -0.07 0.9404
## Year1998 0.1729 0.2102 0.82 0.4115
## Year1999 0.3517 0.2288 1.54 0.1250
## Year2000 0.2867 0.2421 1.18 0.2371
## Year2001 -0.1445 0.2388 -0.60 0.5456
## Year2002 0.3180 0.3068 1.04 0.3005
## Year2003 0.1243 0.2297 0.54 0.5887
## Year2004 0.2351 0.2156 1.09 0.2760
## Year2005 0.2961 0.2207 1.34 0.1805
## Year2006 0.1659 0.2221 0.75 0.4556
```

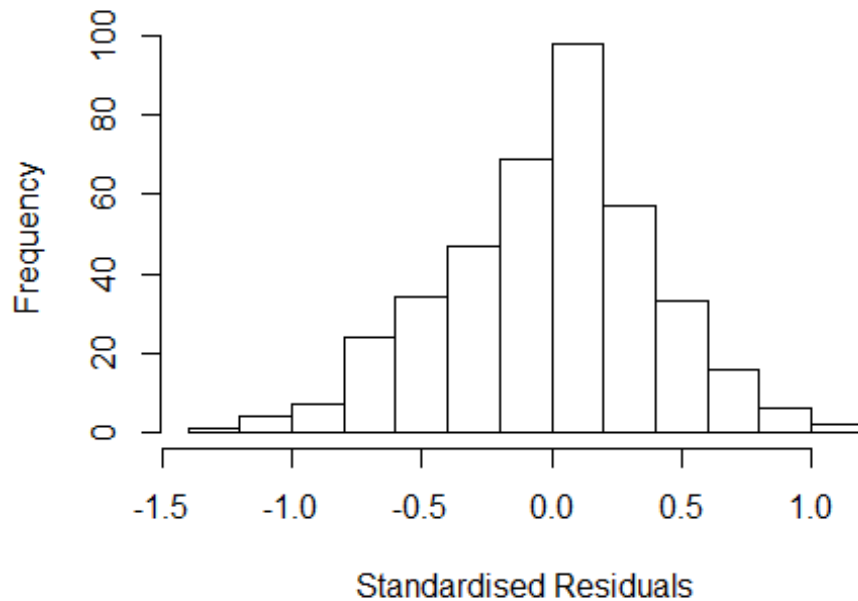


```

## Year2007          0.1977      0.2121      0.93      0.3518
## Year2008          0.1522      0.2148      0.71      0.4790
## Year2009          0.1540      0.2131      0.72      0.4703
## Year2010          0.1226      0.2161      0.57      0.5710
## Year2011          0.1239      0.2140      0.58      0.5629
## Year2012          0.3250      0.2092      1.55      0.1211
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.384
## Multiple R-squared:  0.0838, Adjusted R-squared:  0.0428
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 41 weights are ~= 1. The remaining 357 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.180  0.841   0.950   0.894   0.985   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      2.51e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.111 1      1.054
## Year              1.111 16      1.003

```

Residuals from last author



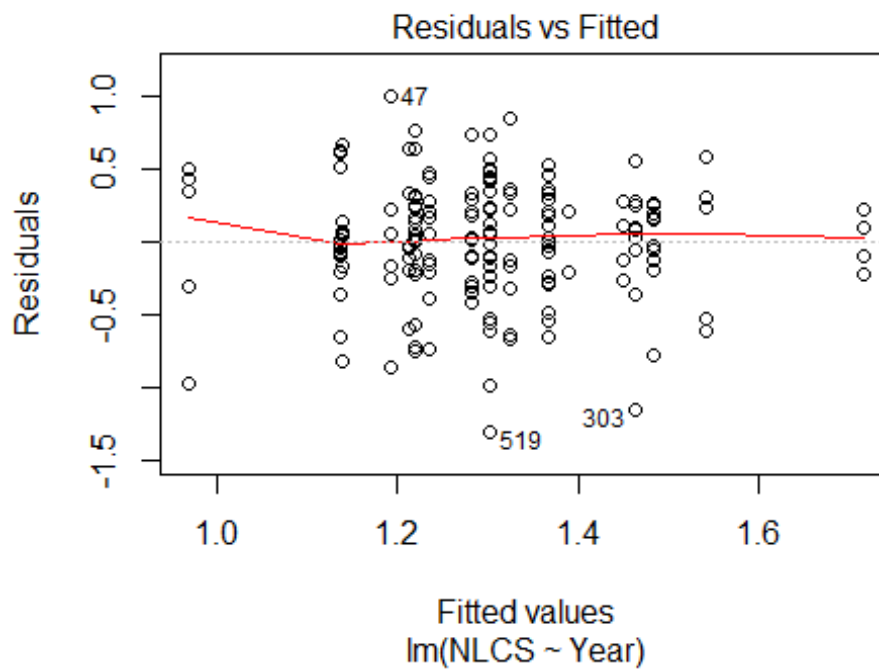
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3923 -0.2595 0.0317 0.2487 1.1937
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.06980 0.21665 4.94 1.2e-06 ***
## LastAuthorFemale1 -0.02287 0.05630 -0.41 0.68
## Year1997 0.00667 0.27392 0.02 0.98
## Year1998 0.17720 0.22392 0.79 0.43
## Year1999 0.33921 0.23893 1.42 0.16
## Year2000 0.31670 0.25216 1.26 0.21
## Year2001 -0.14949 0.25058 -0.60 0.55
## Year2002 0.32060 0.31459 1.02 0.31
## Year2003 0.14319 0.23901 0.60 0.55
## Year2004 0.26126 0.22625 1.15 0.25
## Year2005 0.29136 0.23132 1.26 0.21
## Year2006 0.17799 0.23373 0.76 0.45
```

```

## Year2007      0.21698      0.22543      0.96      0.34
## Year2008      0.16509      0.22659      0.73      0.47
## Year2009      0.14334      0.22468      0.64      0.52
## Year2010      0.12202      0.22828      0.53      0.59
## Year2011      0.13096      0.22647      0.58      0.56
## Year2012      0.32253      0.22155      1.46      0.15
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.384
## Multiple R-squared:  0.0685, Adjusted R-squared:  0.0268
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 34 weights are ~= 1. The remaining 364 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.160  0.845   0.948   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      2.51e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 398"
## [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
##   21   16   26   24   17   31   24   20   40   21   25   21   22   29   22
## 2011 2012
##   22   36
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    5    6    4    4   11    5    7   15   10   11   11    7   19   13
## 2011 2012

```

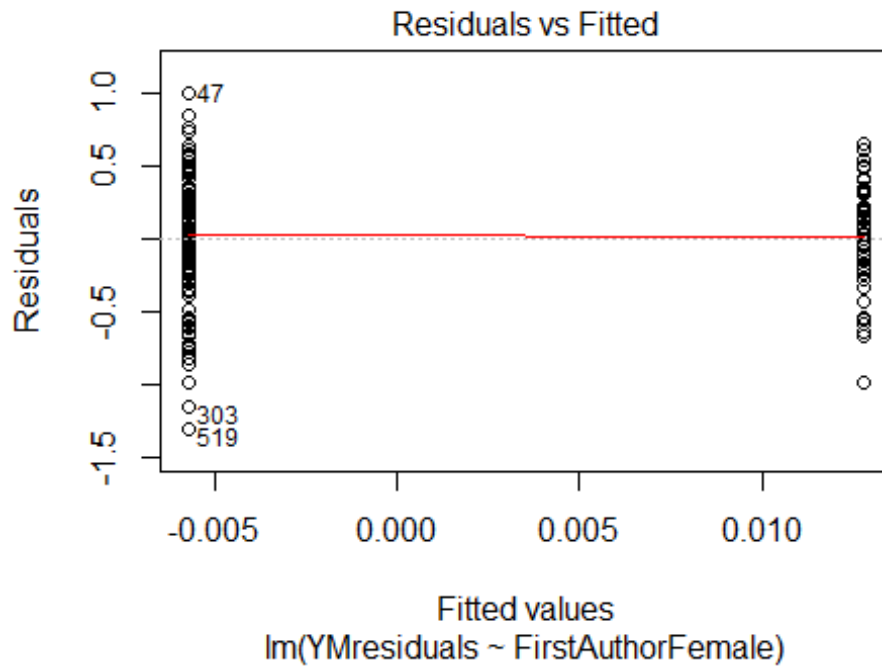
```
## 20 31
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2 5 5 4 4 9 5 6 11 7 7 9 4 14 13
## 2011 2012
## 17 24
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 14, df = 16, p-value = 0.6
```



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

## [1] "Female first author team size 2018 geometric mean: 3.46410161513775"
## [1] "Male first author team size 2018 geometric mean: 4.33788708479079"

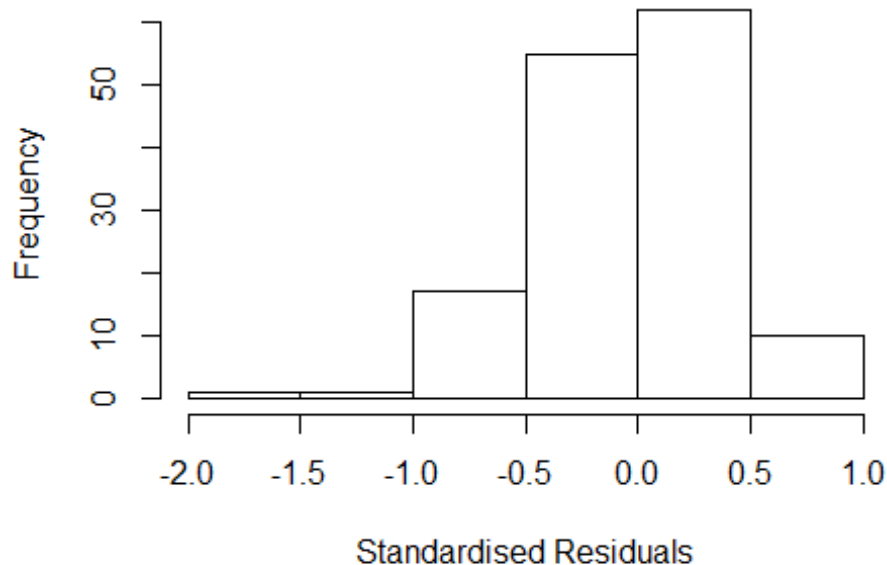
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3, p-value = 0.5
## 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: 4.06787401958886"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
```

| | GVIF | Df | GVIF^(1/(2*Df)) |
|----------------------|--------|----|-----------------|
| ## FirstAuthorFemale | 2.570 | 1 | 1.603 |
| ## LastAuthorFemale | 1.746 | 1 | 1.321 |
| ## UniqueAuthors | 9.760 | 4 | 1.329 |
| ## Year | 27.260 | 16 | 1.109 |

Residuals from first and last author and team size



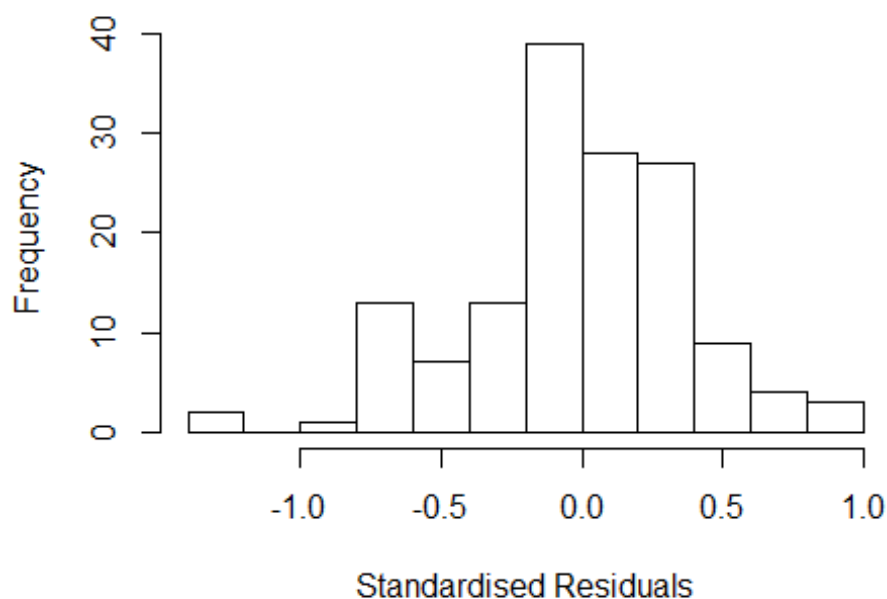
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.8467 -0.1879 -0.0128 0.2440 0.8413
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.73916 0.30574 5.69 8.8e-08 ***
## FirstAuthorFemale1 0.02641 0.08237 0.32 0.749
## LastAuthorFemale1 0.14244 0.07960 1.79 0.076 .
## UniqueAuthors2 -0.42088 0.20944 -2.01 0.047 *
## UniqueAuthors3 -0.46765 0.22712 -2.06 0.042 *
## UniqueAuthors4 -0.48439 0.21274 -2.28 0.025 *
## UniqueAuthors5 -0.46895 0.22096 -2.12 0.036 *
## Year1997 -0.06134 0.33626 -0.18 0.856
## Year1998 -0.33057 0.27664 -1.19 0.234
## Year1999 0.15889 0.25655 0.62 0.537
```

```

## Year2000      0.36297      0.24485      1.48      0.141
## Year2001      0.00937      0.28543      0.03      0.974
## Year2002      0.12630      0.40226      0.31      0.754
## Year2003     -0.13851      0.29928     -0.46      0.644
## Year2004     -0.11375      0.25212     -0.45      0.653
## Year2005      0.17430      0.26074      0.67      0.505
## Year2006      0.01699      0.26890      0.06      0.950
## Year2007      0.15207      0.26082      0.58      0.561
## Year2008     -0.02664      0.26702     -0.10      0.921
## Year2009      0.00594      0.25181      0.02      0.981
## Year2010     -0.19288      0.25717     -0.75      0.455
## Year2011     -0.18904      0.24697     -0.77      0.445
## Year2012      0.07709      0.24452      0.32      0.753
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.373
## Multiple R-squared:  0.179, Adjusted R-squared:  0.0321
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## observation 3 is an outlier with |weight| = 0 ( < 0.00068);
## 17 weights are ~= 1. The remaining 128 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.177  0.877  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           6.85e-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 2.140 1 1.463
## LastAuthorFemale 2.029 1 1.424
## Year 3.937 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
## -1.36997 -0.19016 -0.00669 0.24828 0.85931
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.33163 0.21813 6.10 1.2e-08 ***
## FirstAuthorFemale1 -0.00155 0.07906 -0.02 0.98
## LastAuthorFemale1 0.11574 0.09180 1.26 0.21
## Year1997 -0.18968 0.45082 -0.42 0.67
## Year1998 -0.36439 0.27158 -1.34 0.18
## Year1999 0.11789 0.24340 0.48 0.63
## Year2000 0.35459 0.23312 1.52 0.13
## Year2001 -0.02194 0.28309 -0.08 0.94
## Year2002 0.23910 0.37857 0.63 0.53
## Year2003 -0.16683 0.27304 -0.61 0.54
## Year2004 -0.13803 0.23509 -0.59 0.56
## Year2005 0.20128 0.24068 0.84 0.40
```

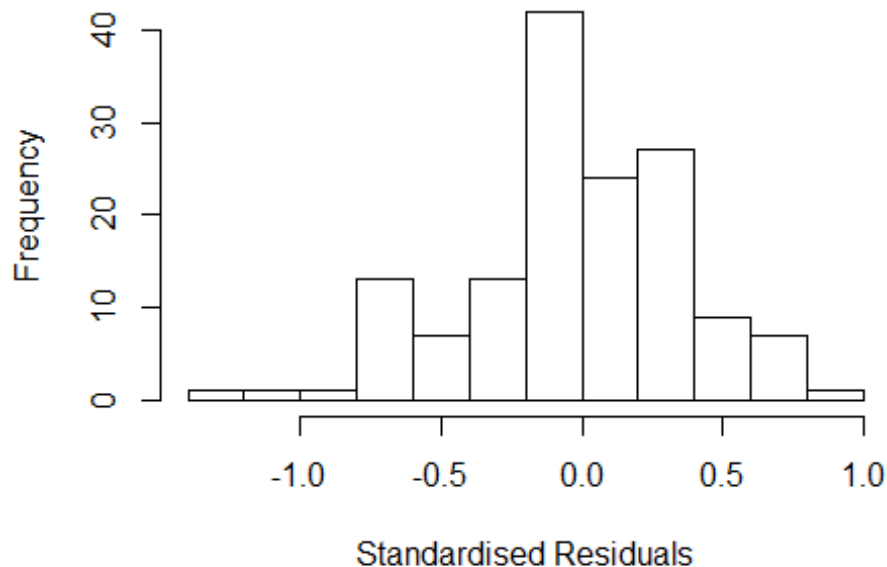


```

## Year2006      -0.00440    0.26390   -0.02    0.99
## Year2007      0.13238    0.25214    0.53    0.60
## Year2008     -0.05363    0.25945   -0.21    0.84
## Year2009     -0.04014    0.23571   -0.17    0.87
## Year2010     -0.24410    0.23629   -1.03    0.30
## Year2011     -0.17364    0.23949   -0.73    0.47
## Year2012      0.03834    0.23369    0.16    0.87
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.363
## Multiple R-squared:  0.152, Adjusted R-squared:  0.0323
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 130 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.124  0.862  0.955  0.896  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      6.85e-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 2.006 1      1.416
## Year              2.006 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
## -1.3851 -0.1961 -0.0147 0.2368 0.8579
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.38950 0.15948 8.71 1.3e-14 ***
## FirstAuthorFemale1 -0.00716 0.07957 -0.09 0.928
## Year1997 -0.25927 0.35153 -0.74 0.462
## Year1998 -0.36235 0.23624 -1.53 0.128
## Year1999 0.06135 0.19258 0.32 0.751
## Year2000 0.33004 0.18785 1.76 0.081 .
## Year2001 -0.07844 0.24069 -0.33 0.745
## Year2002 0.18352 0.34463 0.53 0.595
## Year2003 -0.17567 0.24155 -0.73 0.468
## Year2004 -0.15953 0.18888 -0.84 0.400
## Year2005 0.16274 0.18962 0.86 0.392
## Year2006 -0.02290 0.21684 -0.11 0.916
```

```

## Year2007          0.07926    0.20344    0.39    0.697
## Year2008          -0.08769    0.23066   -0.38    0.704
## Year2009          -0.08126    0.18888   -0.43    0.668
## Year2010          -0.26689    0.19298   -1.38    0.169
## Year2011          -0.19832    0.19469   -1.02    0.310
## Year2012          -0.00438    0.18225   -0.02    0.981
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.365
## Multiple R-squared:  0.142, Adjusted R-squared:  0.0275
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 129 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.118  0.849  0.957  0.895  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      6.85e-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 1.958 1      1.399
## Year      1.958 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
## -1.3691 -0.1899 -0.0066  0.2489  0.8596

```

```

##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.33179    0.21760    6.12 1.1e-08 ***
## LastAuthorFemale1 0.11541    0.09107    1.27  0.21
## Year1997      -0.19376    0.43497   -0.45  0.66
## Year1998      -0.36481    0.27140   -1.34  0.18
## Year1999       0.11736    0.24294    0.48  0.63
## Year2000       0.35370    0.22790    1.55  0.12
## Year2001      -0.02235    0.28235   -0.08  0.94
## Year2002       0.23786    0.37326    0.64  0.53
## Year2003      -0.16704    0.27159   -0.62  0.54
## Year2004      -0.13845    0.23361   -0.59  0.55
## Year2005       0.20079    0.24009    0.84  0.40
## Year2006      -0.00547    0.25841   -0.02  0.98
## Year2007       0.13063    0.23642    0.55  0.58
## Year2008      -0.05477    0.25628   -0.21  0.83
## Year2009      -0.04048    0.23541   -0.17  0.86
## Year2010      -0.24465    0.23411   -1.05  0.30
## Year2011      -0.17415    0.23652   -0.74  0.46
## Year2012       0.03731    0.23238    0.16  0.87
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.365
## Multiple R-squared:  0.152, Adjusted R-squared:  0.0394
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 131 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.128  0.865  0.955  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           6.85e-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: 146"
## [1] ""

```

```

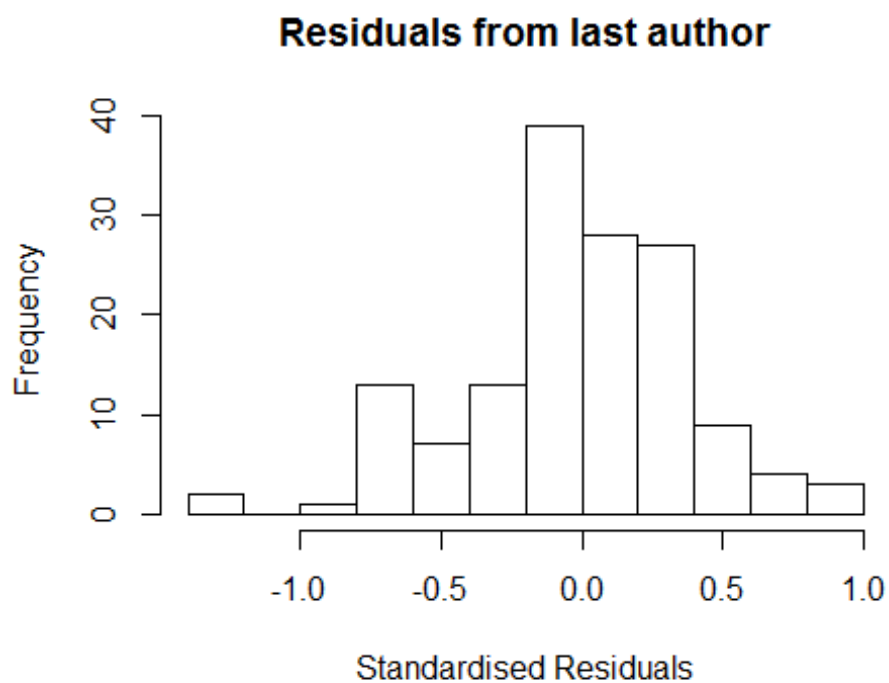
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2507"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6   10    6    2    6    9    7    5    4    8   11    9   13   11    9
## 2011 2012
##   11   20
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    6    0    2    0    3    1    2    2    4    6    5    6    8    7
## 2011 2012
##    6   11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    6    0    2    0    3    1    2    1    3    6    4    6    8    7
## 2011 2012
##    5    9
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.1748021039364"
## [1] "Male first author team size 2018 geometric mean: 4.23386572605092"

## Warning in wilcox.test.default(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"
## [1] "Male last author team size 2018 geometric mean: 3.82771028239322"

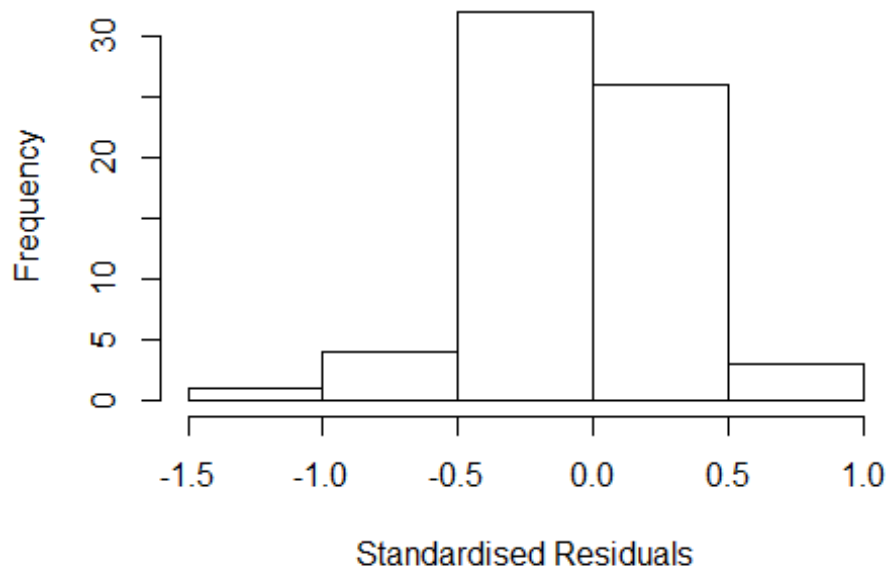
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## 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"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.679e+00  1          1.918
## LastAuthorFemale  2.582e+00  1          1.607
## UniqueAuthors    4.352e+16  4         120.182
## Year              1.557e+17 14          4.112
```

Residuals from first and last author and team size



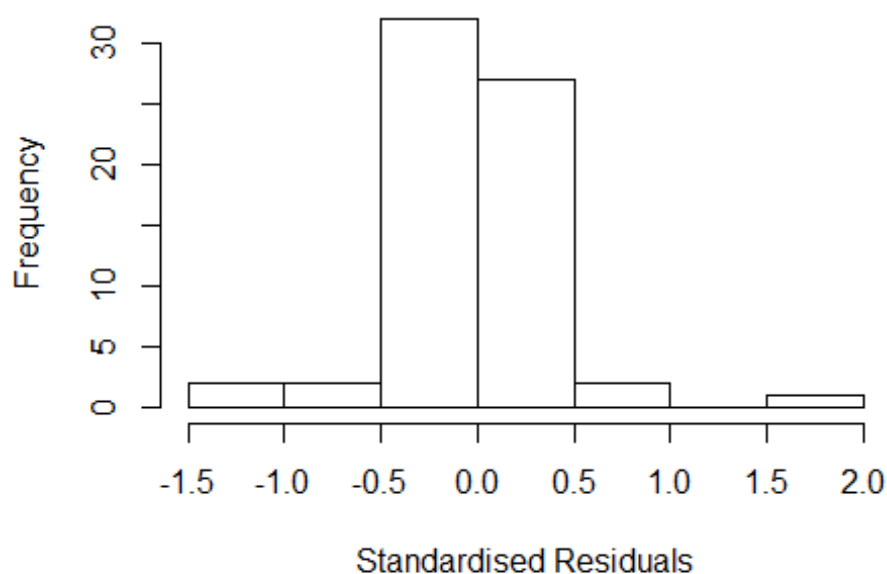
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3785 -0.1594 -0.0282 0.1793 0.9611
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.3799 0.1071 3.55 0.00093 ***
## FirstAuthorFemale1 0.0204 0.1160 0.18 0.86137
## LastAuthorFemale1 0.0617 0.2288 0.27 0.78881
## UniqueAuthors2 1.3193 0.1377 9.58 1.9e-12 ***
## UniqueAuthors3 1.4981 0.2089 7.17 5.7e-09 ***
## UniqueAuthors4 1.1595 0.1890 6.13 2.0e-07 ***
## UniqueAuthors5 1.3410 0.1883 7.12 6.7e-09 ***
## Year1997 -1.0981 0.3140 -3.50 0.00107 **
## Year1999 -0.7531 0.1961 -3.84 0.00038 ***
## Year2001 -1.1621 0.2259 -5.14 5.7e-06 ***
```

```

## Year2002          -0.6482      0.1049    -6.18  1.7e-07 ***
## Year2003          -0.2291      0.1380    -1.66  0.10395
## Year2004          -1.1479      0.1594    -7.20  5.1e-09 ***
## Year2005           0.1826      0.1606     1.14  0.26172
## Year2006          -0.4617      0.2386    -1.94  0.05924 .
## Year2007          -0.6375      0.2126    -3.00  0.00441 **
## Year2008          -0.4103      0.1821    -2.25  0.02917 *
## Year2009          -0.4941      0.1383    -3.57  0.00086 ***
## Year2010          -0.3207      0.1525    -2.10  0.04113 *
## Year2011          -0.3366      0.1675    -2.01  0.05052 .
## Year2012          -0.4176      0.1973    -2.12  0.03985 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.296
## Multiple R-squared:  0.628, Adjusted R-squared:  0.463
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## observation 46 is an outlier with |weight| <= 9.6e-05 ( < 0.0015);
## 6 weights are ~= 1. The remaining 59 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.269  0.906  0.966  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          1.52e-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 2: First author gender, Last author gender, Year as
factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.642 1          1.625
## LastAuthorFemale  1.515 1          1.231
## Year              4.337 14          1.054

```


Residuals from first and last author



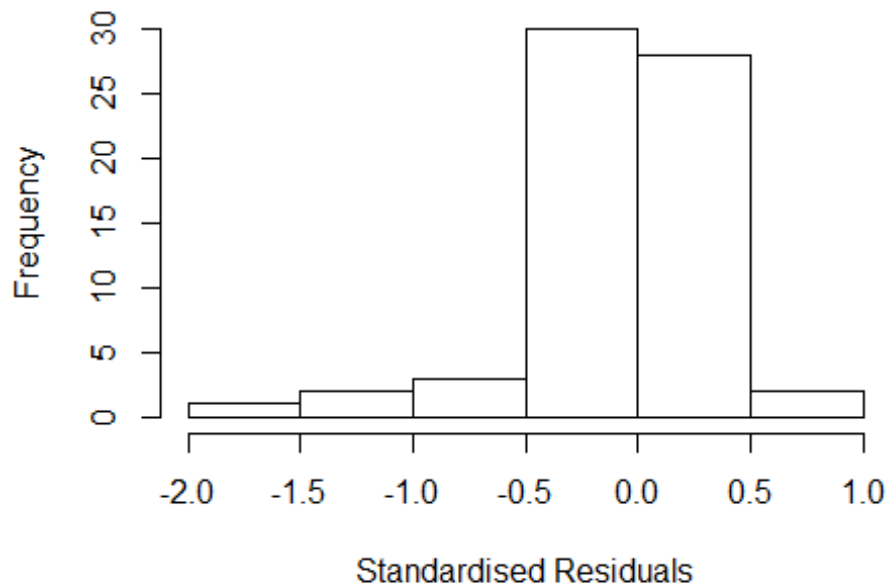
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.24854 -0.19087 -0.00696 0.21321 1.56214
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4641 0.4674 3.13 0.0029 **
## FirstAuthorFemale1 0.0675 0.1363 0.50 0.6225
## LastAuthorFemale1 0.1657 0.1895 0.87 0.3862
## Year1997 -0.7207 0.5062 -1.42 0.1609
## Year1999 -0.7013 0.4481 -1.56 0.1240
## Year2001 -0.9701 0.4985 -1.95 0.0574 .
## Year2002 -0.4131 0.4674 -0.88 0.3811
## Year2003 -1.4052 0.7037 -2.00 0.0514 .
## Year2004 -0.8911 0.4674 -1.91 0.0624 .
## Year2005 -0.5160 0.6050 -0.85 0.3979
## Year2006 -0.2223 0.4557 -0.49 0.6278
## Year2007 -0.4900 0.4661 -1.05 0.2983
```

```

## Year2008          -0.2137      0.4966   -0.43   0.6689
## Year2009          -0.1919      0.4681   -0.41   0.6835
## Year2010          -0.2155      0.4440   -0.49   0.6295
## Year2011          -0.1786      0.4610   -0.39   0.7002
## Year2012          -0.2136      0.4685   -0.46   0.6504
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.371
## Multiple R-squared:  0.414, Adjusted R-squared:  0.223
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 60 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0376 0.9110 0.9660 0.8960 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      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 1.801 1          1.342
## Year              1.801 14          1.021

```

Residuals from first author



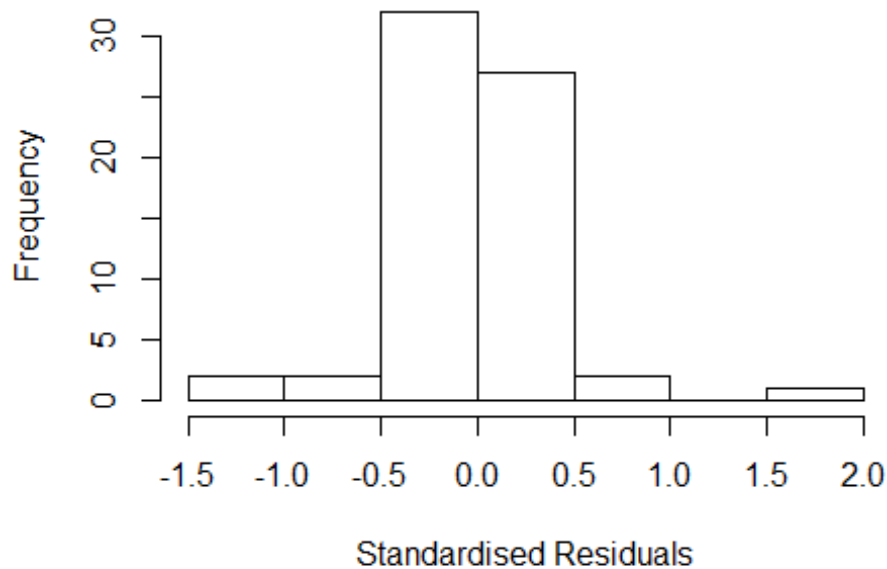
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5969 -0.1923 -0.0174 0.2046 0.9973
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4994 0.4288 3.50 0.001 ***
## FirstAuthorFemale1 0.0487 0.1354 0.36 0.721
## Year1997 -0.7557 0.4703 -1.61 0.114
## Year1999 -0.7272 0.4148 -1.75 0.086 .
## Year2001 -1.0045 0.4624 -2.17 0.035 *
## Year2002 -0.4484 0.4288 -1.05 0.301
## Year2003 0.0975 0.3733 0.26 0.795
## Year2004 -0.9264 0.4288 -2.16 0.036 *
## Year2005 -0.5545 0.5784 -0.96 0.342
## Year2006 -0.2521 0.4172 -0.60 0.548
## Year2007 -0.4790 0.4432 -1.08 0.285
## Year2008 -0.2493 0.4612 -0.54 0.591
```

```

## Year2009          -0.2251      0.4308   -0.52    0.604
## Year2010          -0.2418      0.4052   -0.60    0.553
## Year2011          -0.1705      0.4225   -0.40    0.688
## Year2012          -0.2462      0.4329   -0.57    0.572
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.364
## Multiple R-squared:  0.37,   Adjusted R-squared:  0.181
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 60 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0151 0.9090 0.9640 0.8920 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      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 1.063 1          1.031
## Year            1.063 14          1.002

```

Residuals from last author



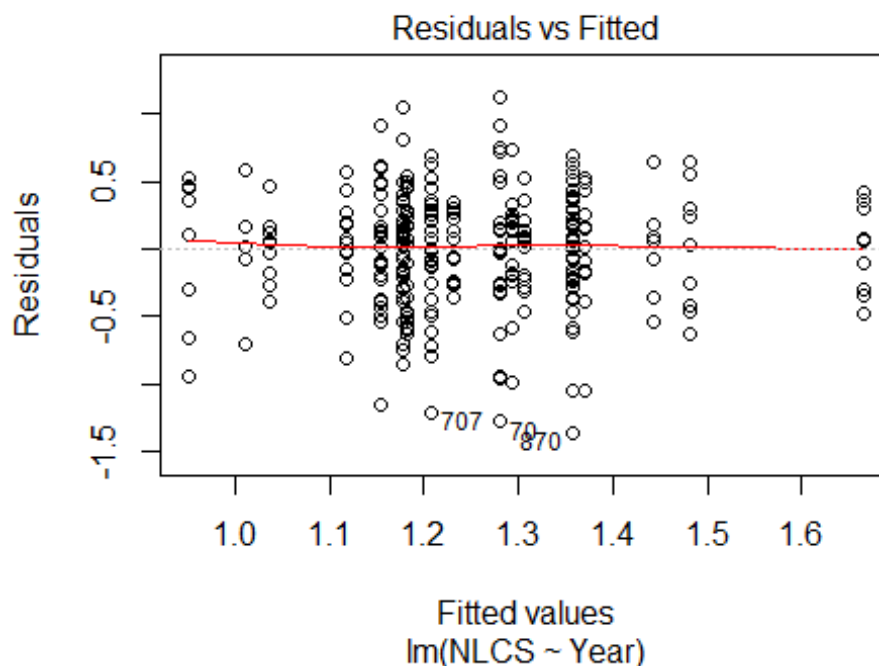
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2682 -0.1795 -0.0208 0.2089 1.6000
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.542 0.361 4.26 8.9e-05 ***
## LastAuthorFemale1 0.140 0.188 0.74 0.4610
## Year1997 -0.787 0.425 -1.85 0.0702 .
## Year1999 -0.745 0.388 -1.92 0.0607 .
## Year2001 -1.047 0.402 -2.60 0.0121 *
## Year2002 -0.491 0.361 -1.36 0.1807
## Year2003 -1.521 0.454 -3.35 0.0015 **
## Year2004 -0.969 0.361 -2.68 0.0099 **
## Year2005 -0.597 0.529 -1.13 0.2647
## Year2006 -0.284 0.371 -0.76 0.4485
## Year2007 -0.549 0.369 -1.49 0.1435
## Year2008 -0.292 0.399 -0.73 0.4677
```

```

## Year2009          -0.261      0.373   -0.70   0.4865
## Year2010          -0.274      0.354   -0.77   0.4430
## Year2011          -0.223      0.394   -0.57   0.5732
## Year2012          -0.285      0.373   -0.76   0.4490
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.363
## Multiple R-squared:  0.433, Adjusted R-squared:  0.262
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 60 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0131 0.9100 0.9690 0.8910 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      1.52e-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 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
##   20   21   40   42   31   24   17   28   37   35   39   47   49   62   54
## 2011 2012
##   50   54
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    8   20    9    7    9    9   14   14   18   13   26   14   30   25
## 2011 2012
##   35   43
##

```

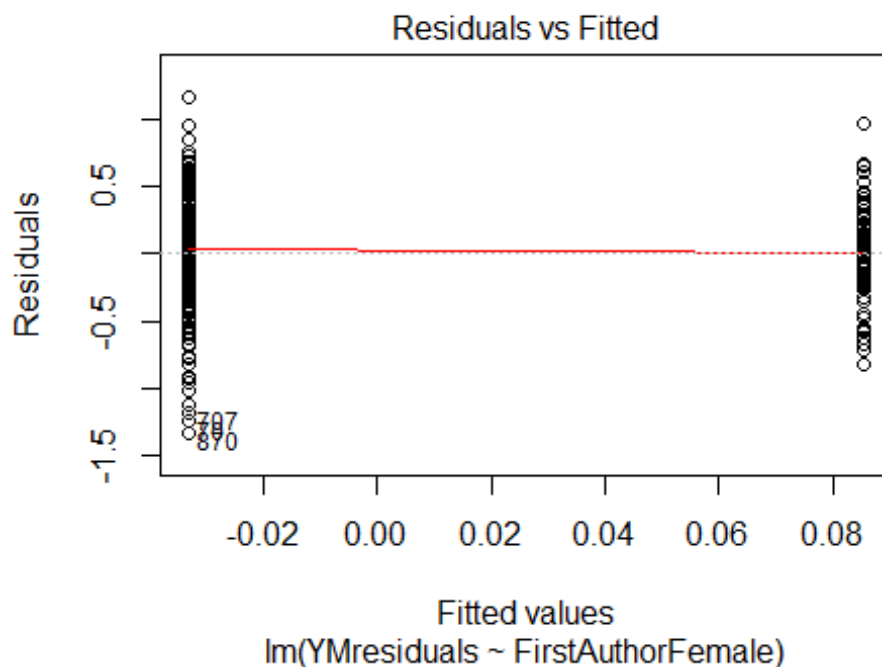
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    8   14    9    7    8    8   12   13   11    9   21   11   25   24
## 2011 2012
##   30   33
## [1] "Heteroscedasticity checks, confirming that there are problems with
##      these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 24, df = 16, p-value = 0.09
```



```
##
## 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: 3.91486764116886"
## [1] "Male first author team size 2018 geometric mean: 4.06432565052329"
##
## Warning in wilcox.test.default(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 = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 7"
## [1] "Male last author team size 2018 geometric mean: 3.8903939311"

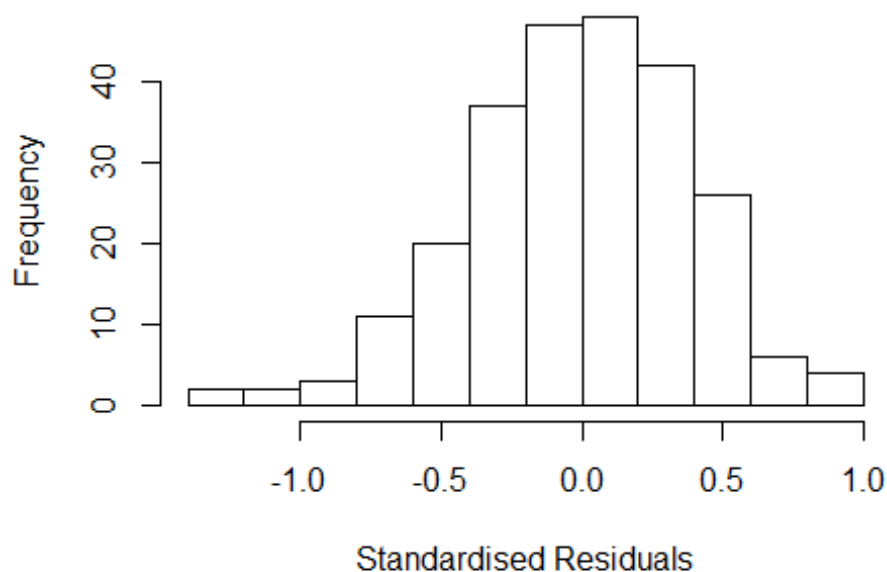
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 14, p-value = 0.2
## 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 | 1.506 | 1 | 1.227 |
| LastAuthorFemale | 1.417 | 1 | 1.190 |
| UniqueAuthors | 3.120 | 4 | 1.153 |
| Year | 4.399 | 16 | 1.047 |

Residuals from first and last author and team size



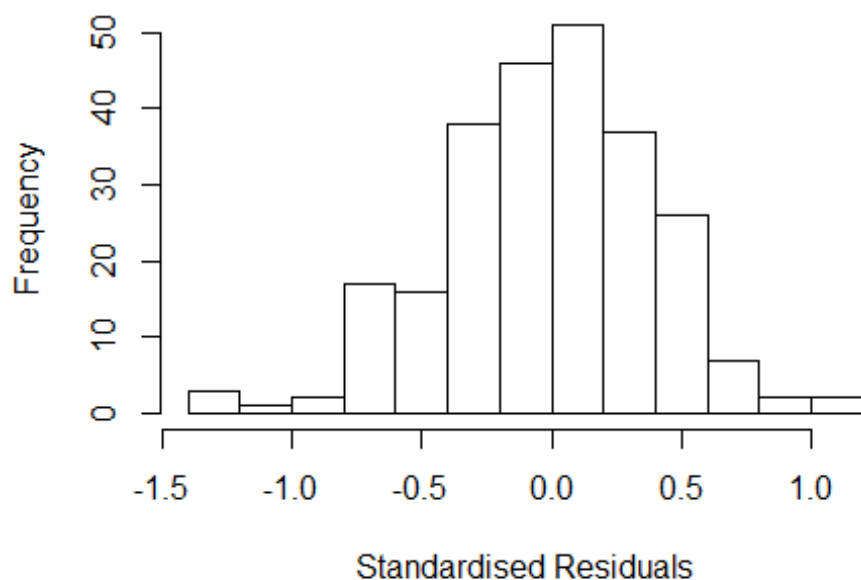
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3197 -0.2628  0.0068  0.2684  0.9907
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.399210   0.249184   5.62 5.8e-08 ***
## FirstAuthorFemale1 0.111277   0.062279   1.79  0.0753 .
## LastAuthorFemale1 0.000905   0.066932   0.01  0.9892
## UniqueAuthors2   -0.418835   0.136768  -3.06  0.0025 **
## UniqueAuthors3   -0.389359   0.133211  -2.92  0.0038 **
## UniqueAuthors4   -0.258337   0.134302  -1.92  0.0557 .
## UniqueAuthors5   -0.254880   0.128954  -1.98  0.0493 *
## Year1997         -0.176376   0.288084  -0.61  0.5410
## Year1998          0.271411   0.295347   0.92  0.3591
## Year1999          0.634074   0.245244   2.59  0.0104 *
```

```

## Year2000      0.358462    0.252598    1.42    0.1573
## Year2001     -0.028399    0.236068   -0.12    0.9044
## Year2002      0.254722    0.277309    0.92    0.3593
## Year2003      0.080700    0.242069    0.33    0.7392
## Year2004      0.103485    0.226527    0.46    0.6482
## Year2005      0.333170    0.240995    1.38    0.1682
## Year2006      0.329132    0.235173    1.40    0.1630
## Year2007      0.047213    0.243186    0.19    0.8462
## Year2008      0.334188    0.240027    1.39    0.1652
## Year2009      0.107372    0.243168    0.44    0.6592
## Year2010      0.098751    0.233295    0.42    0.6725
## Year2011      0.018558    0.232106    0.08    0.9363
## Year2012      0.338392    0.225985    1.50    0.1357
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.382
## Multiple R-squared:  0.188, Adjusted R-squared:  0.108
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 230 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.207  0.867  0.950  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      4.03e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.261 1      1.123
## LastAuthorFemale  1.311 1      1.145
## Year              1.653 16      1.016

```

Residuals from first and last author



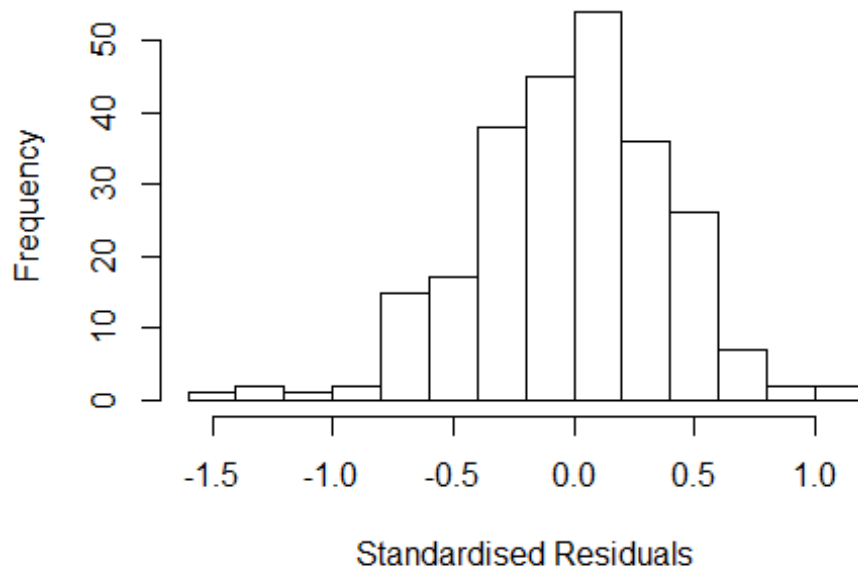
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3880 -0.2632  0.0163  0.2865  1.0732
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.03511    0.20247   5.11 6.7e-07 ***
## FirstAuthorFemale1  0.09530    0.05820   1.64  0.1029
## LastAuthorFemale1 -0.02153    0.06667  -0.32  0.7471
## Year1997        -0.05104    0.29044  -0.18  0.8607
## Year1998         0.29468    0.27899   1.06  0.2920
## Year1999         0.61366    0.23076   2.66  0.0084 **
## Year2000         0.36214    0.24779   1.46  0.1452
## Year2001        -0.00574    0.21996  -0.03  0.9792
## Year2002         0.34280    0.26274   1.30  0.1933
## Year2003         0.06676    0.23201   0.29  0.7738
## Year2004         0.13799    0.21503   0.64  0.5217
## Year2005         0.32902    0.22193   1.48  0.1396
```

```

## Year2006          0.33271      0.21986      1.51      0.1316
## Year2007          0.11677      0.22709      0.51      0.6076
## Year2008          0.44218      0.22124      2.00      0.0468 *
## Year2009          0.18176      0.22916      0.79      0.4285
## Year2010          0.13880      0.21880      0.63      0.5265
## Year2011          0.07335      0.21596      0.34      0.7344
## Year2012          0.37442      0.21214      1.77      0.0789 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.41
## Multiple R-squared:  0.141, Adjusted R-squared:  0.0736
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 234 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.228  0.888  0.953  0.910  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      4.03e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##      trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## 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
## -1.4069 -0.2602 0.0108 0.2914 1.0822
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.02931 0.19864 5.18 4.8e-07 ***
## FirstAuthorFemale1 0.09459 0.05804 1.63 0.1045
## Year1997 -0.04921 0.29061 -0.17 0.8657
## Year1998 0.29150 0.27737 1.05 0.2944
## Year1999 0.61954 0.22735 2.73 0.0069 **
## Year2000 0.36642 0.24550 1.49 0.1369
## Year2001 -0.00275 0.21765 -0.01 0.9899
## Year2002 0.34541 0.25972 1.33 0.1849
## Year2003 0.06689 0.23062 0.29 0.7721
## Year2004 0.14046 0.21324 0.66 0.5107
## Year2005 0.33074 0.22051 1.50 0.1350
## Year2006 0.33419 0.21800 1.53 0.1267
```

```

## Year2007          0.12159    0.22438    0.54    0.5884
## Year2008          0.44608    0.21839    2.04    0.0422 *
## Year2009          0.18520    0.22588    0.82    0.4131
## Year2010          0.14061    0.21683    0.65    0.5173
## Year2011          0.07365    0.21455    0.34    0.7317
## Year2012          0.37761    0.20976    1.80    0.0731 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.411
## Multiple R-squared:  0.141, Adjusted R-squared:  0.0773
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 235 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.218  0.891  0.954  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      4.03e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## 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
## -1.4257 -0.2688  0.0119  0.2896  1.0960

```

```

##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.03277    0.20174    5.12 6.5e-07 ***
## LastAuthorFemale1 -0.01602    0.06808   -0.24  0.8142
## Year1997        -0.00695    0.29996   -0.02  0.9815
## Year1998         0.32067    0.28499    1.13  0.2617
## Year1999         0.63379    0.22965    2.76  0.0062 **
## Year2000         0.40624    0.24427    1.66  0.0977 .
## Year2001         0.00881    0.22003    0.04  0.9681
## Year2002         0.36962    0.26095    1.42  0.1580
## Year2003         0.10579    0.22655    0.47  0.6410
## Year2004         0.18152    0.21267    0.85  0.3943
## Year2005         0.36412    0.22198    1.64  0.1023
## Year2006         0.36885    0.21615    1.71  0.0893 .
## Year2007         0.15759    0.22826    0.69  0.4907
## Year2008         0.46667    0.21916    2.13  0.0343 *
## Year2009         0.20208    0.22783    0.89  0.3760
## Year2010         0.15963    0.21825    0.73  0.4653
## Year2011         0.10028    0.21450    0.47  0.6406
## Year2012         0.40893    0.20903    1.96  0.0516 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.418
## Multiple R-squared:  0.131, Adjusted R-squared:  0.0671
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 228 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.220  0.885   0.954   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           4.03e-04           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01           5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500           50           2           1           1000           200
## trace.lev      mts      compute.rd
##           0           1000           0
##           psi           subsampling           cov
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 248"
## [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 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
## 16 9 1 12 12 14 22 13 16 25 23 23 26 32 21
## 2012
## 23
##
## 1996 1997 1998 1999 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
## 12 5 1 11 7 7 15 10 10 19 21 22 21 23 20
## 2012
## 21
##
## 1996 1997 1998 1999 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
## 12 5 1 10 7 7 15 10 9 16 20 21 20 22 20
## 2012
## 20
## [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: 1.19623119885132"

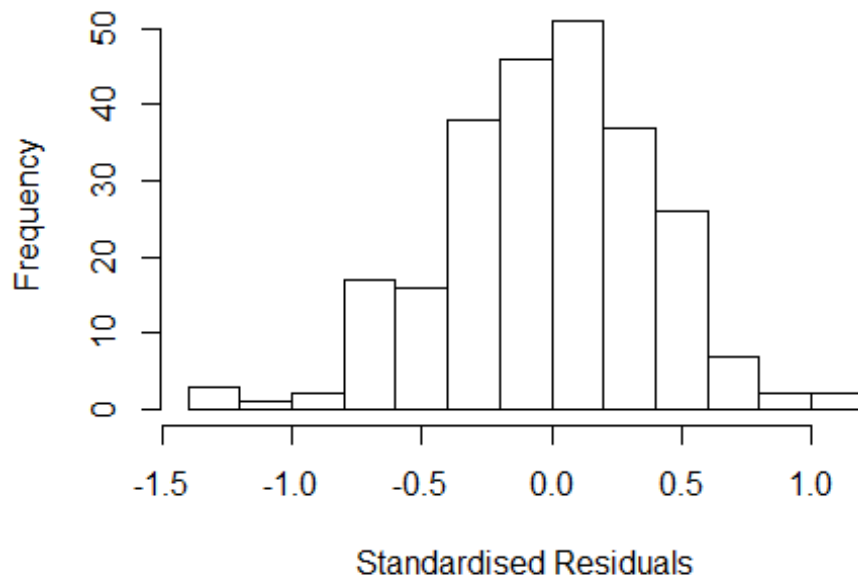
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## 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: 2"
## [1] "Male last author team size 2018 geometric mean: 1.12983096390975"

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

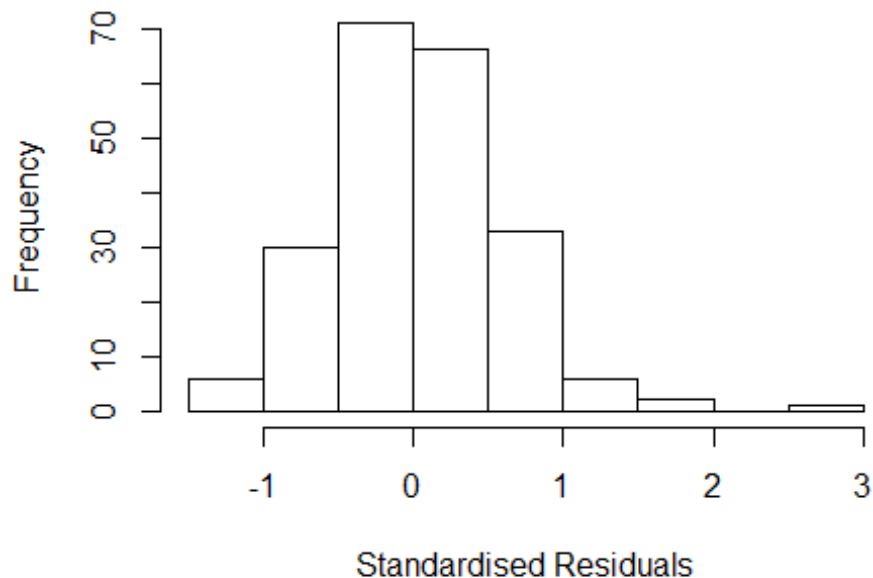
```


Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 16, 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 2.687  1      1.639
## LastAuthorFemale  2.722  1      1.650
## UniqueAuthors    3.101  3      1.208
## Year              4.094 15      1.048
```

Residuals from first and last author and team size



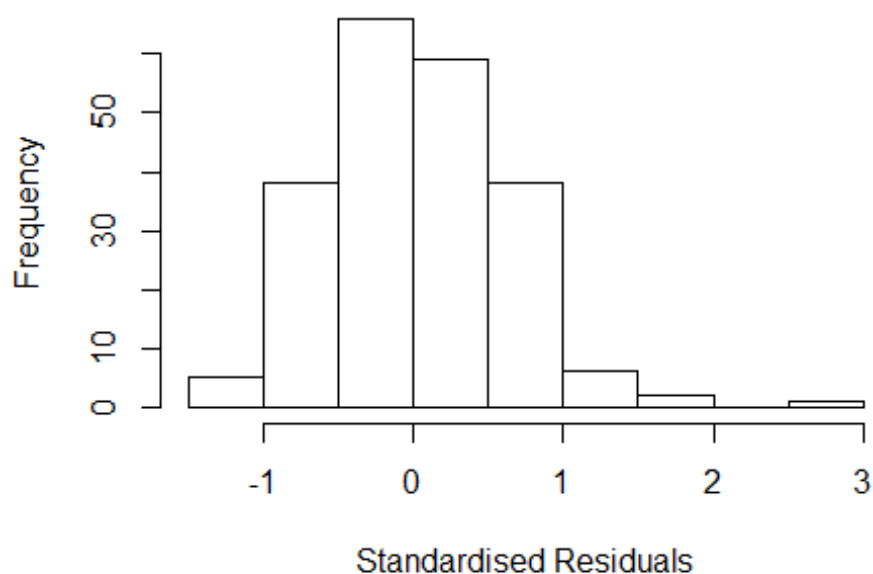
```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 163 33750526877 3.609 2006      2600      1      2.984
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.30397 -0.36025  0.00297  0.35620  2.98371
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1600    0.1551   7.48 2.6e-12 ***
## FirstAuthorFemale1 -0.2428    0.1422  -1.71 0.08941 .
## LastAuthorFemale1 -0.1415    0.1391  -1.02 0.31056
## UniqueAuthors2     0.2202    0.0997   2.21 0.02841 *
## UniqueAuthors3     0.6078    0.1316   4.62 7.0e-06 ***
## UniqueAuthors4     0.1348    0.1863   0.72 0.47028
## Year1997          -0.0491    0.4933  -0.10 0.92083
## Year1998           0.5900    0.1551   3.80 0.00019 ***
## Year1999          -0.2699    0.2826  -0.96 0.34068
## Year2001           0.1440    0.4361   0.33 0.74159
```

```

## Year2002          -0.2477      0.2005    -1.24    0.21824
## Year2003          -0.4250      0.2170    -1.96    0.05157 .
## Year2004          -0.6985      0.1900    -3.68    0.00031 ***
## Year2005          -0.6882      0.2062    -3.34    0.00101 **
## Year2006          -0.5347      0.2139    -2.50    0.01324 *
## Year2007          -0.4610      0.1947    -2.37    0.01887 *
## Year2008          -0.3952      0.2154    -1.83    0.06811 .
## Year2009          -0.1399      0.1993    -0.70    0.48332
## Year2010          -0.2866      0.1960    -1.46    0.14537
## Year2011          -0.2666      0.2006    -1.33    0.18540
## Year2012          -0.2404      0.2086    -1.15    0.25046
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.593
## Multiple R-squared:  0.186, Adjusted R-squared:  0.102
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## observation 79 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.
##   0.365  0.889  0.961  0.921  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          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 2.145  1          1.465
## LastAuthorFemale  2.283  1          1.511
## Year              1.717 15          1.018

```

Residuals from first and last author



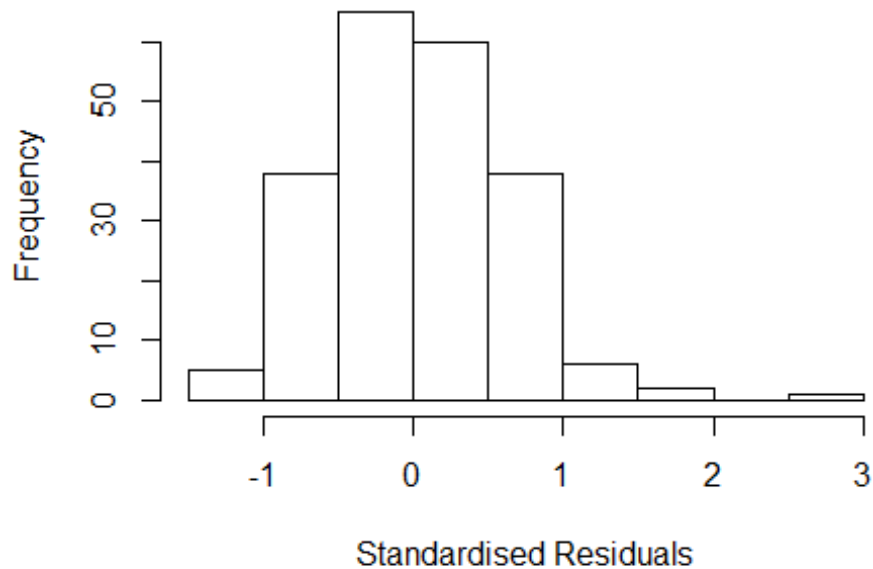
```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 163 33750526877 3.609 2006      2600      1      2.909
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.38170 -0.36913 -0.00158  0.40234  2.90853
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1491     0.1552   7.40 3.7e-12 ***
## FirstAuthorFemale1 -0.2658     0.1350  -1.97  0.05028 .
## LastAuthorFemale1  -0.0158     0.1330  -0.12  0.90569
## Year1997          -0.0380     0.4883  -0.08  0.93808
## Year1998           0.6009     0.1552   3.87  0.00015 ***
## Year1999          -0.1744     0.2726  -0.64  0.52312
## Year2001           0.2326     0.4202   0.55  0.58044
## Year2002          -0.1927     0.2090  -0.92  0.35757
## Year2003          -0.3595     0.2237  -1.61  0.10960
## Year2004          -0.6458     0.1837  -3.52  0.00054 ***
## Year2005          -0.5290     0.1992  -2.66  0.00856 **
## Year2006          -0.4486     0.2050  -2.19  0.02980 *
```

```

## Year2007          -0.4076      0.1978    -2.06  0.04064 *
## Year2008          -0.2475      0.2099    -1.18  0.23969
## Year2009           0.0154      0.2125     0.07  0.94234
## Year2010          -0.1817      0.2050    -0.89  0.37632
## Year2011          -0.1522      0.2030    -0.75  0.45427
## Year2012          -0.1573      0.2021    -0.78  0.43737
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.603
## Multiple R-squared:  0.128, Adjusted R-squared:  0.0533
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## observation 79 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.
##   0.422  0.878  0.953  0.916  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      4.65e-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 1.401 1          1.183
## Year              1.401 15          1.011

```

Residuals from first author



```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 163 33750526877 3.609 2006      2600      1      2.909
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q    Median        3Q        Max
## -1.39e+00 -3.68e-01 -9.99e-16  4.02e-01  2.91e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1486     0.1548   7.42 3.4e-12 ***
## FirstAuthorFemale1 -0.2747     0.1104  -2.49  0.01366 *
## Year1997         -0.0377     0.4950  -0.08  0.93932
## Year1998          0.6014     0.1548   3.89  0.00014 ***
## Year1999         -0.1770     0.2695  -0.66  0.51204
## Year2001          0.2385     0.4310   0.55  0.58063
## Year2002         -0.1929     0.2087  -0.92  0.35657
## Year2003         -0.3589     0.2243  -1.60  0.11124
## Year2004         -0.6475     0.1832  -3.53  0.00051 ***
## Year2005         -0.5276     0.1980  -2.67  0.00833 **
## Year2006         -0.4511     0.2051  -2.20  0.02899 *
## Year2007         -0.4095     0.1987  -2.06  0.04068 *
```

```

## Year2008          -0.2502      0.2099   -1.19  0.23472
## Year2009           0.0175      0.2125    0.08  0.93461
## Year2010          -0.1816      0.2053   -0.88  0.37751
## Year2011          -0.1538      0.2032   -0.76  0.45008
## Year2012          -0.1559      0.2022   -0.77  0.44155
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.59
## Multiple R-squared:  0.13,   Adjusted R-squared:  0.0596
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## observation 79 is an outlier with |weight| = 0 ( < 0.00047);
## 19 weights are ~= 1. The remaining 195 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.401  0.872  0.951  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      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 1.498 1          1.224
## Year            1.498 15          1.014

## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 163 33750526877 3.609 2006      2600      1      2.909
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q    Median      3Q      Max
## -1.34e+00 -3.57e-01 -8.88e-16  4.08e-01  2.92e+00

```

```

##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.13794    0.15564    7.31 6.4e-12 ***
## LastAuthorFemale1 -0.17356    0.11115   -1.56 0.12000
## Year1997        -0.02691    0.49007   -0.05 0.95626
## Year1998         0.61206    0.15564    3.93 0.00012 ***
## Year1999        -0.15673    0.27488   -0.57 0.56922
## Year2001         0.19769    0.39536    0.50 0.61762
## Year2002        -0.18175    0.20924   -0.87 0.38610
## Year2003        -0.35744    0.22436   -1.59 0.11272
## Year2004        -0.62885    0.18373   -3.42 0.00075 ***
## Year2005        -0.54539    0.20910   -2.61 0.00980 **
## Year2006        -0.44680    0.19983   -2.24 0.02647 *
## Year2007        -0.40435    0.19678   -2.05 0.04121 *
## Year2008        -0.25918    0.21842   -1.19 0.23680
## Year2009         0.00523    0.20993    0.02 0.98017
## Year2010        -0.21142    0.20685   -1.02 0.30797
## Year2011        -0.15939    0.20793   -0.77 0.44426
## Year2012        -0.19112    0.20603   -0.93 0.35472
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.6
## Multiple R-squared:  0.113, Adjusted R-squared:  0.0415
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## observation 79 is an outlier with |weight| = 0 ( < 0.00047);
## 21 weights are ~= 1. The remaining 193 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.412  0.868  0.951  0.914  0.985  0.999
## Algorithmic parameters:
##           tuning.chi           bb           tuning.psi           refine.tol
##           1.55e+00           5.00e-01           4.69e+00           1.00e-07
##           rel.tol           solve.tol           eps.outlier           eps.x
##           1.00e-07           1.00e-07           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 2601"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1999 2002 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    1    2    1    3    2    4    1    5    1    2    5
##
## 1999 2002 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    1    2    2    1    0    1    1    0    2
##
## 1999 2002 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    1    2    2    1    0    1    1    0    2
## [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: 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 = 3.5, 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: 2.05976714390712"

## Warning in wilcox.test.default(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"
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
## [1] "Regression 3: First author gender, Year as factors"

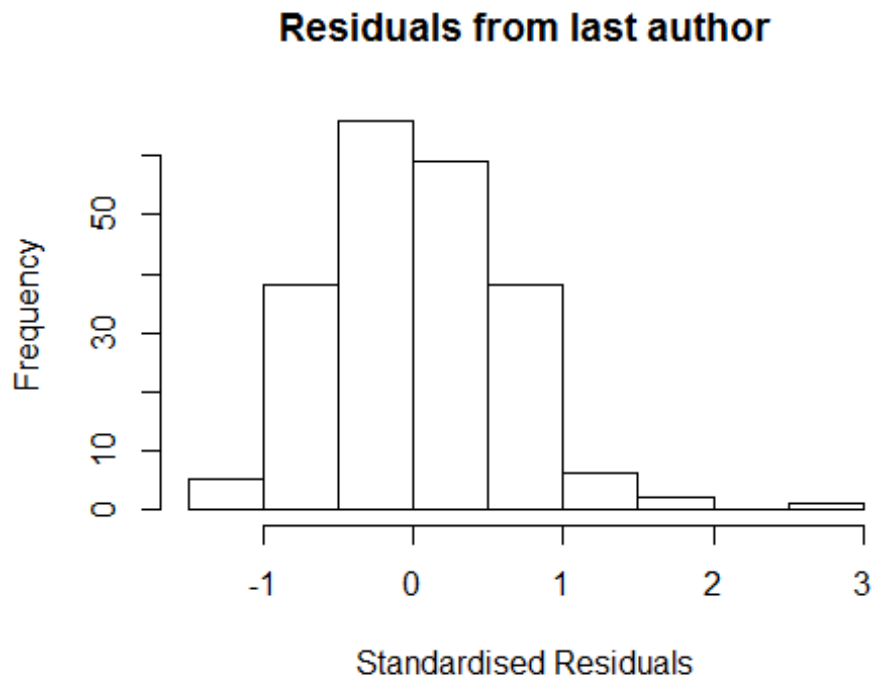
## Warning in lmrob.S(x, y, control = control, mf = mf): find_scale() did not
## converge in 'maxit.scale' (= 200) iterations

```

```
## [1] "Regression 4: Last author gender, Year as factors"

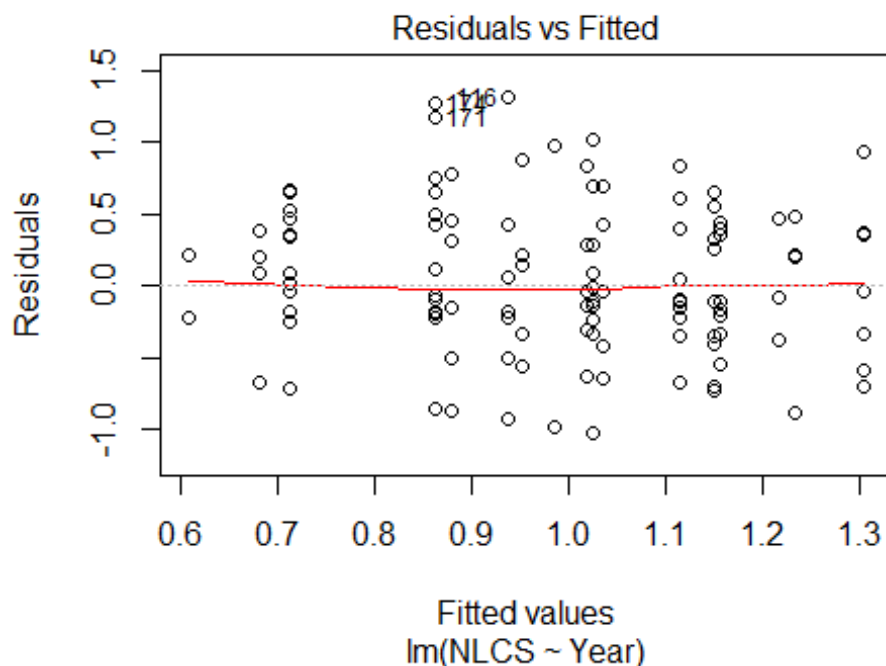
## Warning in lmrob.S(x, y, control = control, mf = mf): find_scale() did not
## converge in 'maxit.scale' (= 200) iterations

## Warning in lmrob.S(x, y, control = control, mf = mf): find_scale() did not
## converge in 'maxit.scale' (= 200) iterations
```



```
## [1] "Sample size for the above analysis: 12"
## [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    7    5    7    6    9    9    4    7    4   13   10    9   13   13
## 2011 2012
##   26   22
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    6    4    6    4    5    7    2    6    3   11   10    8   12   11
## 2011 2012
##   20   18
##
```

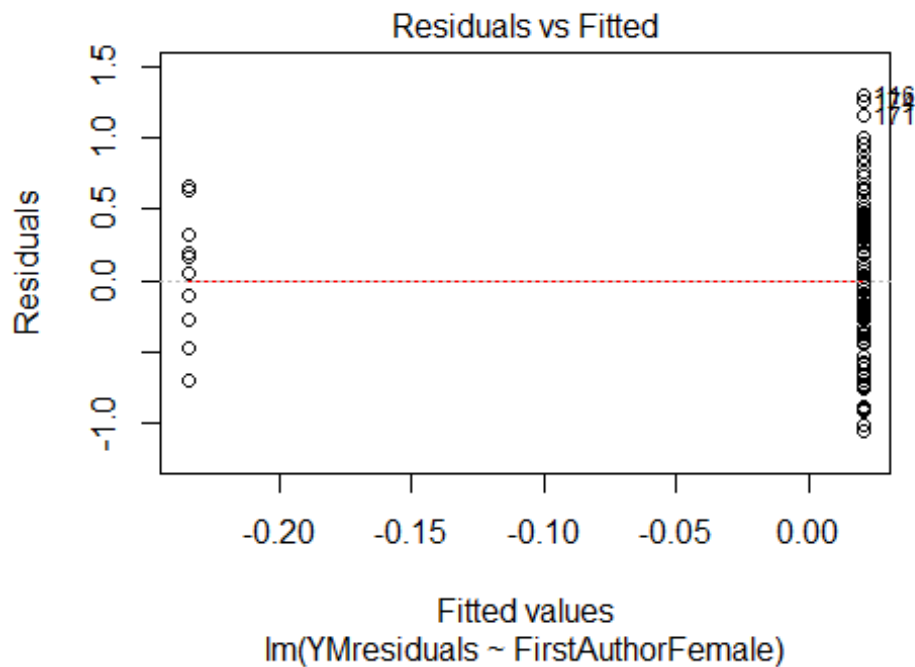
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      6      4      6      4      5      5      2      6      3     11     10      7     11     11
## 2011 2012
##     19     17
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 9.7, df = 16, p-value = 0.9
```



```
##
## 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: 1.41421356237309"
## [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 = 12, p-value = 1
## 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: 1.41421356237309"

## 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 = 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 1.808  1      1.345
## LastAuthorFemale  2.691  1      1.640
## UniqueAuthors    2.811  3      1.188
## Year              4.198 16      1.046

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
```

```

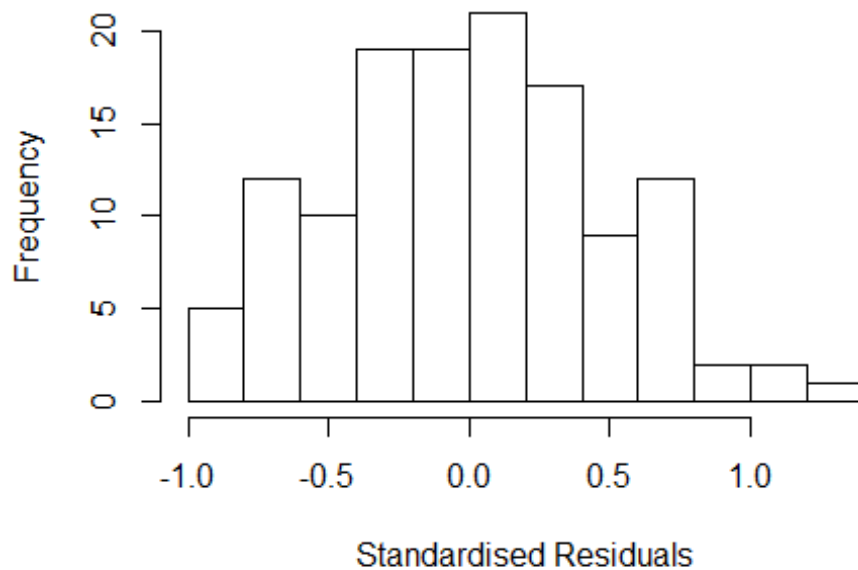
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.9850 -0.3147 -0.0138  0.3305  1.2426
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.9850     2.2224   0.44    0.659
## FirstAuthorFemale1 -0.3738     0.1755  -2.13    0.035 *
## LastAuthorFemale1  0.2056     0.1370   1.50    0.136
## UniqueAuthors2     0.2146     0.1195   1.80    0.075 .
## UniqueAuthors3     0.1840     0.2440   0.75    0.453
## UniqueAuthors4     0.8266     0.7968   1.04    0.302
## Year1997          -0.0910     2.2397  -0.04    0.968
## Year1998           0.1183     2.2367   0.05    0.958
## Year1999          -0.0341     2.2275  -0.02    0.988
## Year2000          -0.2938     2.2335  -0.13    0.896
## Year2001           0.0884     2.2393   0.04    0.969
## Year2002           0.0130     2.2275   0.01    0.995
## Year2003          -0.3785     2.2280  -0.17    0.865
## Year2004          -0.0543     2.2336  -0.02    0.981
## Year2005           0.1589     2.2281   0.07    0.943
## Year2006          -0.2007     2.2288  -0.09    0.928
## Year2007           0.0108     2.2261   0.00    0.996
## Year2008          -0.2865     2.2283  -0.13    0.898
## Year2009          -0.1395     2.2292  -0.06    0.950
## Year2010           0.1216     2.2268   0.05    0.957
## Year2011          -0.2992     2.2279  -0.13    0.893
## Year2012          -0.3185     2.2274  -0.14    0.887
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.549
## Multiple R-squared:  0.18, Adjusted R-squared:  0.0191
## Convergence in 44 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 122 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.588  0.871  0.959   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          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"

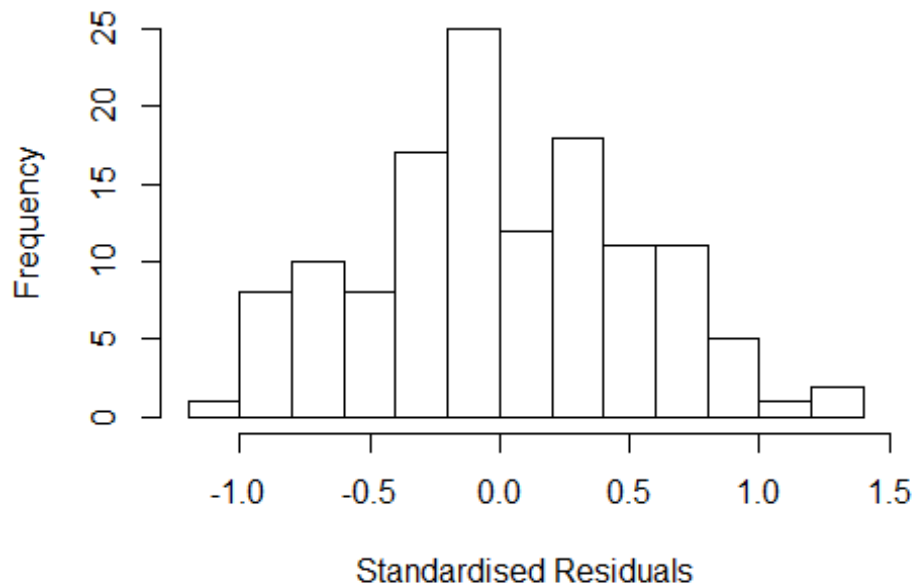
## Warning in lmrob.fit(x, y, control, init = init, mf = mf): M-step did NOT
## converge. Returning unconverged SM-estimate
```

Residuals from first and last author and team size



```
## [1] "Regression 3: First author gender, Year as factors"
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.342 1          1.158
## Year              1.342 16          1.009
```

Residuals from first author



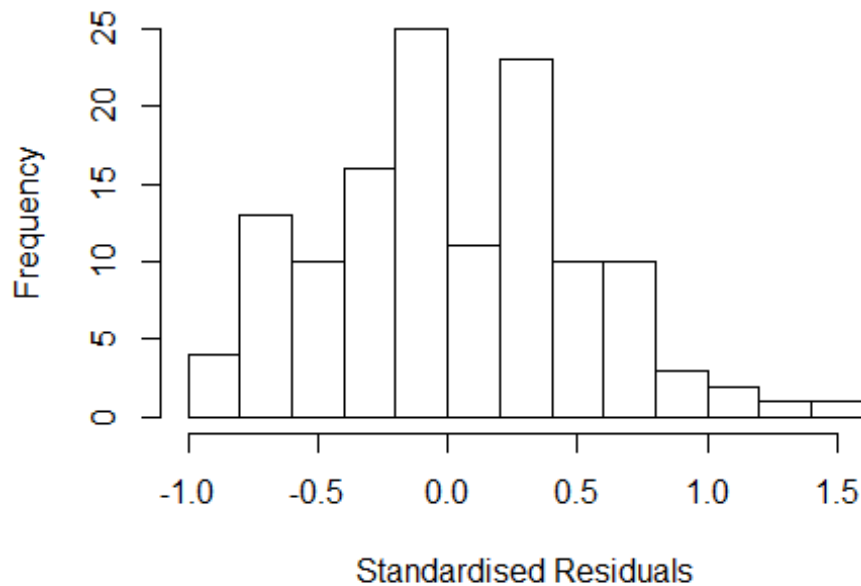
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0166 -0.3200 -0.0393 0.3316 1.3307
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9850 2.0677 0.48 0.63
## FirstAuthorFemale1 -0.2558 0.1622 -1.58 0.12
## Year1997 -0.0915 2.0861 -0.04 0.97
## Year1998 0.3000 2.0885 0.14 0.89
## Year1999 0.0163 2.0766 0.01 0.99
## Year2000 -0.2109 2.0822 -0.10 0.92
## Year2001 0.1103 2.0860 0.05 0.96
## Year2002 0.0490 2.0754 0.02 0.98
## Year2003 -0.3785 2.0738 -0.18 0.86
## Year2004 -0.0085 2.0767 0.00 1.00
## Year2005 0.2258 2.0789 0.11 0.91
## Year2006 0.0316 2.0730 0.02 0.99
```

```

## Year2007          0.1924      2.0714      0.09      0.93
## Year2008         -0.0596      2.0760     -0.03      0.98
## Year2009          0.0782      2.0724      0.04      0.97
## Year2010          0.1932      2.0728      0.09      0.93
## Year2011         -0.1727      2.0735     -0.08      0.93
## Year2012         -0.2294      2.0718     -0.11      0.91
##
## Robust residual standard error: 0.558
## Multiple R-squared:  0.111, Adjusted R-squared:  -0.0255
## Convergence in 42 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 120 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.548  0.886  0.960  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      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 1.812 1      1.346
## Year            1.812 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
## -0.9919 -0.3200 -0.0338 0.3411 1.4898
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.98500 2.06462 0.48 0.63
## LastAuthorFemale1 0.13647 0.12363 1.10 0.27
## Year1997 -0.09154 2.08299 -0.04 0.97
## Year1998 0.26199 2.08284 0.13 0.90
## Year1999 0.00304 2.07204 0.00 1.00
## Year2000 -0.32371 2.07646 -0.16 0.88
## Year2001 0.01098 2.08098 0.01 1.00
## Year2002 0.04899 2.07233 0.02 0.98
## Year2003 -0.37850 2.07067 -0.18 0.86
## Year2004 -0.08322 2.07477 -0.04 0.97
## Year2005 0.22581 2.07585 0.11 0.91
## Year2006 0.00693 2.06941 0.00 1.00
```

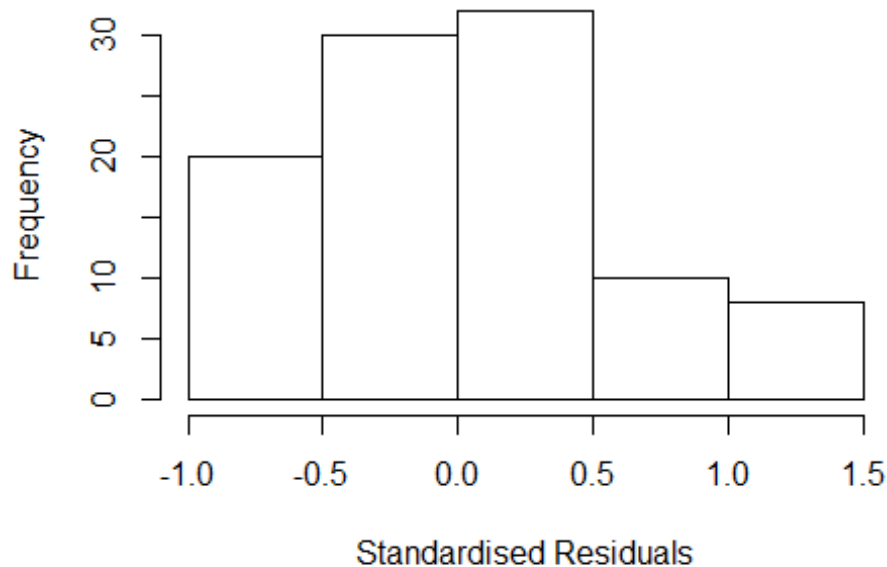
```

## Year2007      0.11587      2.06693      0.06      0.96
## Year2008     -0.22282      2.07833     -0.11      0.91
## Year2009      0.01876      2.07042      0.01      0.99
## Year2010      0.15889      2.07035      0.08      0.94
## Year2011     -0.23123      2.06894     -0.11      0.91
## Year2012     -0.26512      2.06934     -0.13      0.90
##
## Robust residual standard error: 0.558
## Multiple R-squared: 0.105, Adjusted R-squared: -0.0318
## Convergence in 42 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 119 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.455  0.877  0.958  0.921  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      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 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
##    9    8    5    6    3    8    5    4    6    3    6    10   11   17   10
## 2011 2012
##   14   13
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7    7    4    5    2    1    4    1    4    3    2   10    8   15    9
## 2011 2012
##   13    7
##

```

```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7    7    4    5    2    1    4    1    4    3    2   10    8   14    9
## 2011 2012
##   12    7
## [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: 1.317980629213"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.840e+00  1      1.357e+00
## LastAuthorFemale  2.838e+00  1      1.685e+00
## UniqueAuthors    9.970e+13  1      9.985e+06
## Year              3.888e+14 16      2.857e+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"
```

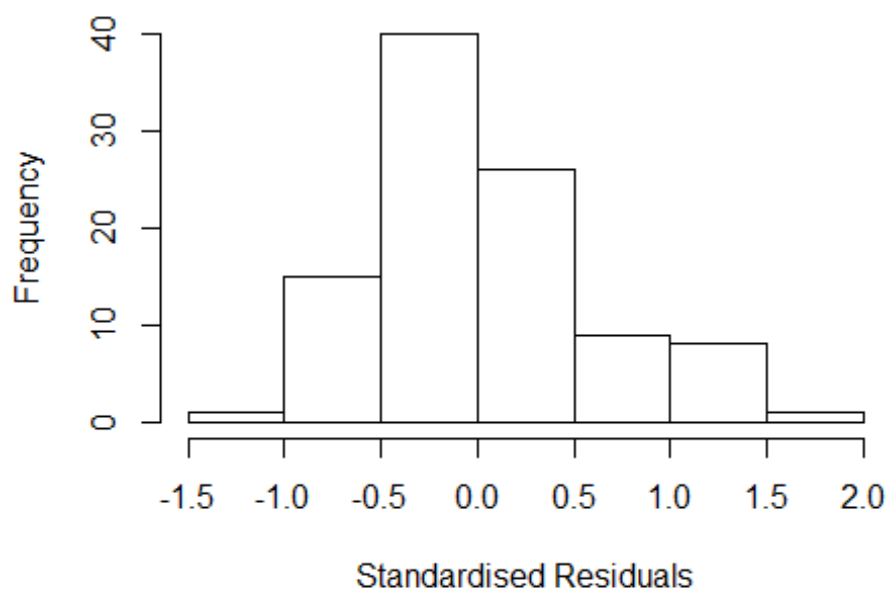
```

## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.93240 -0.38454  0.00775  0.32355  1.45738
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9324    0.2948   3.16 0.00221 **
## FirstAuthorFemale1 -0.3205    0.2373  -1.35 0.18060
## LastAuthorFemale1 -0.5399    0.2200  -2.45 0.01629 *
## UniqueAuthors2    0.5773    0.1679   3.44 0.00093 ***
## Year1997         -0.0950    0.4876  -0.19 0.84599
## Year1998         -0.2665    0.4226  -0.63 0.53005
## Year1999         -0.0670    0.3398  -0.20 0.84410
## Year2000          0.3830    0.2921   1.31 0.19358
## Year2001          0.2663    0.2979   0.89 0.37405
## Year2002          0.0264    0.3596   0.07 0.94175
## Year2003         -0.1084    0.2948  -0.37 0.71412
## Year2004         -0.0605    0.3380  -0.18 0.85835
## Year2005         -0.0124    0.6821  -0.02 0.98551
## Year2006         -0.1369    0.2951  -0.46 0.64393
## Year2007         -0.4400    0.3114  -1.41 0.16161
## Year2008          0.1608    0.4198   0.38 0.70271
## Year2009         -0.2932    0.3239  -0.91 0.36809
## Year2010         -0.2853    0.4109  -0.69 0.48948
## Year2011         -0.5479    0.3238  -1.69 0.09453 .
## Year2012         -0.9748    0.3513  -2.77 0.00688 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.575
## Multiple R-squared:  0.319, Adjusted R-squared:  0.157
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 83 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.500  0.888  0.940  0.908  0.984  0.999
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          1.00e-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 3.636 1      1.907
## LastAuthorFemale  5.822 1      2.413
## Year              2.975 16      1.035
```

Residuals from first and last author



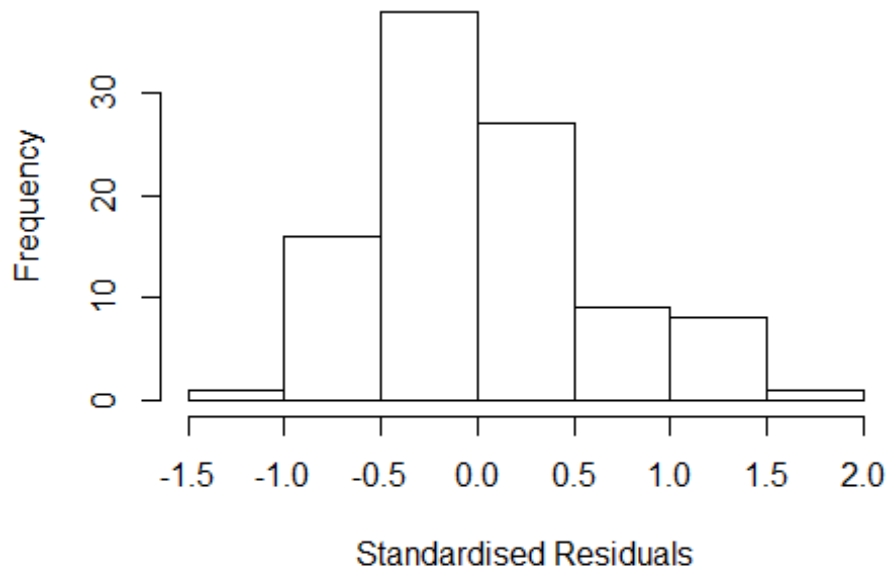
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.0479 -0.4259 -0.0425  0.3937  1.5294
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.0479     0.3141   3.34  0.0013 **
## FirstAuthorFemale1  0.0117     0.2677   0.04  0.9653
```

```

## LastAuthorFemale1 -0.2194 0.2443 -0.90 0.3718
## Year1997 -0.1833 0.5876 -0.31 0.7559
## Year1998 -0.3348 0.4735 -0.71 0.4816
## Year1999 -0.0897 0.3901 -0.23 0.8187
## Year2000 0.5561 0.3439 1.62 0.1097
## Year2001 0.7281 0.3141 2.32 0.0230 *
## Year2002 -0.0216 0.3929 -0.06 0.9562
## Year2003 -0.2239 0.3141 -0.71 0.4781
## Year2004 -0.1095 0.3703 -0.30 0.7683
## Year2005 0.2500 0.5345 0.47 0.6412
## Year2006 -0.2524 0.3143 -0.80 0.4243
## Year2007 -0.4443 0.3306 -1.34 0.1828
## Year2008 0.1438 0.4472 0.32 0.7487
## Year2009 -0.2965 0.3690 -0.80 0.4240
## Year2010 -0.3951 0.4279 -0.92 0.3585
## Year2011 -0.5566 0.3690 -1.51 0.1353
## Year2012 -0.8058 0.3699 -2.18 0.0323 *
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.588
## Multiple R-squared: 0.213, Adjusted R-squared: 0.038
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 90 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.479 0.883 0.947 0.908 0.985 0.999
## Algorithmic parameters:
## tuning.chi bb tuning.psi refine.tol
## 1.55e+00 5.00e-01 4.69e+00 1.00e-07
## rel.tol solve.tol eps.outlier eps.x
## 1.00e-07 1.00e-07 1.00e-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 0.3062 1 0.5534
## Year 0.3062 16 0.9637

```

Residuals from first author



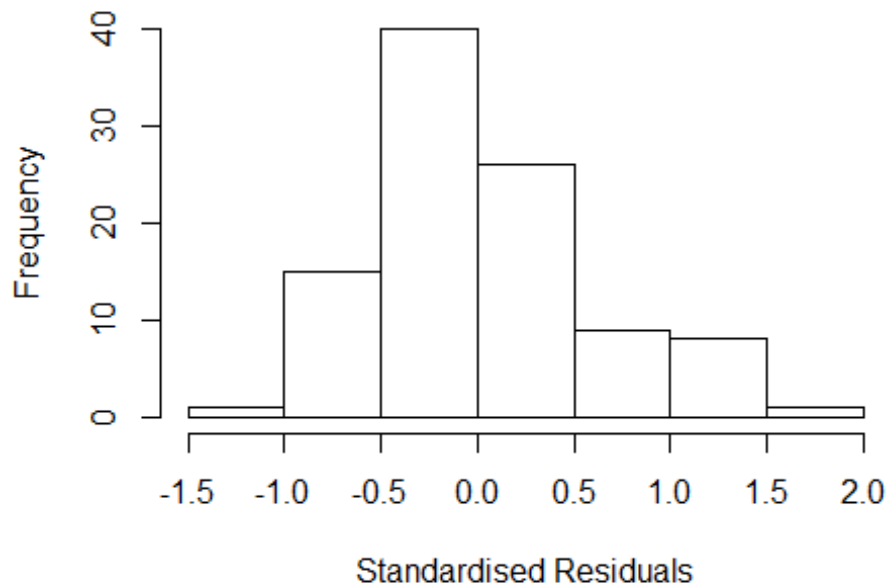
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.048 -0.409 -0.032 0.393 1.584
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0476 0.3133 3.34 0.0012 **
## FirstAuthorFemale1 0.0107 0.2775 0.04 0.9694
## Year1997 -0.1814 0.5849 -0.31 0.7573
## Year1998 -0.3342 0.4745 -0.70 0.4832
## Year1999 -0.0890 0.3906 -0.23 0.8202
## Year2000 0.5564 0.3431 1.62 0.1087
## Year2001 0.7284 0.3133 2.33 0.0225 *
## Year2002 -0.0211 0.3931 -0.05 0.9572
## Year2003 -0.2236 0.3133 -0.71 0.4773
## Year2004 -0.1663 0.3510 -0.47 0.6370
## Year2005 0.2505 0.5332 0.47 0.6398
## Year2006 -0.2521 0.3134 -0.80 0.4235
```

```

## Year2007          -0.4441      0.3298   -1.35    0.1819
## Year2008           0.0894      0.4671    0.19    0.8487
## Year2009          -0.3526      0.3480   -1.01    0.3138
## Year2010          -0.3937      0.4269   -0.92    0.3591
## Year2011          -0.5557      0.3682   -1.51    0.1350
## Year2012          -0.8049      0.3692   -2.18    0.0321 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.592
## Multiple R-squared:  0.208, Adjusted R-squared:  0.044
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 88 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.453  0.877  0.948  0.906  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.00e-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 2.077 1          1.441
## Year             2.077 16          1.023

```


Residuals from last author



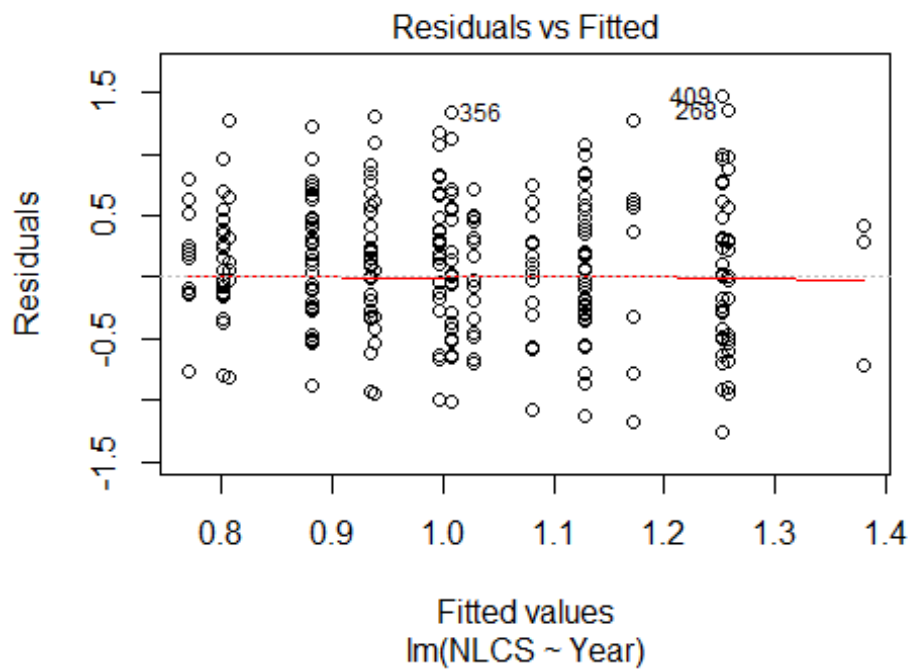
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0468 -0.4335 -0.0428 0.3886 1.5189
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0468 0.3100 3.38 0.0011 **
## LastAuthorFemale1 -0.2236 0.2412 -0.93 0.3565
## Year1997 -0.1742 0.5734 -0.30 0.7620
## Year1998 -0.3309 0.4795 -0.69 0.4920
## Year1999 -0.0839 0.3825 -0.22 0.8269
## Year2000 0.5572 0.3401 1.64 0.1052
## Year2001 0.7292 0.3100 2.35 0.0211 *
## Year2002 -0.0174 0.3779 -0.05 0.9634
## Year2003 -0.2228 0.3100 -0.72 0.4744
## Year2004 -0.1078 0.3664 -0.29 0.7693
## Year2005 0.2542 0.5312 0.48 0.6335
## Year2006 -0.2513 0.3102 -0.81 0.4203
```

```

## Year2007          -0.4432      0.3268    -1.36    0.1787
## Year2008           0.1553      0.4383     0.35    0.7241
## Year2009          -0.2917      0.3611    -0.81    0.4215
## Year2010          -0.3884      0.4224    -0.92    0.3605
## Year2011          -0.5525      0.3645    -1.52    0.1334
## Year2012          -0.8016      0.3662    -2.19    0.0314 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.604
## Multiple R-squared:  0.211, Adjusted R-squared:  0.0471
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 91 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.507  0.890  0.951  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      1.00e-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 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
##   24   25   26    5   25   28   18   21   28   29   36   38   56   38   54
## 2012
##   44
##
## 1996 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##   13   10   14    3    9   15    9   14   17   18   24   26   43   31   44
## 2012

```

```
## 27
##
## 1996 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
## 13 9 12 3 7 12 8 12 14 17 22 24 41 30 40
## 2012
## 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 = 15, df = 15, p-value = 0.5
```



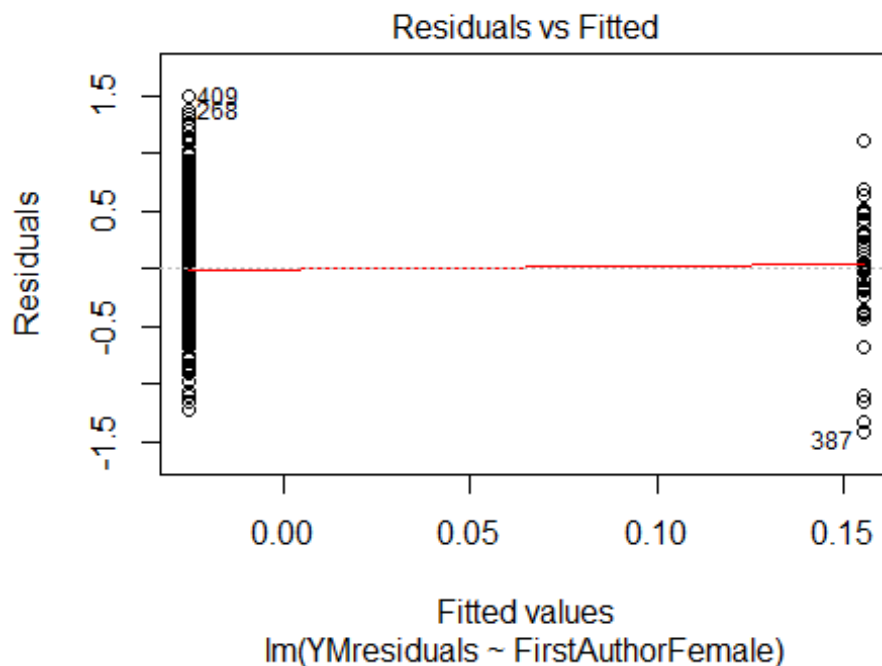
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.4, df = 1, p-value = 0.2

## [1] "Female first author team size 2018 geometric mean: 2.49146187923103"
## [1] "Male first author team size 2018 geometric mean: 1.58775960786737"

## Warning in wilcox.test.default(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 = 0.09
## 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: 1.62142906467649"

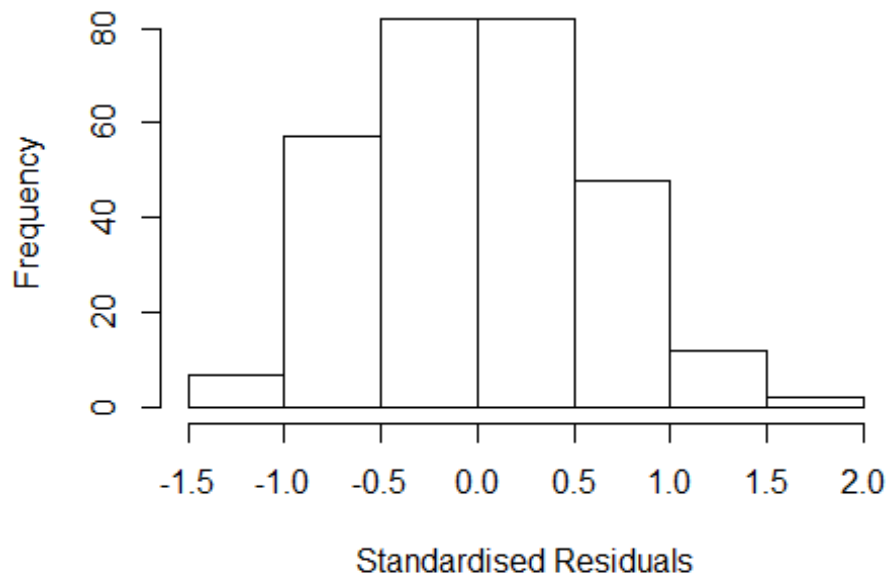
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 54, p-value = 0.07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
## Year as factors"
##
```

| | GVIF | Df | GVIF^(1/(2*Df)) |
|-------------------|-------|----|-----------------|
| FirstAuthorFemale | 1.564 | 1 | 1.251 |
| LastAuthorFemale | 1.340 | 1 | 1.158 |
| UniqueAuthors | 2.537 | 4 | 1.123 |
| Year | 3.219 | 15 | 1.040 |

Residuals from first and last author and team size



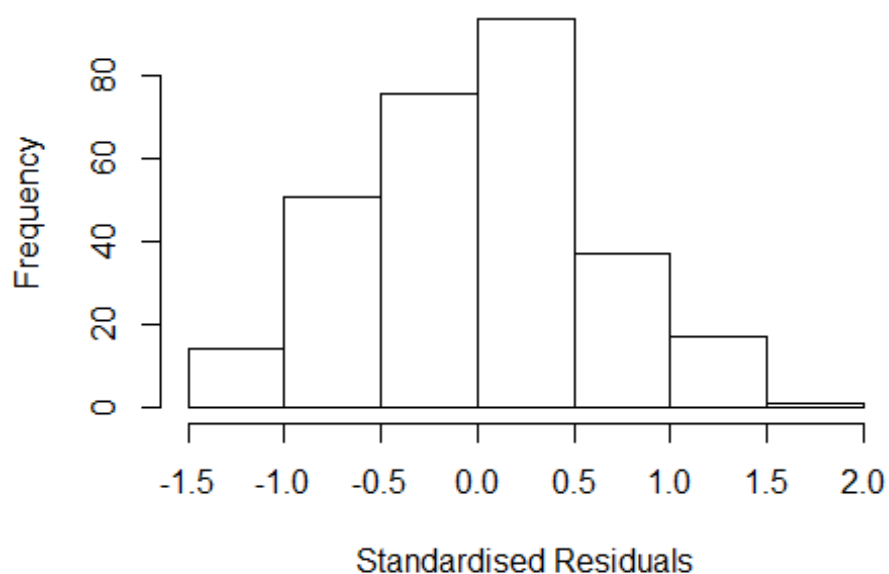
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3466 -0.4126 -0.0128 0.3988 1.6640
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9877 0.2458 4.02 7.6e-05 ***
## FirstAuthorFemale1 0.0898 0.1001 0.90 0.3705
## LastAuthorFemale1 0.1055 0.1165 0.91 0.3660
## UniqueAuthors2 0.2782 0.0880 3.16 0.0017 **
## UniqueAuthors3 0.2145 0.1178 1.82 0.0698 .
## UniqueAuthors4 0.5624 0.1258 4.47 1.1e-05 ***
## UniqueAuthors5 0.6630 0.2107 3.15 0.0018 **
## Year1997 -0.0769 0.3647 -0.21 0.8332
## Year1998 -0.2589 0.2867 -0.90 0.3672
## Year2000 0.2815 0.4342 0.65 0.5173
```

```

## Year2001          -0.0933      0.5009   -0.19   0.8524
## Year2002          -0.0708      0.2766   -0.26   0.7983
## Year2003          -0.4322      0.3176   -1.36   0.1748
## Year2004          -0.1834      0.3106   -0.59   0.5553
## Year2005          -0.0814      0.2970   -0.27   0.7843
## Year2006          -0.0417      0.3013   -0.14   0.8901
## Year2007          -0.2030      0.2732   -0.74   0.4582
## Year2008          -0.0644      0.2894   -0.22   0.8240
## Year2009          -0.0310      0.2701   -0.11   0.9086
## Year2010          -0.3285      0.2730   -1.20   0.2300
## Year2011          -0.3377      0.2621   -1.29   0.1988
## Year2012          -0.4095      0.2619   -1.56   0.1191
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.625
## Multiple R-squared:  0.148, Adjusted R-squared:  0.0809
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 27 weights are ~= 1. The remaining 263 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.459  0.894  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      3.45e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.433 1 1.197
## LastAuthorFemale 1.437 1 1.199
## Year 1.590 15 1.016

```

Residuals from first and last author



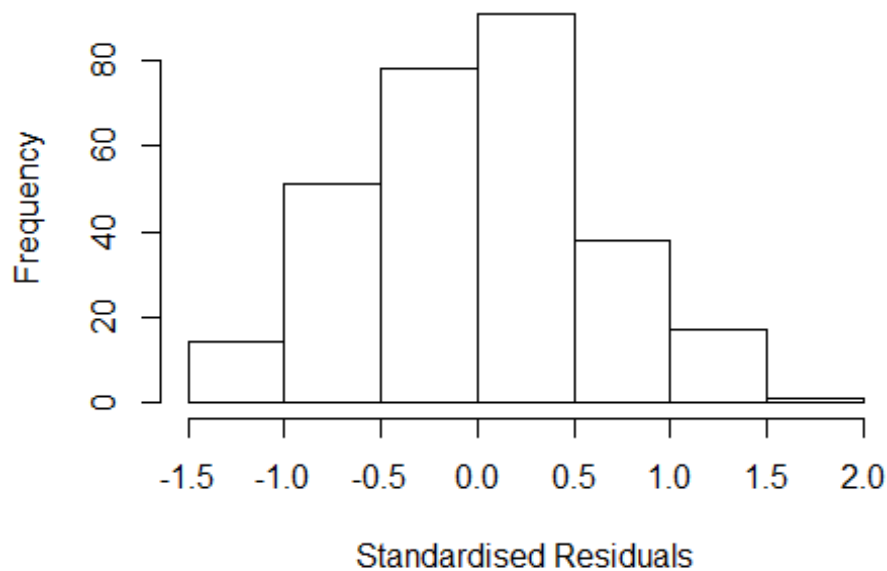
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.458 -0.442 0.012 0.415 1.564
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0156 0.2485 4.09 5.8e-05 ***
## FirstAuthorFemale1 0.2286 0.1047 2.18 0.03 *
## LastAuthorFemale1 0.0723 0.1218 0.59 0.55
## Year1997 -0.0543 0.3964 -0.14 0.89
## Year1998 -0.2596 0.2927 -0.89 0.38
## Year2000 0.3851 0.4035 0.95 0.34
## Year2001 -0.0640 0.5400 -0.12 0.91
## Year2002 0.0257 0.2815 0.09 0.93
## Year2003 -0.3364 0.3534 -0.95 0.34
## Year2004 0.0827 0.2979 0.28 0.78
## Year2005 0.0345 0.3015 0.11 0.91
## Year2006 0.1093 0.2958 0.37 0.71
```

```

## Year2007          -0.0785      0.2824   -0.28      0.78
## Year2008           0.1412      0.2887    0.49      0.63
## Year2009           0.0930      0.2705    0.34      0.73
## Year2010          -0.1444      0.2763   -0.52      0.60
## Year2011          -0.1945      0.2678   -0.73      0.47
## Year2012          -0.2689      0.2636   -1.02      0.31
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.647
## Multiple R-squared:  0.0691, Adjusted R-squared:  0.0109
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 270 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.538  0.879   0.956   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      3.45e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.257 1          1.121
## Year              1.257 15          1.008

```


Residuals from first author



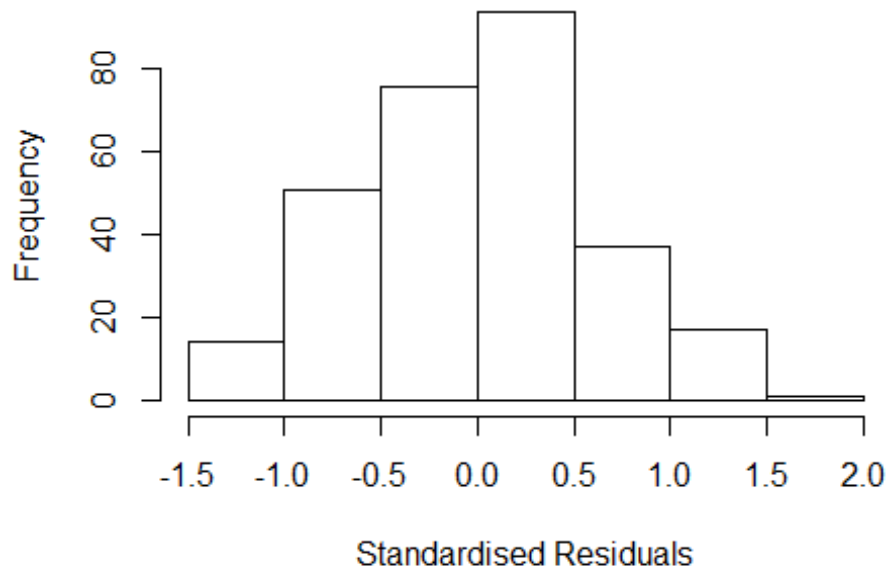
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.4148 -0.4488  0.0146  0.4119  1.5611
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0156     0.2486   4.09 5.8e-05 ***
## FirstAuthorFemale1 0.2549     0.0978   2.61 0.0096 **
## Year1997        -0.0544     0.3964  -0.14 0.8910
## Year1998        -0.2571     0.2919  -0.88 0.3791
## Year2000         0.3851     0.4036   0.95 0.3409
## Year2001        -0.0663     0.5422  -0.12 0.9028
## Year2002         0.0232     0.2815   0.08 0.9343
## Year2003        -0.3365     0.3535  -0.95 0.3420
## Year2004         0.0966     0.2941   0.33 0.7428
## Year2005         0.0362     0.3019   0.12 0.9047
## Year2006         0.1161     0.2946   0.39 0.6937
## Year2007        -0.0673     0.2811  -0.24 0.8110
```

```

## Year2008          0.1443      0.2894      0.50      0.6185
## Year2009          0.0982      0.2701      0.36      0.7165
## Year2010         -0.1435      0.2762     -0.52      0.6040
## Year2011         -0.1906      0.2679     -0.71      0.4776
## Year2012         -0.2681      0.2639     -1.02      0.3105
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.646
## Multiple R-squared:  0.0681, Adjusted R-squared:  0.0134
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 269 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.539  0.878  0.955  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      3.45e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##  nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##  trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.212  1          1.101
## Year            1.212 15          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
## -1.3382 -0.4532 0.0241 0.4225 1.5588
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0152 0.2476 4.10 5.4e-05 ***
## LastAuthorFemale1 0.1760 0.1112 1.58 0.11
## Year1997 -0.0531 0.3949 -0.13 0.89
## Year1998 -0.2310 0.2956 -0.78 0.44
## Year2000 0.3851 0.4026 0.96 0.34
## Year2001 -0.0438 0.5367 -0.08 0.94
## Year2002 0.0474 0.2812 0.17 0.87
## Year2003 -0.3352 0.3526 -0.95 0.34
## Year2004 0.0634 0.2992 0.21 0.83
## Year2005 0.0662 0.3009 0.22 0.83
## Year2006 0.1292 0.2947 0.44 0.66
## Year2007 -0.0643 0.2829 -0.23 0.82
```

```

## Year2008          0.1470      0.2860      0.51      0.61
## Year2009          0.1079      0.2698      0.40      0.69
## Year2010         -0.0863      0.2740     -0.31      0.75
## Year2011         -0.1733      0.2658     -0.65      0.51
## Year2012         -0.2394      0.2626     -0.91      0.36
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.652
## Multiple R-squared:  0.058, Adjusted R-squared:  0.00275
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 272 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.547  0.875  0.954  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      3.45e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##   nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##   trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 290"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 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
##    2    9    8    4    4    6    3    8    5    7    6    6    7   12    8
## 2011 2012
##   17   13
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    1    3    0    1    1    2    4    1    3    4    3    6    8    5
## 2011 2012
##   13    9

```

```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    1    3    0    1    1    2    4    1    1    4    3    5    7    4
## 2011 2012
##   12    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: 1.5157165665104"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 9.124e+13  1      9.552e+06
## LastAuthorFemale  2.189e+00  1      1.480e+00
## UniqueAuthors    9.733e+15  4      9.966e+01
## Year              6.000e+29 14      1.157e+01

## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##      Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##      k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q      Median        3Q       Max
## -1.57e+00 -2.37e-01 -9.99e-16  2.81e-01  8.51e-01
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.19e+00   3.16e-08  3.76e+07 < 2e-16 ***
## FirstAuthorFemale1 2.26e-02   1.82e-01  1.20e-01  0.90195
## LastAuthorFemale1 1.94e-01   1.52e-01  1.28e+00  0.21006
## UniqueAuthors2    1.73e-01   2.93e-01  5.90e-01  0.55741
## UniqueAuthors3    3.15e-01   2.63e-01  1.20e+00  0.23808
## UniqueAuthors4    3.37e-01   3.48e-01  9.70e-01  0.33972
## UniqueAuthors5    3.53e-01   4.40e-01  8.00e-01  0.42719
## Year1998          -3.00e-01   2.80e-01 -1.07e+00  0.29057
## Year2000           -8.38e-01   2.63e-01 -3.19e+00  0.00295 **
## Year2001           -5.85e-01   2.87e-01 -2.04e+00  0.04895 *
## Year2002           -1.13e+00   3.15e-01 -3.58e+00  0.00099 ***
## Year2003           -7.02e-01   3.76e-01 -1.87e+00  0.06971 .
## Year2004           -2.52e-01   2.63e-01 -9.60e-01  0.34347
## Year2005           -1.06e-01   2.63e-01 -4.00e-01  0.68835
## Year2006           -9.42e-02   2.64e-01 -3.60e-01  0.72284
## Year2007           -1.44e-01   3.68e-01 -3.90e-01  0.69738
```

```

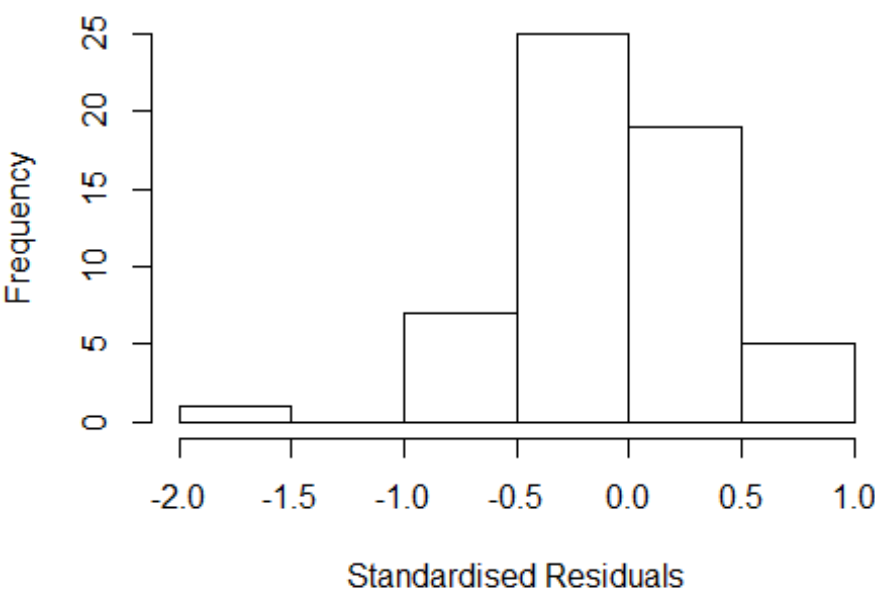
## Year2008          -5.82e-01    2.41e-01 -2.42e+00    0.02090 *
## Year2009          -3.16e-01    3.98e-01 -7.90e-01    0.43272
## Year2010           3.79e-01    5.32e-01  7.10e-01    0.48111
## Year2011          -4.91e-01    2.83e-01 -1.74e+00    0.09046 .
## Year2012          -2.29e-01    2.93e-01 -7.80e-01    0.44020
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.488
## Multiple R-squared:  0.329, Adjusted R-squared:  -0.043
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 48 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.279  0.899  0.967   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      1.75e-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"

## 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

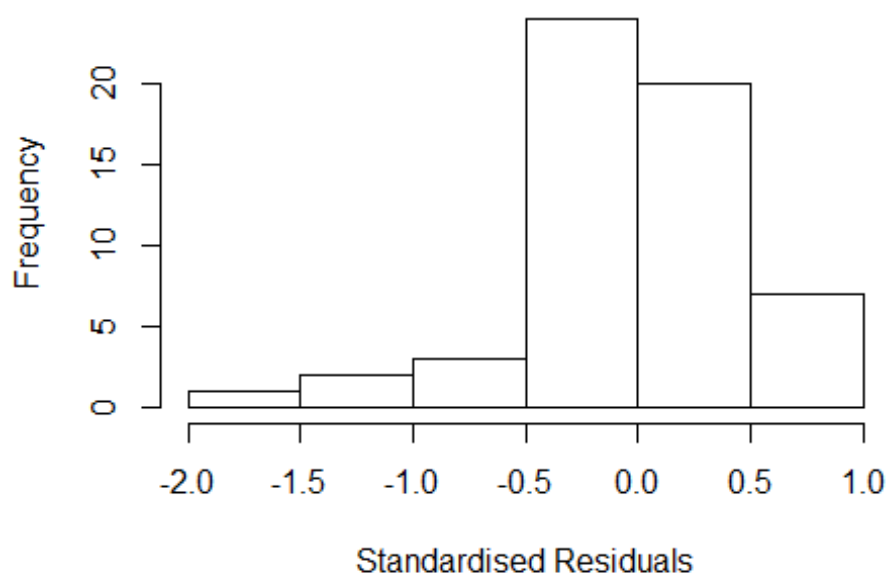
```

Residuals from first and last author and team size



| ## | | GVIF | Df | $GVIF^{1/(2 \cdot Df)}$ |
|----|-------------------|------|----|-------------------------|
| ## | FirstAuthorFemale | NaN | 1 | NaN |
| ## | LastAuthorFemale | NaN | 1 | NaN |
| ## | Year | NaN | 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
## -1.66e+00 -3.08e-01 -2.50e-16 2.54e-01 8.13e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1910 0.0000 Inf <2e-16 ***
## FirstAuthorFemale1 0.0733 0.1884 0.39 0.6992
## LastAuthorFemale1 0.2438 0.1806 1.35 0.1846
## Year1998 -0.1807 0.1200 -1.51 0.1397
## Year2000 -0.5230 0.0000 -Inf <2e-16 ***
## Year2001 -0.3203 0.1884 -1.70 0.0968 .
## Year2002 -0.8830 0.2331 -3.79 0.0005 ***
## Year2003 -0.4702 0.2429 -1.94 0.0600 .
## Year2004 0.0630 0.0000 Inf <2e-16 ***
## Year2005 0.2090 0.0000 Inf <2e-16 ***
## Year2006 0.1125 0.2196 0.51 0.6113
## Year2007 -0.2182 0.3269 -0.67 0.5083
```



```

## Year2008          -0.3608      0.1359   -2.66   0.0113 *
## Year2009          -0.1136      0.2712   -0.42   0.6776
## Year2010           0.4696      0.3308    1.42   0.1635
## Year2011          -0.3872      0.1918   -2.02   0.0502 .
## Year2012          -0.0842      0.1716   -0.49   0.6261
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.517
## Multiple R-squared:  0.294, Adjusted R-squared:  0.0122
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 44 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.281  0.899  0.962  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      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 2.722e+15  1      5.217e+07
## Year              2.722e+15 14      3.558e+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
## -1.66e+00 -2.37e-01 -1.39e-16  2.69e-01  7.62e-01
##

```

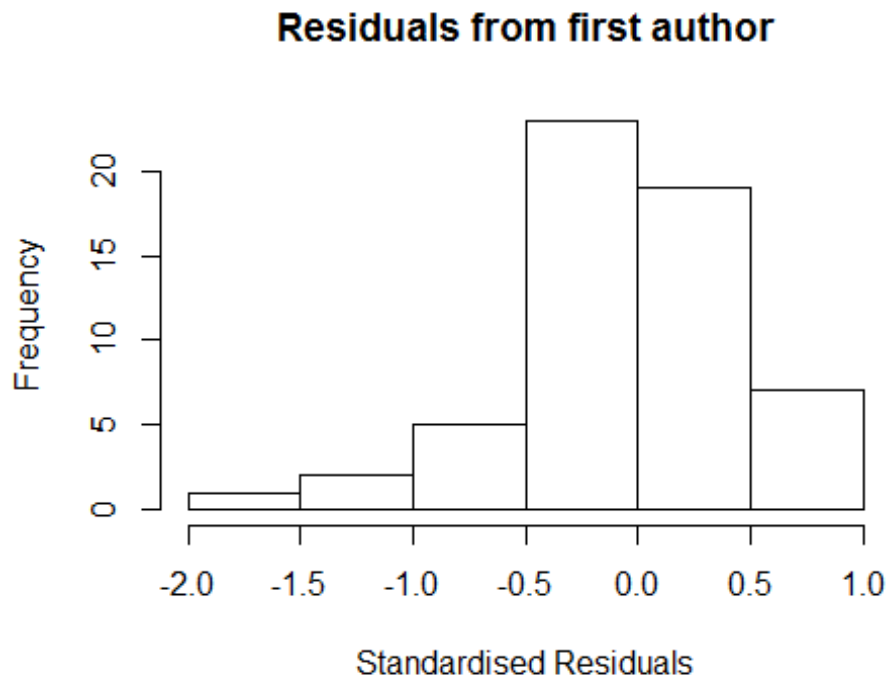
```

## Coefficients:
##               Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.19e+00  4.50e-08  2.65e+07 < 2e-16 ***
## FirstAuthorFemale1 1.34e-01  2.10e-01  6.40e-01  0.52666
## Year1998        -1.81e-01  1.20e-01 -1.51e+00  0.13995
## Year2000         -5.23e-01  3.73e-08 -1.40e+07 < 2e-16 ***
## Year2001         -3.81e-01  2.10e-01 -1.81e+00  0.07700 .
## Year2002         -8.83e-01  2.34e-01 -3.77e+00  0.00052 ***
## Year2003         -4.69e-01  2.44e-01 -1.92e+00  0.06168 .
## Year2004          6.30e-02  3.27e-08  1.93e+06 < 2e-16 ***
## Year2005          2.09e-01  4.18e-08  5.00e+06 < 2e-16 ***
## Year2006          9.56e-02  2.24e-01  4.30e-01  0.67222
## Year2007         -8.40e-02  3.63e-01 -2.30e-01  0.81829
## Year2008         -3.61e-01  1.36e-01 -2.66e+00  0.01123 *
## Year2009         -6.29e-02  2.45e-01 -2.60e-01  0.79853
## Year2010          4.68e-01  3.17e-01  1.47e+00  0.14827
## Year2011         -2.96e-01  1.77e-01 -1.68e+00  0.10136
## Year2012         -4.46e-02  1.72e-01 -2.60e-01  0.79745
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.5
## Multiple R-squared:  0.285, Adjusted R-squared:  0.0229
## 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.
##  0.249  0.887  0.965  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           1.75e-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"

## 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 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
## -1.70e+00 -2.85e-01  1.11e-16  2.56e-01  8.15e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1910     0.0000    Inf <2e-16 ***
## LastAuthorFemale1  0.2605     0.1923     1.35  0.1829
## Year1998         -0.1808     0.1200    -1.51  0.1396
## Year2000         -0.5230     0.0000   -Inf <2e-16 ***
## Year2001         -0.2470     0.0000   -Inf <2e-16 ***
```

```

## Year2002          -0.8830      0.2334    -3.78    0.0005 ***
## Year2003          -0.4699      0.2432    -1.93    0.0602 .
## Year2004           0.0630      0.0000      Inf    <2e-16 ***
## Year2005           0.2090      0.0000      Inf    <2e-16 ***
## Year2006           0.1338      0.1933      0.69    0.4926
## Year2007          -0.1776      0.2927    -0.61    0.5474
## Year2008          -0.3609      0.1359    -2.66    0.0112 *
## Year2009          -0.1163      0.2728    -0.43    0.6722
## Year2010           0.5071      0.2890      1.75    0.0868 .
## Year2011          -0.3745      0.1807    -2.07    0.0445 *
## Year2012          -0.0861      0.1714    -0.50    0.6180
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.513
## Multiple R-squared:  0.298, Adjusted R-squared:  0.041
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 43 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.251  0.900  0.960  0.910  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.75e-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 2606"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1998 2001 2002 2005 2006 2008 2012
##    3    1    1    1    1    3    1
##

```

```

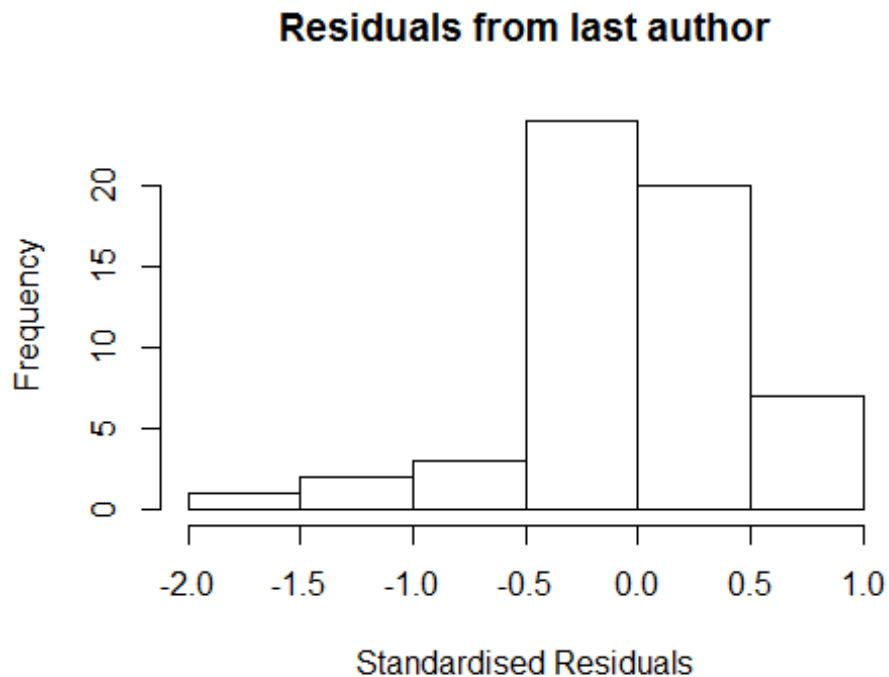
## 1998 2001 2002 2005 2006 2008 2012
##    3    1    1    1    1    2    1
##
## 1998 2001 2002 2005 2006 2008 2012
##    1    0    0    1    1    1    1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: NaN"
## [1] "Male first author team size 2018 geometric mean: 3.66284150148471"
## [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: 5"
## [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]"
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    3    4    2    5    2    1    3    3    8    8    8    17    11    18    11
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    3    1    1    3    2    0    2    3    6    8    7    16    11    13    9
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    3    1    1    2    2    0    2    3    6    8    7    16    11    13    9
## [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: 1.41421356237309"

## 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 = 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"
## [1] "Male last author team size 2018 geometric mean: 1.48599428913695"

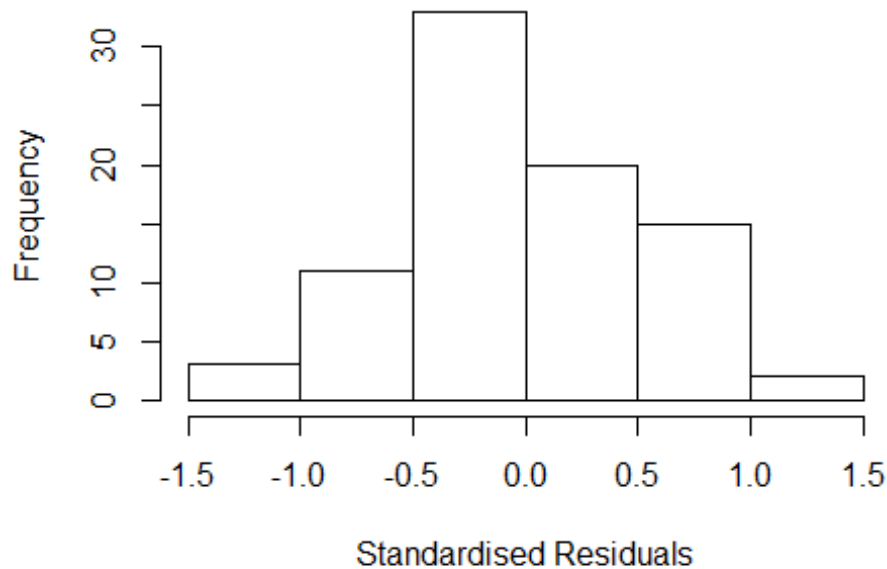
```

```
## Warning in wilcox.test.default(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 2.920e+13  1      5.404e+06
## LastAuthorFemale  4.175e+14  1      2.043e+07
## UniqueAuthors    1.942e+28  3      5.185e+04
## Year              6.726e+29 13      1.404e+01
```

Residuals from first and last author and team size



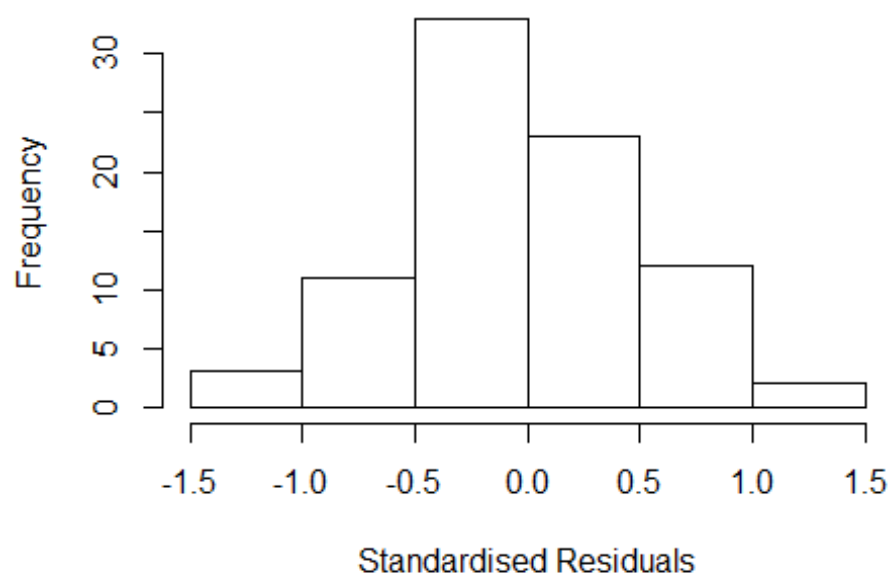
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1443 -0.2771 -0.0406 0.3214 1.3045
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.550 0.419 1.31 0.1939
## FirstAuthorFemale1 0.161 0.138 1.17 0.2474
## LastAuthorFemale1 0.284 0.177 1.60 0.1139
## UniqueAuthors2 -0.195 0.191 -1.02 0.3116
## UniqueAuthors3 -0.322 0.273 -1.18 0.2430
## UniqueAuthors4 0.174 0.151 1.15 0.2560
## Year1999 1.217 0.455 2.67 0.0095 **
## Year2000 -0.461 0.459 -1.01 0.3183
## Year2001 1.675 0.505 3.32 0.0015 **
## Year2002 1.052 0.506 2.08 0.0416 *
```

```

## Year2004          0.782      0.433      1.81      0.0755 .
## Year2005          0.327      0.469      0.70      0.4885
## Year2006          0.767      0.641      1.20      0.2353
## Year2007          0.527      0.471      1.12      0.2677
## Year2008          0.251      0.527      0.48      0.6355
## Year2009          0.594      0.463      1.28      0.2037
## Year2010          0.650      0.441      1.47      0.1453
## Year2011          0.372      0.466      0.80      0.4274
## Year2012          0.111      0.450      0.25      0.8055
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.593
## Multiple R-squared:  0.267, Adjusted R-squared:  0.0645
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 73 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.608  0.870  0.968   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      1.19e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.817e+14 1      1.348e+07
## LastAuthorFemale  3.376e+01 1      5.810e+00
## Year              8.536e+15 13      4.100e+00

```


Residuals from first and last author



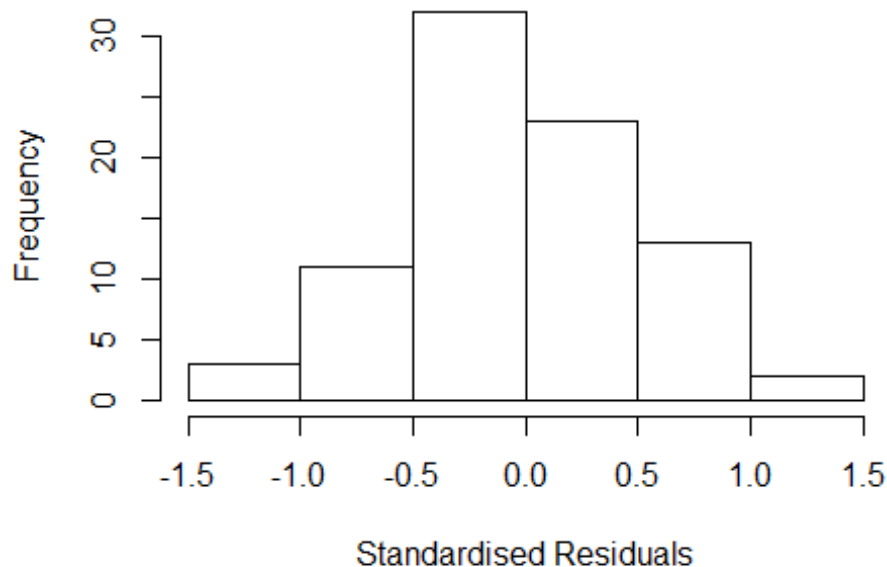
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1917 -0.2808 -0.0397 0.3637 1.3438
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5522 0.4152 1.33 0.188
## FirstAuthorFemale1 0.1533 0.1333 1.15 0.254
## LastAuthorFemale1 0.1733 0.1507 1.15 0.254
## Year1999 1.1314 0.4419 2.56 0.013 *
## Year2000 -0.3449 0.4525 -0.76 0.449
## Year2001 1.4821 0.4603 3.22 0.002 **
## Year2002 0.8586 0.4593 1.87 0.066 .
## Year2004 0.8394 0.4246 1.98 0.052 .
## Year2005 0.3244 0.4650 0.70 0.488
## Year2006 0.6394 0.6456 0.99 0.325
## Year2007 0.4336 0.4798 0.90 0.369
## Year2008 0.1411 0.4878 0.29 0.773
```

```

## Year2009          0.4934      0.4458      1.11      0.272
## Year2010          0.6335      0.4382      1.45      0.153
## Year2011          0.3310      0.4580      0.72      0.472
## Year2012          0.0458      0.4370      0.10      0.917
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.606
## Multiple R-squared:  0.252, Adjusted R-squared:  0.0868
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 76 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.602  0.876  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      1.19e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.308e+14 1      1.519e+07
## Year              2.308e+14 13      3.568e+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
## -1.2607 -0.2791 -0.0268 0.3811 1.3030
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5501 0.4180 1.32 0.1926
## FirstAuthorFemale1 0.1953 0.1408 1.39 0.1697
## Year1999 1.3069 0.4180 3.13 0.0026 **
## Year2000 -0.2114 0.4397 -0.48 0.6322
## Year2001 1.4633 0.4571 3.20 0.0021 **
## Year2002 0.8398 0.4695 1.79 0.0781 .
## Year2004 0.9073 0.4246 2.14 0.0362 *
## Year2005 0.3269 0.4679 0.70 0.4871
## Year2006 0.7106 0.6362 1.12 0.2678
## Year2007 0.5151 0.4717 1.09 0.2786
## Year2008 0.1346 0.4912 0.27 0.7849
## Year2009 0.5468 0.4409 1.24 0.2191
```

```

## Year2010          0.6584      0.4467      1.47      0.1450
## Year2011          0.3739      0.4619      0.81      0.4210
## Year2012          0.0899      0.4363      0.21      0.8374
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.594
## Multiple R-squared:  0.243, Adjusted R-squared:  0.0891
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 74 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.609  0.876  0.958  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      1.19e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 12.9 1          3.592
## Year              12.9 13          1.103
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.1866 -0.3071 -0.0352  0.3553  1.3172
##
## Coefficients:
##      Estimate Std. Error t value Pr(>|t|)

```

```

## (Intercept)          0.5522      0.4152      1.33      0.188
## LastAuthorFemale1    0.1871      0.1493      1.25      0.214
## Year1999              1.1176      0.4412      2.53      0.014 *
## Year2000             -0.2054      0.4412     -0.47      0.643
## Year2001              1.5588      0.4852      3.21      0.002 **
## Year2002              0.9353      0.4345      2.15      0.035 *
## Year2004              0.9092      0.4235      2.15      0.035 *
## Year2005              0.3244      0.4649      0.70      0.488
## Year2006              0.6343      0.6463      0.98      0.330
## Year2007              0.4274      0.4792      0.89      0.376
## Year2008              0.1696      0.4813      0.35      0.726
## Year2009              0.4893      0.4458      1.10      0.276
## Year2010              0.6315      0.4379      1.44      0.154
## Year2011              0.3576      0.4552      0.79      0.435
## Year2012              0.0425      0.4370      0.10      0.923
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.606
## Multiple R-squared:  0.248, Adjusted R-squared:  0.0949
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 75 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.616  0.866  0.963   0.923  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           1.19e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01           5.00e-01
##   nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500          50          2          1          1000          200
##   trace.lev      mts      compute.rd
##           0          1000          0
##           psi           subsampling           cov
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 84"
## [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]"

```

```

##
## 1997 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    3    2    1    1    1    1    1    2    6   10    4   11    6   10    9
##
## 1997 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    3    2    1    1    1    1    0    2    5    9    4   11    5    6    7
##
## 1997 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    3    2    1    1    1    1    0    2    5    9    4   11    3    6    7
## [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: 1.25992104989487"

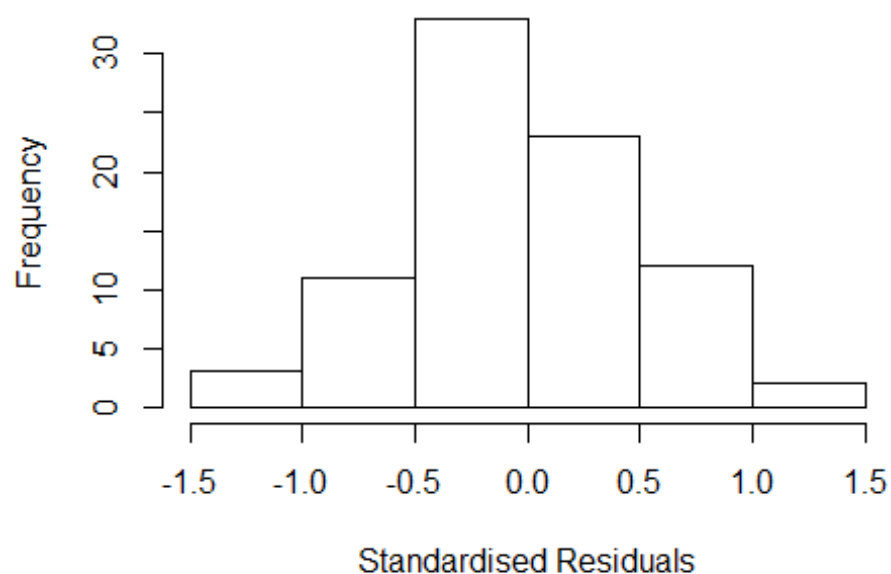
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2.5, 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: 1.25992104989487"

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

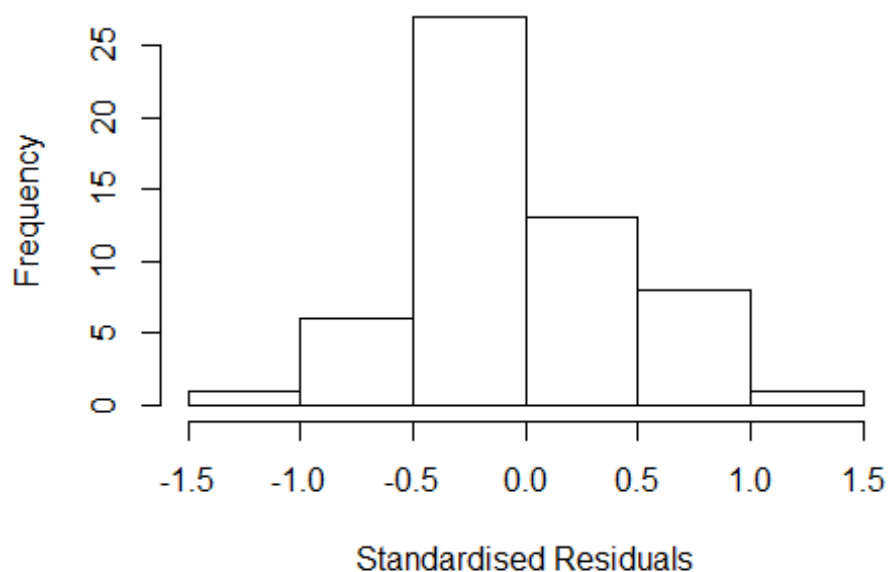
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2.5, p-value = 0.6
## 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"
##
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.412e+00  1      1.188e+00
## LastAuthorFemale  7.378e+14  1      2.716e+07
## Year              9.593e+14 13      3.769e+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
## -1.04811 -0.24964 -0.00753 0.26044 1.12298
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9589 0.1954 4.91 1.6e-05 ***
## FirstAuthorFemale1 -0.7503 0.2114 -3.55 0.00101 **
## LastAuthorFemale1 0.3969 0.1900 2.09 0.04312 *
## Year1998 -0.0699 0.5477 -0.13 0.89904
## Year1999 0.5011 0.2742 1.83 0.07503 .
## Year2000 -0.2909 0.1954 -1.49 0.14431
## Year2002 0.7071 0.1954 3.62 0.00082 ***
## Year2003 -0.6099 0.1954 -3.12 0.00333 **
## Year2005 -0.0384 0.2469 -0.16 0.87706
## Year2006 0.0892 0.4130 0.22 0.83014
## Year2007 -0.3825 0.2458 -1.56 0.12750
## Year2008 -0.4905 0.4100 -1.20 0.23862
```

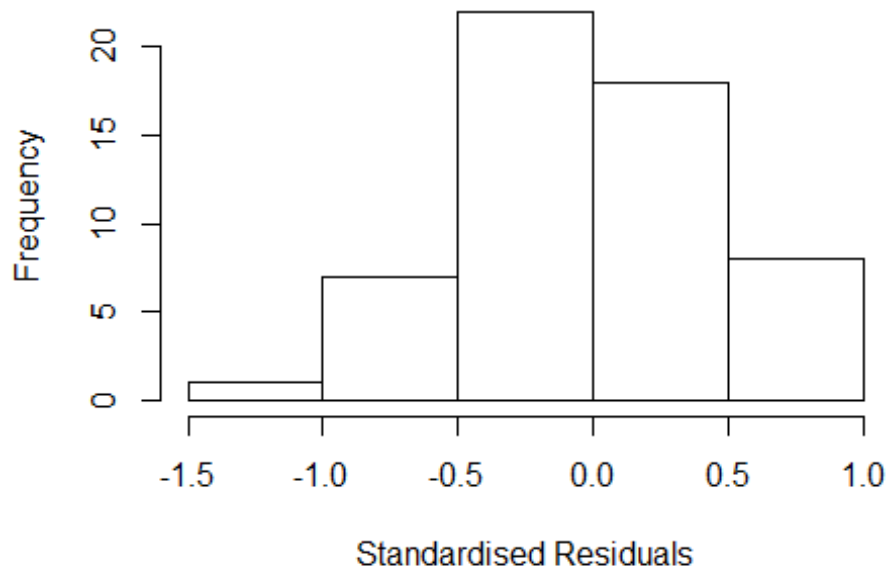


```

## Year2009          -0.3559      0.3044    -1.17  0.24923
## Year2010           0.5362      0.2495     2.15  0.03776 *
## Year2011           0.0761      0.2290     0.33  0.74135
## Year2012          -0.3860      0.3044    -1.27  0.21206
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.493
## Multiple R-squared:  0.391, Adjusted R-squared:  0.163
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 42 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.583  0.880   0.944   0.909   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      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"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.723 1          1.313
## Year              1.723 13          1.021

```

Residuals from first author



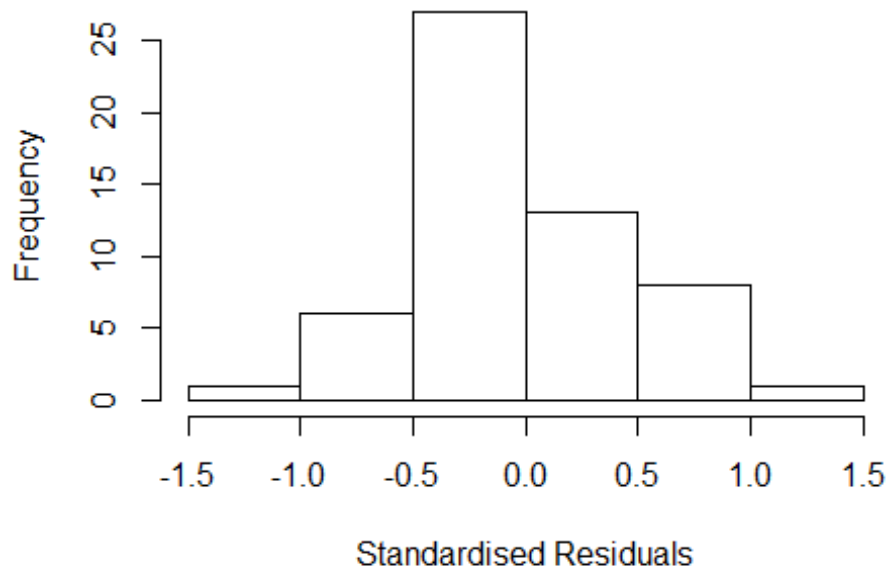
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.15e+00 -3.44e-01 -1.67e-16 3.55e-01 9.04e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9585 0.1924 4.98 1.2e-05 ***
## FirstAuthorFemale1 -0.5401 0.2583 -2.09 0.04272 *
## Year1998 -0.0695 0.5113 -0.14 0.89251
## Year1999 0.8985 0.1924 4.67 3.2e-05 ***
## Year2000 -0.2905 0.1924 -1.51 0.13864
## Year2002 0.7075 0.1924 3.68 0.00068 ***
## Year2003 -0.6095 0.1924 -3.17 0.00289 **
## Year2005 -0.0380 0.2437 -0.16 0.87682
## Year2006 0.1908 0.4368 0.44 0.66456
## Year2007 -0.2626 0.2506 -1.05 0.30090
## Year2008 -0.4269 0.4011 -1.06 0.29332
## Year2009 -0.1364 0.2501 -0.55 0.58844
```

```

## Year2010          0.4699      0.2812      1.67  0.10230
## Year2011          0.1732      0.2343      0.74  0.46402
## Year2012         -0.3192      0.2759     -1.16  0.25397
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.568
## Multiple R-squared:  0.318, Adjusted R-squared:  0.085
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 50 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.661  0.916  0.961  0.937  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      1.79e-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"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale -5.108e+14 1      NaN
## Year             -5.108e+14 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
## -1.10e+00 -2.93e-01 -3.33e-16 2.22e-01 1.12e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9587 0.1940 4.94 1.3e-05 ***
## LastAuthorFemale1 0.1964 0.1799 1.09 0.28130
## Year1998 -0.0697 0.5302 -0.13 0.89600
## Year1999 0.7019 0.2658 2.64 0.01167 *
## Year2000 -0.2907 0.1940 -1.50 0.14155
## Year2002 0.7073 0.1940 3.65 0.00074 ***
## Year2003 -0.6097 0.1940 -3.14 0.00310 **
## Year2005 -0.0382 0.2454 -0.16 0.87694
## Year2006 0.1461 0.4266 0.34 0.73372
## Year2007 -0.3223 0.2463 -1.31 0.19795
## Year2008 -0.6657 0.3595 -1.85 0.07126 .
## Year2009 -0.3498 0.3057 -1.14 0.25908
```

```

## Year2010          0.3344      0.4502      0.74  0.46191
## Year2011          0.0301      0.2116      0.14  0.88746
## Year2012         -0.3531      0.2895     -1.22  0.22954
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.524
## Multiple R-squared:  0.293, Adjusted R-squared:  0.0517
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 48 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.629  0.882  0.963  0.918  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      1.79e-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 2609"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 2005 2009 2010 2012
##    1    1    1    2
##
## 2005 2009 2010 2012
##    1    0    0    2
##
## 2005 2009 2010 2012
##    1    0    0    2
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: NaN"

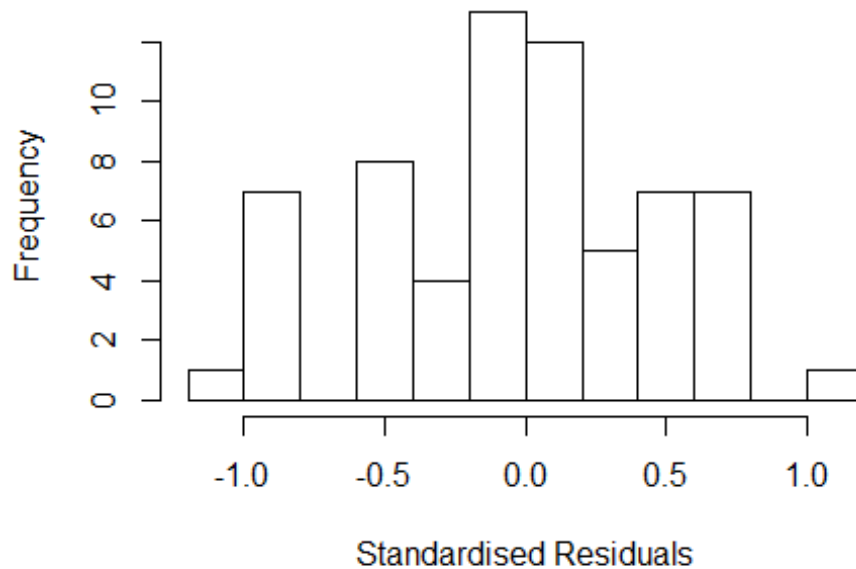
```

```

## [1] "Male first author team size 2018 geometric mean: 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 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
##    7   10    7    6   12   10    7    6   12    5   11    9   16   10    4
## 2011 2012
##   12    8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    2    2    3    1    2    3    1    6    3    6    6   11    8    3
## 2011 2012
##    8    4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    2    2    3    1    2    3    1    6    3    4    5   11    7    2
## 2011 2012
##    7    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: 1.25992104989487"
## [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 26.13 1          5.112
## Year              26.13 16          1.107

```

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
## -1.03e+00 -2.70e-01 3.33e-16 3.34e-01 1.11e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9348 0.2644 3.54 0.00093 ***
## FirstAuthorFemale1 -0.3711 0.2688 -1.38 0.17386
## Year1997 -0.2888 0.2942 -0.98 0.33132
## Year1998 -0.4783 0.2903 -1.65 0.10610
## Year1999 -0.1785 0.4061 -0.44 0.66223
## Year2000 0.2612 0.2644 0.99 0.32822
## Year2001 1.0152 0.4342 2.34 0.02368 *
## Year2002 -0.1292 0.5008 -0.26 0.79760
## Year2003 0.4652 0.2644 1.76 0.08498 .
## Year2004 0.0911 0.3759 0.24 0.80968
## Year2005 0.0537 0.2811 0.19 0.84929
## Year2006 -0.1156 0.3962 -0.29 0.77178
```

```

## Year2007          -0.1103      0.3597   -0.31  0.76058
## Year2008           0.0643      0.3792    0.17  0.86603
## Year2009          -0.1227      0.3292   -0.37  0.71104
## Year2010          -0.0283      0.5359   -0.05  0.95810
## Year2011           0.4273      0.2973    1.44  0.15728
## Year2012           0.5665      0.2891    1.96  0.05597 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.645
## Multiple R-squared:  0.241, Adjusted R-squared:  -0.0337
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 56 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.747  0.905  0.957  0.939  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      1.54e-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 7.388e+14 1      2.718e+07
## Year              7.388e+14 16      2.915e+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
## -1.03e+00 -2.70e-01  1.67e-16  3.34e-01  1.11e+00

```



```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9348    0.2644    3.54 0.00093 ***
## LastAuthorFemale1 -0.3711    0.2688   -1.38 0.17386
## Year1997        -0.2888    0.2942   -0.98 0.33132
## Year1998        -0.4783    0.2903   -1.65 0.10610
## Year1999        -0.1785    0.4061   -0.44 0.66223
## Year2000         0.2612    0.2644    0.99 0.32822
## Year2001         1.0152    0.4342    2.34 0.02368 *
## Year2002        -0.1292    0.5008   -0.26 0.79760
## Year2003         0.4652    0.2644    1.76 0.08498 .
## Year2004         0.0911    0.3759    0.24 0.80968
## Year2005         0.0537    0.2811    0.19 0.84929
## Year2006        -0.1156    0.3962   -0.29 0.77178
## Year2007        -0.1103    0.3597   -0.31 0.76058
## Year2008         0.0643    0.3792    0.17 0.86603
## Year2009        -0.1227    0.3292   -0.37 0.71104
## Year2010        -0.0283    0.5359   -0.05 0.95810
## Year2011         0.4273    0.2973    1.44 0.15728
## Year2012         0.5665    0.2891    1.96 0.05597 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.645
## Multiple R-squared:  0.241, Adjusted R-squared:  -0.0337
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 56 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.747  0.905   0.957   0.939   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      1.54e-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 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
##      8      2     23     11     13      8     14     10     20     17     18     18     22     28     20
## 2011 2012
##     28     33
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      1      1      7      2      6      3      6      6      7      6     12      8     12     22     12
## 2011 2012
##     21     26
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      1      1      6      2      6      2      5      4      4      4     12      6     11     17     11
## 2011 2012
##     18     23
## [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: 2.2837538219638"

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

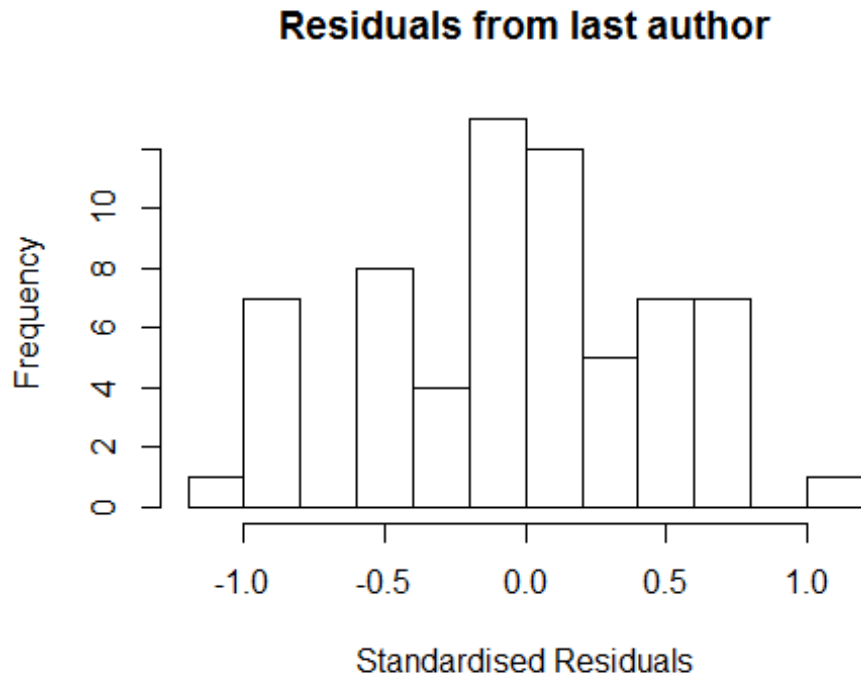
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 6, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4"
## [1] "Male last author team size 2018 geometric mean: 2.2837538219638"

## Warning in wilcox.test.default(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 = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"

```

```
## Warning in outlierStats(ret, x, control): Detected possible local
breakdown of SM-estimate in coefficient 'Year1999'.
## Use lmrob argument 'setting="KS2014"' to avoid this problem.
```



```
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale  3.421  1         1.850
## LastAuthorFemale   1.603  1         1.266
## UniqueAuthors     38.263  4         1.577
## Year               53.495 16         1.132

## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId NLCS Year OneField Fields residuals
## 50 53149135804  0 1999   2611      1    -2.85
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##      Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##      k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q   Median      3Q      Max
## -2.85e+00 -2.52e-01 -4.44e-16  3.00e-01  1.29e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
```

```

## (Intercept)          4.62e-01    6.80e-08    6.80e+06    < 2e-16 ***
## FirstAuthorFemale1 -3.09e-02    1.37e-01   -2.20e-01    0.82264
## LastAuthorFemale1  1.56e-01    1.90e-01    8.20e-01    0.41396
## UniqueAuthors2     2.13e-01    1.43e-01    1.49e+00    0.13813
## UniqueAuthors3     1.63e-01    1.66e-01    9.80e-01    0.32838
## UniqueAuthors4     3.39e-01    1.81e-01    1.87e+00    0.06377 .
## UniqueAuthors5     1.82e-01    2.06e-01    8.80e-01    0.37963
## Year1997            -1.16e-01    9.16e-08   -1.27e+06    < 2e-16 ***
## Year1998            1.02e+00    9.27e-02    1.10e+01    < 2e-16 ***
## Year1999            2.39e+00    5.88e-08    4.06e+07    < 2e-16 ***
## Year2000            4.86e-01    2.48e-01    1.96e+00    0.05293 .
## Year2001            1.64e-01    1.83e+00    9.00e-02    0.92866
## Year2002            1.02e-01    5.72e-01    1.80e-01    0.85940
## Year2003            4.77e-01    1.40e-01    3.42e+00    0.00089 ***
## Year2004            4.85e-01    1.37e-01    3.53e+00    0.00061 ***
## Year2005            8.29e-01    2.78e-01    2.99e+00    0.00349 **
## Year2006            6.86e-01    1.31e-01    5.23e+00    8e-07 ***
## Year2007            2.12e-01    2.17e-01    9.70e-01    0.33210
## Year2008            4.25e-01    2.35e-01    1.81e+00    0.07307 .
## Year2009            4.55e-01    1.61e-01    2.83e+00    0.00550 **
## Year2010            1.69e-01    2.05e-01    8.20e-01    0.41223
## Year2011            3.58e-01    1.82e-01    1.97e+00    0.05170 .
## Year2012            4.55e-01    1.74e-01    2.62e+00    0.01005 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.472
## Multiple R-squared:  0.263, Adjusted R-squared:  0.115
## Convergence in 33 IRWLS iterations
##
## Robustness weights:
## observation 9 is an outlier with |weight| = 0 ( < 0.00075);
## 15 weights are ~= 1. The remaining 117 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.436  0.862  0.957   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           7.52e-04           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01           5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500           50           2           1           1000           200
## trace.lev      mts      compute.rd
##           0           1000           0
##           psi           subsampling           cov
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"

```

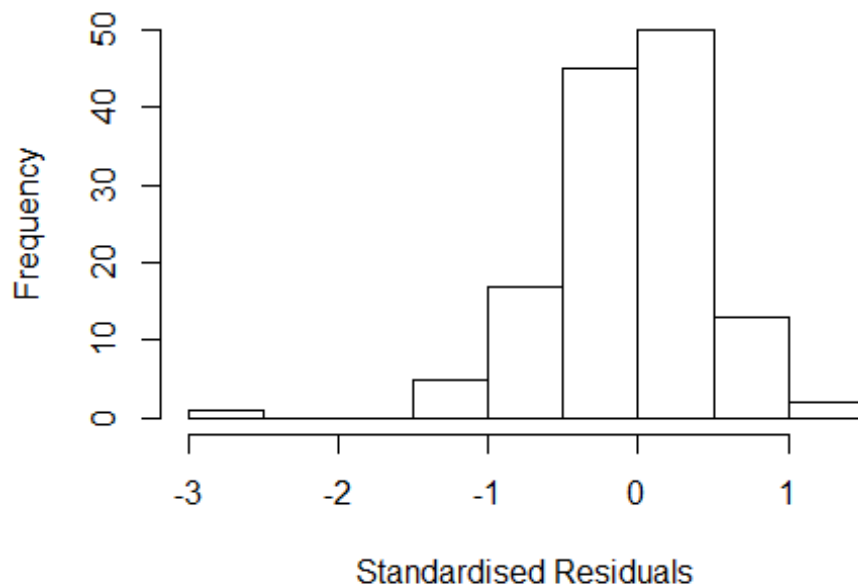
```
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"

## Warning in outlierStats(ret, x, control): Detected possible local
breakdown of SM-estimate in coefficient 'Year1999'.
## Use lmrob argument 'setting="KS2014"' to avoid this problem.

## 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 and team size



```
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1 NaN
## LastAuthorFemale NaN 1 NaN
## Year NaN 16 NaN

## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId NLCS Year OneField Fields residuals
## 50 53149135804 0 1999 2611 1 -2.85
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
```

```

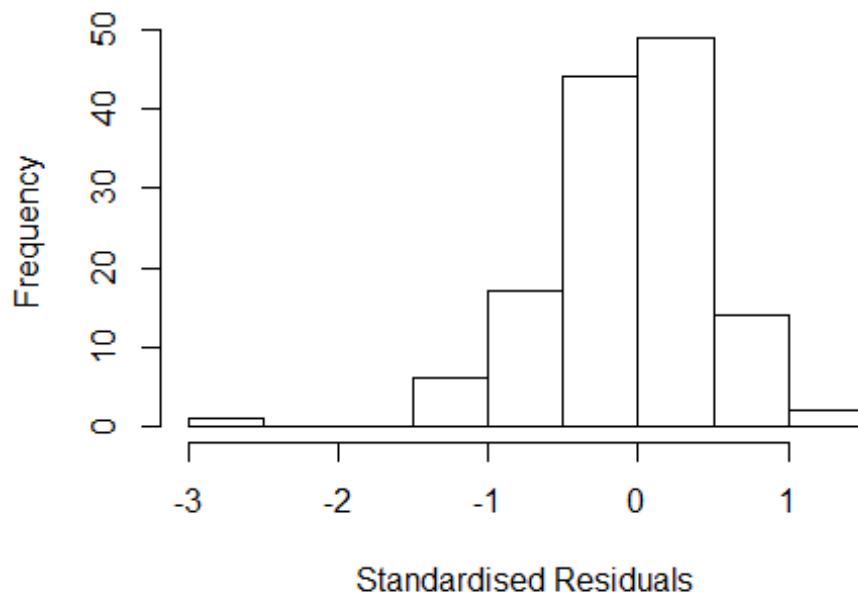
## Residuals:
##      Min      1Q      Median      3Q      Max
## -2.85e+00 -3.37e-01  1.33e-15  3.31e-01  1.20e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   4.62e-01   2.01e-09  2.30e+08 < 2e-16 ***
## FirstAuthorFemale1 3.00e-02   1.19e-01  2.50e-01  0.8018
## LastAuthorFemale1 1.59e-01   1.79e-01  8.90e-01  0.3752
## Year1997      -1.16e-01   0.00e+00   -Inf < 2e-16 ***
## Year1998       1.09e+00   1.29e-01  8.40e+00  1.4e-13 ***
## Year1999       2.39e+00   0.00e+00   Inf < 2e-16 ***
## Year2000       5.51e-01   2.30e-01  2.40e+00  0.0181 *
## Year2001       3.77e-01   1.46e+00  2.60e-01  0.7959
## Year2002       1.87e-01   4.01e-01  4.70e-01  0.6421
## Year2003       6.22e-01   1.18e-01  5.27e+00  6.5e-07 ***
## Year2004       6.34e-01   4.13e-02  1.54e+01 < 2e-16 ***
## Year2005       8.97e-01   3.03e-01  2.97e+00  0.0037 **
## Year2006       8.17e-01   9.00e-02  9.08e+00  3.9e-15 ***
## Year2007       3.73e-01   1.87e-01  1.99e+00  0.0486 *
## Year2008       5.86e-01   1.99e-01  2.94e+00  0.0040 **
## Year2009       6.04e-01   1.43e-01  4.23e+00  4.8e-05 ***
## Year2010       3.18e-01   1.97e-01  1.62e+00  0.1085
## Year2011       4.96e-01   1.23e-01  4.02e+00  0.0001 ***
## Year2012       6.51e-01   1.26e-01  5.15e+00  1.1e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.499
## Multiple R-squared:  0.227, Adjusted R-squared:  0.105
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## observation 9 is an outlier with |weight| = 0 ( < 0.00075);
## 17 weights are ~ 1. The remaining 115 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.540  0.863  0.952  0.906  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      7.52e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500      50      2      1      1000      200
## trace.lev      mts      compute.rd
##      0      1000      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 outlierStats(ret, x, control): Detected possible local
breakdown of SM-estimate in coefficient 'Year1999'.
## Use lmrob argument 'setting="KS2014"' to avoid this problem.
```

Residuals from first and last author



```
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.094 1 1.447
## Year 2.094 16 1.023

## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId NLCS Year OneField Fields residuals
## 50 53149135804 0 1999 2611 1 -2.85
##
## Call:
## lmrob(formula = NLCS ~ 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 -3.18e-01  1.08e-15  3.15e-01  2.85e+00
##
## Coefficients:
```

```

##               Estimate Std. Error   t value Pr(>|t|)
## (Intercept)      4.62e-01   7.34e-08  6.29e+06 < 2e-16 ***
## FirstAuthorFemale1 4.49e-02   1.18e-01   3.80e-01   0.7044
## Year1997         -1.16e-01   6.31e-08 -1.84e+06 < 2e-16 ***
## Year1998          1.09e+00   1.29e-01   8.41e+00  1.3e-13 ***
## Year1999         -4.62e-01   5.41e-08 -8.55e+06 < 2e-16 ***
## Year2000          5.49e-01   2.30e-01   2.38e+00   0.0187 *
## Year2001          3.77e-01   1.41e+00   2.70e-01   0.7889
## Year2002          1.86e-01   4.10e-01   4.50e-01   0.6505
## Year2003          6.58e-01   1.49e-01   4.41e+00  2.3e-05 ***
## Year2004          6.34e-01   4.13e-02   1.54e+01 < 2e-16 ***
## Year2005          8.94e-01   3.01e-01   2.97e+00   0.0037 **
## Year2006          8.36e-01   8.03e-02   1.04e+01 < 2e-16 ***
## Year2007          3.88e-01   2.06e-01   1.88e+00   0.0620 .
## Year2008          6.12e-01   1.91e-01   3.20e+00   0.0018 **
## Year2009          5.99e-01   1.43e-01   4.20e+00  5.3e-05 ***
## Year2010          3.29e-01   2.05e-01   1.60e+00   0.1120
## Year2011          5.12e-01   1.18e-01   4.34e+00  3.1e-05 ***
## Year2012          6.74e-01   1.27e-01   5.33e+00  5.0e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.504
## Multiple R-squared:  0.177, Adjusted R-squared:  0.0555
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## observation 10 is an outlier with |weight| = 0 ( < 0.00075);
## 14 weights are ~= 1. The remaining 118 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.548  0.869  0.955  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      7.52e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"

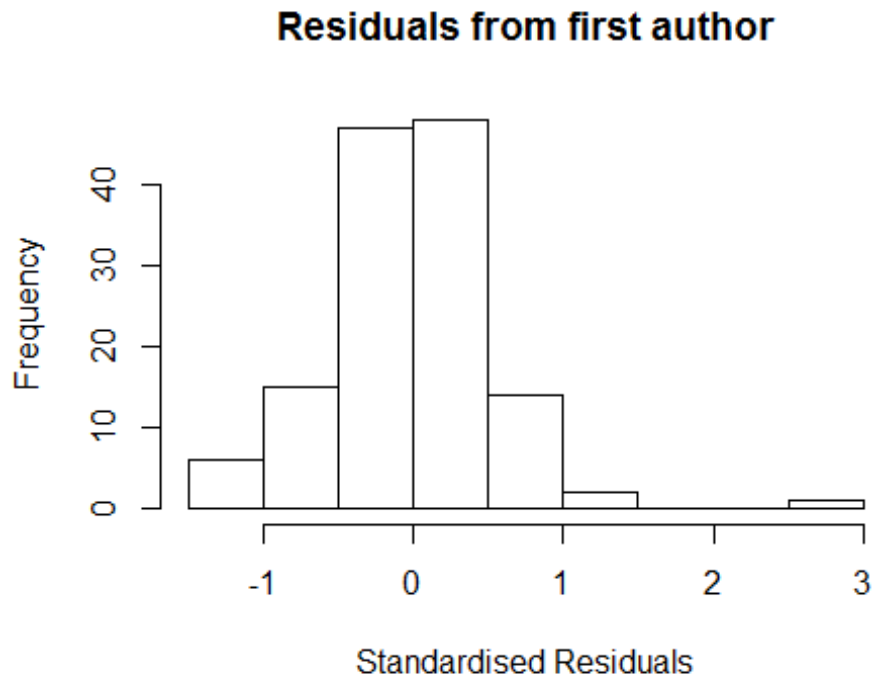
```



```
## Warning in outlierStats(ret, x, control): Detected possible local
breakdown of SM-estimate in coefficient 'Year1999'.
## Use lmrob argument 'setting="KS2014"' to avoid this problem.

## 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 16           NaN

## [1] "List of 1 outliers with residuals above 2.5"
##       ScopusId NLCS Year OneField Fields residuals
## 50 53149135804 0 1999 2611 1 -2.85
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.12e+00 -3.28e-01 -4.30e-16  3.28e-01  2.85e+00
##
```

```

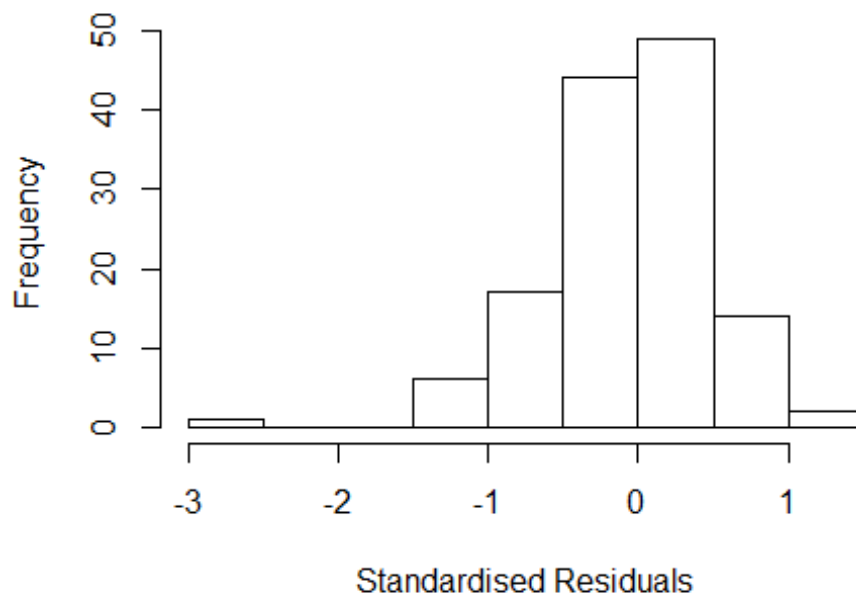
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.4620    0.0000    Inf < 2e-16 ***
## LastAuthorFemale1 0.1650    0.1747    0.94  0.3470
## Year1997       -0.1160    0.0000   -Inf < 2e-16 ***
## Year1998        1.0869    0.1293    8.40  1.3e-13 ***
## Year1999       -0.4620    0.0000   -Inf < 2e-16 ***
## Year2000        0.5501    0.2264    2.43  0.0167 *
## Year2001        0.3775    1.4282    0.26  0.7920
## Year2002        0.1941    0.4004    0.48  0.6287
## Year2003        0.6212    0.1170    5.31  5.5e-07 ***
## Year2004        0.6337    0.0413   15.36 < 2e-16 ***
## Year2005        0.9087    0.2917    3.12  0.0023 **
## Year2006        0.8233    0.0880    9.35  8.4e-16 ***
## Year2007        0.3759    0.1914    1.96  0.0519 .
## Year2008        0.5958    0.1856    3.21  0.0017 **
## Year2009        0.6117    0.1352    4.52  1.5e-05 ***
## Year2010        0.3214    0.1949    1.65  0.1019
## Year2011        0.5042    0.1181    4.27  4.0e-05 ***
## Year2012        0.6550    0.1238    5.29  5.9e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.502
## Multiple R-squared:  0.184, Adjusted R-squared:  0.0629
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## observation 10 is an outlier with |weight| = 0 ( < 0.00075);
## 17 weights are ~= 1. The remaining 115 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.550  0.864  0.954  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           7.52e-04           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01           5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500           50           2           1           1000           200
## trace.lev      mts      compute.rd
##           0           1000           0
##           psi           subsampling           cov
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 133"
## [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]"
##
## 1997 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2    3    1    2    2    3    3    3    6    7    5    8    3    9    7
##
## 1997 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2    3    1    0    2    3    1    1    6    6    5    8    3    3    5
##
## 1997 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2    3    1    0    2    3    1    1    6    6    3    8    3    2    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: 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 refinements did
not
## converge (to refine.tol=1e-07) in 1000 (= k.max) steps
```

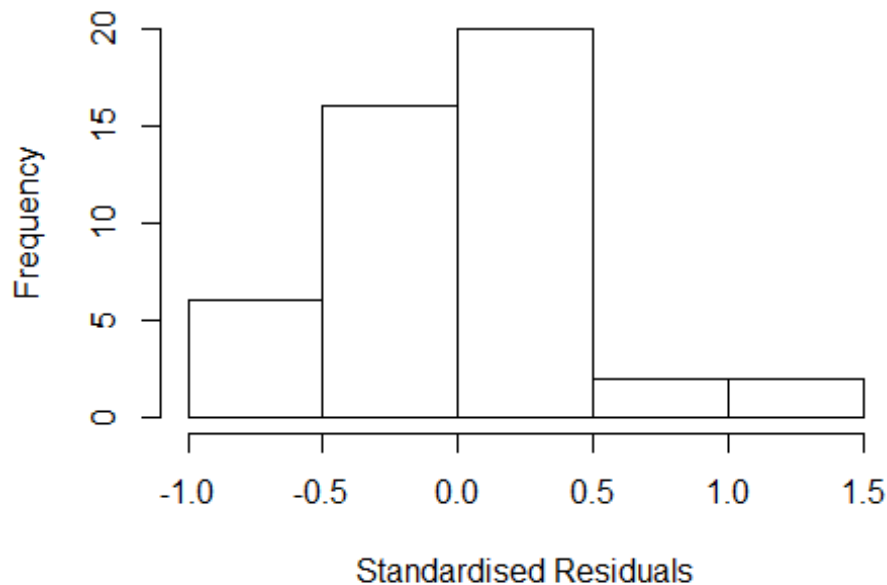
Residuals from last author



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

```
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.608e+00  1      1.899e+00
## LastAuthorFemale  2.202e+13  1      4.693e+06
## Year              6.237e+13 13      3.393e+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
## -0.8968 -0.1637  0.0257  0.1773  1.1056
```

```
##
```

```
## Coefficients:
```

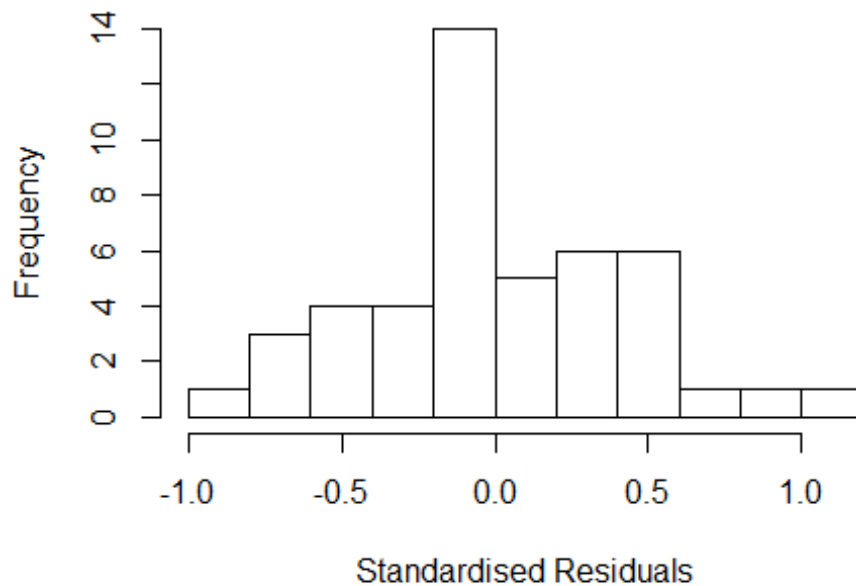
```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0395     0.4899   2.12    0.042 *
## FirstAuthorFemale1  0.0984     0.1197   0.82    0.418
## LastAuthorFemale1  0.5483     0.1138   4.82  3.9e-05 ***
## Year1998          -0.1192     0.5769  -0.21    0.838
## Year1999           0.2692     0.5004   0.54    0.595
```

```

## Year2002          0.2980      0.5530      0.54      0.594
## Year2003         -0.4689      0.5546     -0.85      0.405
## Year2004          0.3135      0.4899      0.64      0.527
## Year2005          0.0875      0.4899      0.18      0.859
## Year2006         -0.0728      0.5275     -0.14      0.891
## Year2007         -0.1652      0.4962     -0.33      0.742
## Year2008         -0.5759      0.5082     -1.13      0.266
## Year2009         -0.6036      0.5048     -1.20      0.241
## Year2010         -0.3431      0.8252     -0.42      0.681
## Year2011         -0.3500      0.5019     -0.70      0.491
## Year2012         -0.1427      0.5465     -0.26      0.796
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.461
## Multiple R-squared:  0.399, Adjusted R-squared:  0.0988
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.545  0.895  0.980  0.926  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      2.17e-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 0.9143 1      0.9562
## Year              0.9143 13      0.9966

```

Residuals from first author



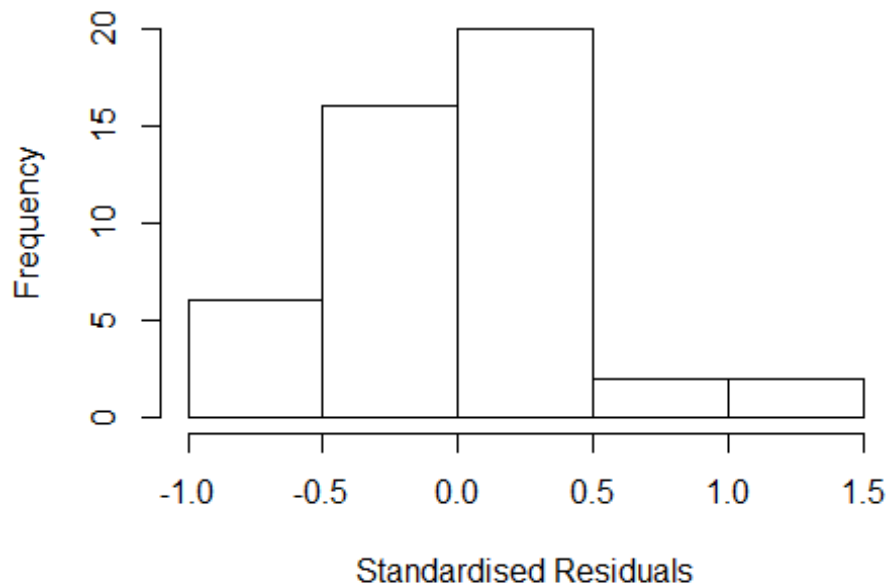
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8796 -0.2032 -0.0212 0.2858 1.0442
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0395 0.4406 2.36 0.025 *
## FirstAuthorFemale1 0.3193 0.1256 2.54 0.016 *
## Year1998 -0.1221 0.5260 -0.23 0.818
## Year1999 0.8175 0.4406 1.86 0.073 .
## Year2002 0.2980 0.5055 0.59 0.560
## Year2003 -0.4764 0.5045 -0.94 0.352
## Year2004 0.3135 0.4406 0.71 0.482
## Year2005 0.0875 0.4406 0.20 0.844
## Year2006 0.0510 0.5176 0.10 0.922
## Year2007 0.1019 0.4611 0.22 0.827
## Year2008 -0.2872 0.4517 -0.64 0.530
## Year2009 -0.3486 0.4629 -0.75 0.457
```

```

## Year2010          -0.2817      0.7239   -0.39    0.700
## Year2011          0.0878      0.4598    0.19    0.850
## Year2012         -0.1599      0.4962   -0.32    0.749
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.563
## Multiple R-squared:  0.254, Adjusted R-squared:  -0.0825
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 41 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.711  0.923  0.969  0.943  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      2.17e-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 1.631e+13 1      4.038e+06
## Year            1.631e+13 13      3.222e+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
## -9.03e-01 -1.67e-01 1.87e-16 1.68e-01 1.14e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0395 0.5112 2.03 0.051 .
## LastAuthorFemale1 0.5672 0.1117 5.08 1.7e-05 ***
## Year1998 -0.1180 0.5955 -0.20 0.844
## Year1999 0.2503 0.5266 0.48 0.638
## Year2002 0.2980 0.5735 0.52 0.607
## Year2003 -0.4660 0.5681 -0.82 0.418
## Year2004 0.3135 0.5112 0.61 0.544
## Year2005 0.0875 0.5112 0.17 0.865
## Year2006 -0.0856 0.5690 -0.15 0.881
## Year2007 -0.1738 0.5189 -0.33 0.740
## Year2008 -0.5547 0.5335 -1.04 0.306
## Year2009 -0.6006 0.5287 -1.14 0.265
```

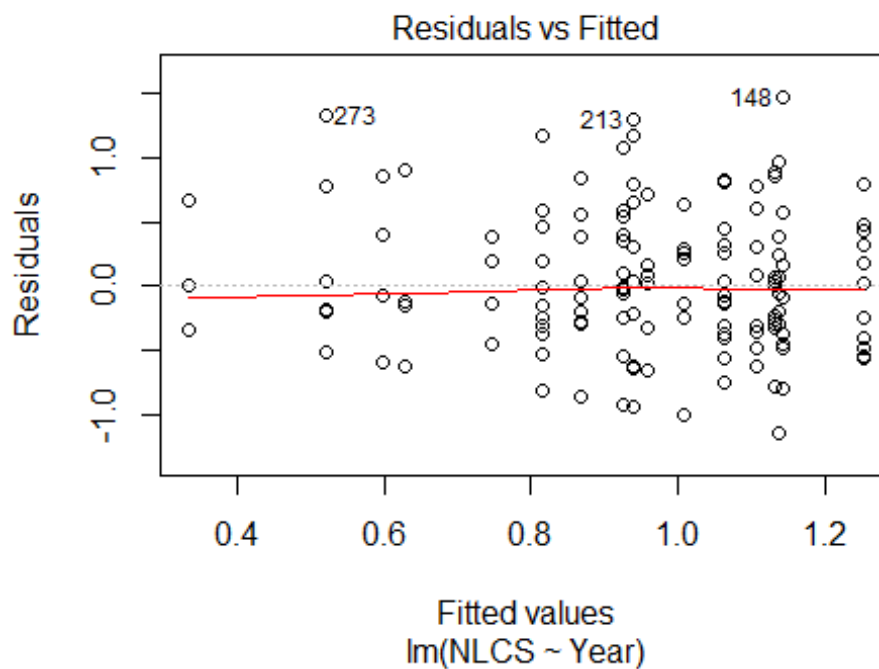


```

## Year2010          -0.3761      0.9300   -0.40      0.689
## Year2011          -0.3197      0.5266   -0.61      0.548
## Year2012          -0.1360      0.5493   -0.25      0.806
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.436
## Multiple R-squared:  0.408, Adjusted R-squared:  0.141
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 38 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.476  0.880  0.972  0.916  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      2.17e-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 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
##    7   12   14    6    7   12    6    9   15   15   10   17   21   17   15
## 2011 2012
##   18   22
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    4    8    6    5    8    4    7   10   11    8   11   12   14    9
## 2011 2012
##   10   15
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010

```

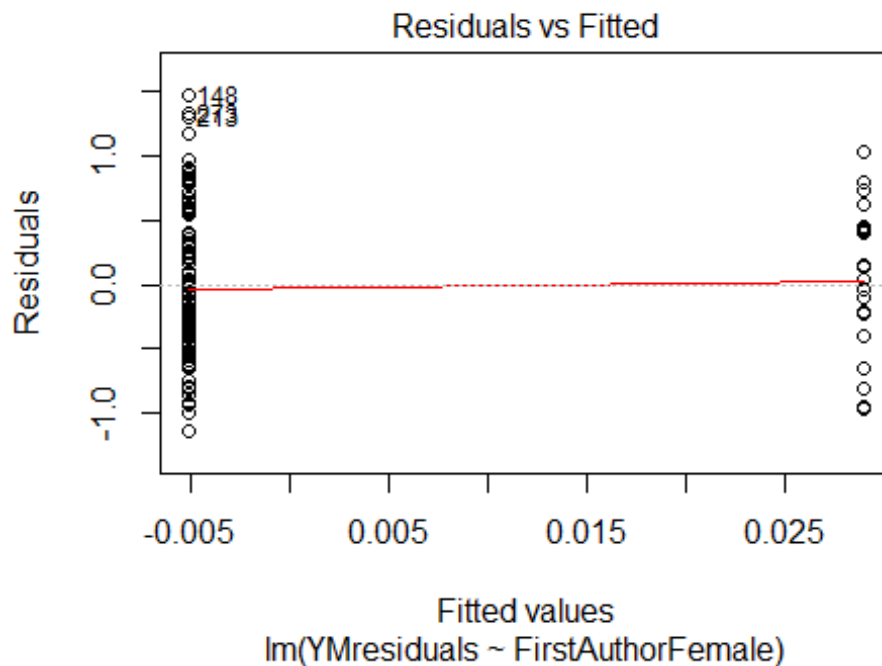
```
##      5      4      8      6      5      7      4      7      9      9      8      9      11      12      9
## 2011 2012
##    10    15
## [1] "Heteroscedasticity checks, confirming that there are problems with
##      these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 9.2, df = 16, p-value = 0.9
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.33, df = 1, p-value = 0.6
## [1] "Female first author team size 2018 geometric mean: 2.23606797749979"
## [1] "Male first author team size 2018 geometric mean: 2.55970759249787"
## Warning in wilcox.test.default(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"
## [1] "Male last author team size 2018 geometric mean: 2.78315768371374"

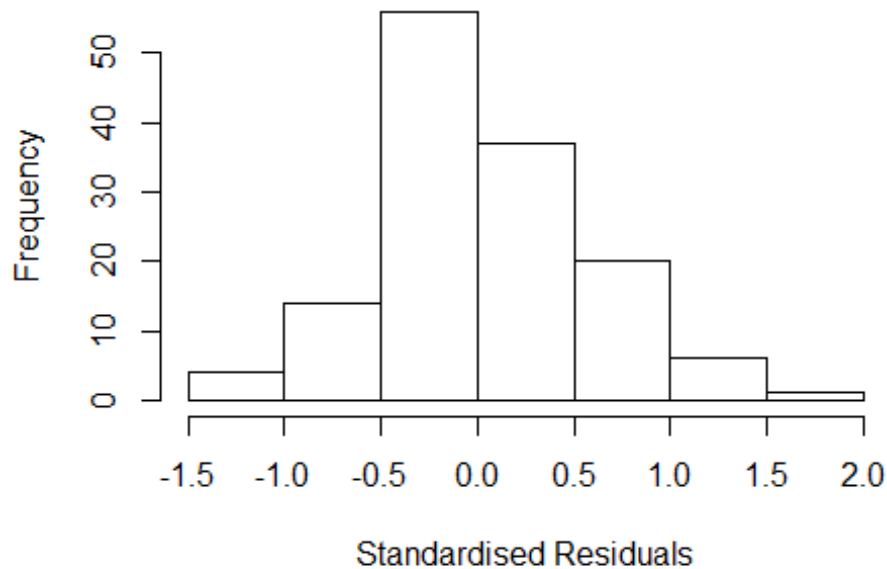
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0.5, 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 | 2.483 | 1 | 1.576 |
| LastAuthorFemale | 1.982 | 1 | 1.408 |
| UniqueAuthors | 5.788 | 4 | 1.245 |
| Year | 12.675 | 16 | 1.083 |

Residuals from first and last author and team size



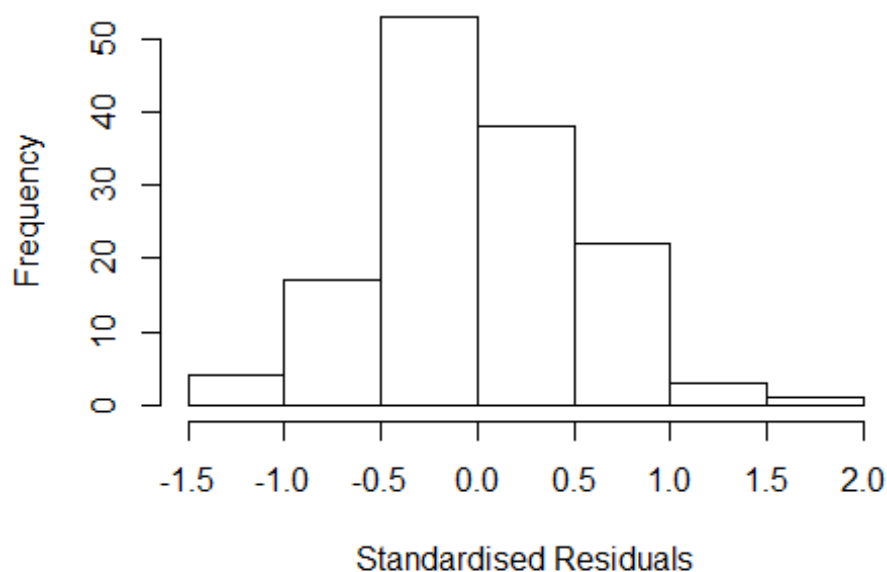
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1705 -0.3461 -0.0676 0.3176 1.7318
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.2866 0.1798 1.59 0.11374
## FirstAuthorFemale1 -0.0750 0.1938 -0.39 0.69942
## LastAuthorFemale1 0.2156 0.2115 1.02 0.31011
## UniqueAuthors2 0.2475 0.1405 1.76 0.08088 .
## UniqueAuthors3 0.2903 0.1944 1.49 0.13810
## UniqueAuthors4 0.1869 0.1424 1.31 0.19214
## UniqueAuthors5 0.4799 0.2641 1.82 0.07175 .
## Year1997 0.2939 0.3502 0.84 0.40315
## Year1998 0.6136 0.2479 2.47 0.01479 *
## Year1999 0.4909 0.2908 1.69 0.09410 .
```

```

## Year2000          0.1906      0.3205      0.59  0.55325
## Year2001          0.7752      0.2984      2.60  0.01060 *
## Year2002          0.3863      0.2760      1.40  0.16435
## Year2003          0.4953      0.2833      1.75  0.08307 .
## Year2004          0.4374      0.2579      1.70  0.09252 .
## Year2005          0.3113      0.2256      1.38  0.17028
## Year2006          0.5916      0.3152      1.88  0.06311 .
## Year2007          0.7996      0.2103      3.80  0.00023 ***
## Year2008          0.5936      0.3985      1.49  0.13908
## Year2009          0.5739      0.2689      2.13  0.03497 *
## Year2010          0.0654      0.2369      0.28  0.78313
## Year2011          0.6849      0.2426      2.82  0.00560 **
## Year2012          0.4437      0.2777      1.60  0.11281
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.53
## Multiple R-squared:  0.213, Adjusted R-squared:  0.0629
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 128 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.264  0.893  0.960  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          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 2.116 1 1.455
## LastAuthorFemale 1.618 1 1.272
## Year 2.537 16 1.030

```

Residuals from first and last author



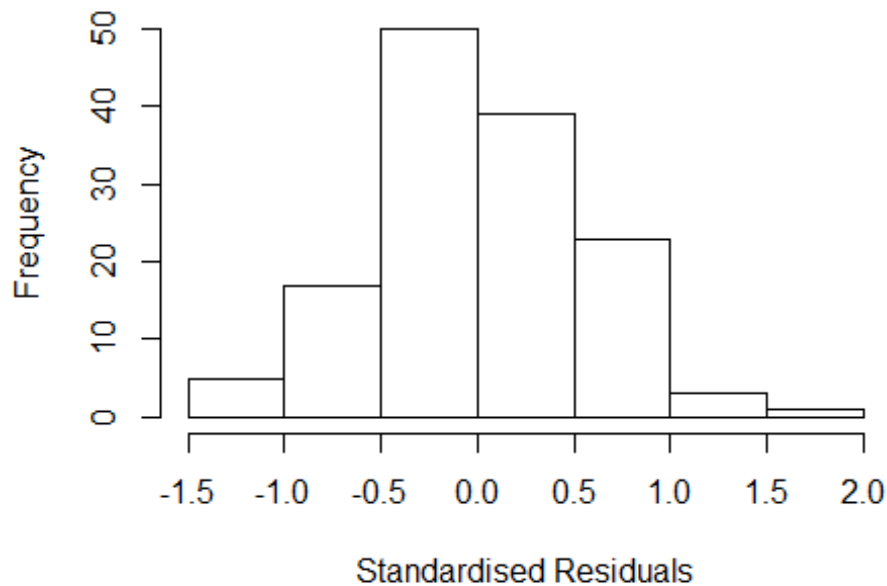
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1324 -0.3024 -0.0587 0.3786 1.6165
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.275 0.182 1.51 0.13379
## FirstAuthorFemale1 0.026 0.188 0.14 0.89033
## LastAuthorFemale1 0.169 0.200 0.85 0.39914
## Year1997 0.307 0.351 0.88 0.38280
## Year1998 0.770 0.272 2.83 0.00544 **
## Year1999 0.674 0.253 2.67 0.00866 **
## Year2000 0.248 0.353 0.70 0.48328
## Year2001 0.858 0.297 2.88 0.00466 **
## Year2002 0.475 0.248 1.92 0.05774 .
## Year2003 0.747 0.250 2.99 0.00338 **
## Year2004 0.579 0.237 2.44 0.01602 *
## Year2005 0.446 0.233 1.91 0.05821 .
```

```

## Year2006          0.719      0.302      2.38  0.01888 *
## Year2007          0.904      0.206      4.39  2.5e-05 ***
## Year2008          0.708      0.386      1.84  0.06887 .
## Year2009          0.814      0.236      3.46  0.00076 ***
## Year2010          0.141      0.254      0.56  0.57931
## Year2011          0.822      0.238      3.46  0.00075 ***
## Year2012          0.649      0.258      2.51  0.01325 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.537
## Multiple R-squared:  0.173, Adjusted R-squared:  0.0485
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 129 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.344  0.860  0.952  0.907  0.984  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      7.25e-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 1.825 1      1.351
## Year              1.825 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
## -1.1431 -0.3192 -0.0407 0.3787 1.6042
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.3007 0.1750 1.72 0.08825 .
## FirstAuthorFemale1 0.0836 0.1774 0.47 0.63836
## Year1997 0.2834 0.3470 0.82 0.41566
## Year1998 0.7822 0.2637 2.97 0.00364 **
## Year1999 0.6375 0.2421 2.63 0.00957 **
## Year2000 0.2543 0.3449 0.74 0.46237
## Year2001 0.8424 0.2960 2.85 0.00520 **
## Year2002 0.4488 0.2422 1.85 0.06641 .
## Year2003 0.7421 0.2562 2.90 0.00448 **
## Year2004 0.5843 0.2382 2.45 0.01562 *
## Year2005 0.4146 0.2243 1.85 0.06705 .
## Year2006 0.7051 0.2854 2.47 0.01488 *
```



```

## Year2007          0.8949      0.2082      4.30  3.5e-05 ***
## Year2008          0.6749      0.3769      1.79  0.07592 .
## Year2009          0.7835      0.2271      3.45  0.00077 ***
## Year2010          0.1071      0.2452      0.44  0.66303
## Year2011          0.7870      0.2327      3.38  0.00097 ***
## Year2012          0.6373      0.2581      2.47  0.01493 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.547
## Multiple R-squared:  0.166, Adjusted R-squared:  0.0476
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 125 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.370  0.862  0.955  0.907  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      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 1.48  1      1.217
## Year              1.48 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
## -1.139 -0.310 -0.059  0.379  1.606

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.279      0.178    1.57 0.11963
## LastAuthorFemale1 0.181      0.191    0.95 0.34584
## Year1997          0.306      0.349    0.88 0.38225
## Year1998          0.766      0.265    2.89 0.00458 **
## Year1999          0.675      0.253    2.67 0.00854 **
## Year2000          0.250      0.349    0.72 0.47518
## Year2001          0.860      0.290    2.96 0.00367 **
## Year2002          0.470      0.244    1.93 0.05651 .
## Year2003          0.739      0.243    3.04 0.00292 **
## Year2004          0.575      0.233    2.46 0.01529 *
## Year2005          0.445      0.233    1.91 0.05861 .
## Year2006          0.725      0.301    2.41 0.01757 *
## Year2007          0.906      0.206    4.39 2.5e-05 ***
## Year2008          0.710      0.381    1.87 0.06456 .
## Year2009          0.813      0.235    3.46 0.00075 ***
## Year2010          0.148      0.257    0.58 0.56569
## Year2011          0.825      0.238    3.47 0.00072 ***
## Year2012          0.643      0.248    2.59 0.01065 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.55
## Multiple R-squared:  0.171, Adjusted R-squared:  0.0531
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 128 ones are summarized as
##   Min. 1st Qu. Median      Mean 3rd Qu.      Max.
## 0.374 0.865 0.954 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      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] ""

```

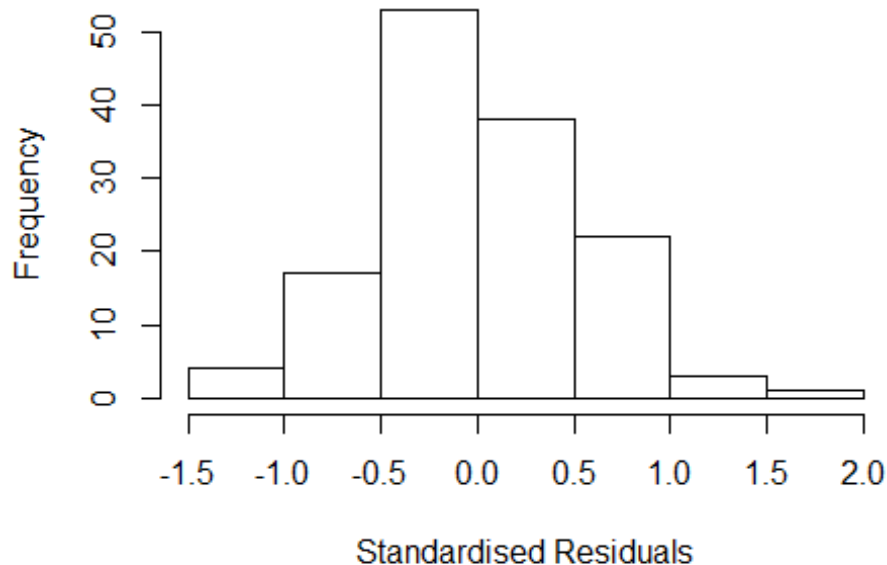
```

## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 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
##    2    2    4    3    4    5    9    4    8   11   11   11    9   21   17
## 2011 2012
##   17   11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    2    3    3    3    4    4    2    7    9   10    8    8   18   13
## 2011 2012
##   17    9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    2    3    3    3    4    3    2    7    8    8    8    8   16   11
## 2011 2012
##   16    8
## [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"

## 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 = NA
## 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"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale  2.282  1          1.511
## LastAuthorFemale  1.812  1          1.346
## UniqueAuthors    21.850  4          1.470
## Year              45.556 16          1.127
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##      Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##      k.max = 1000))
```

```

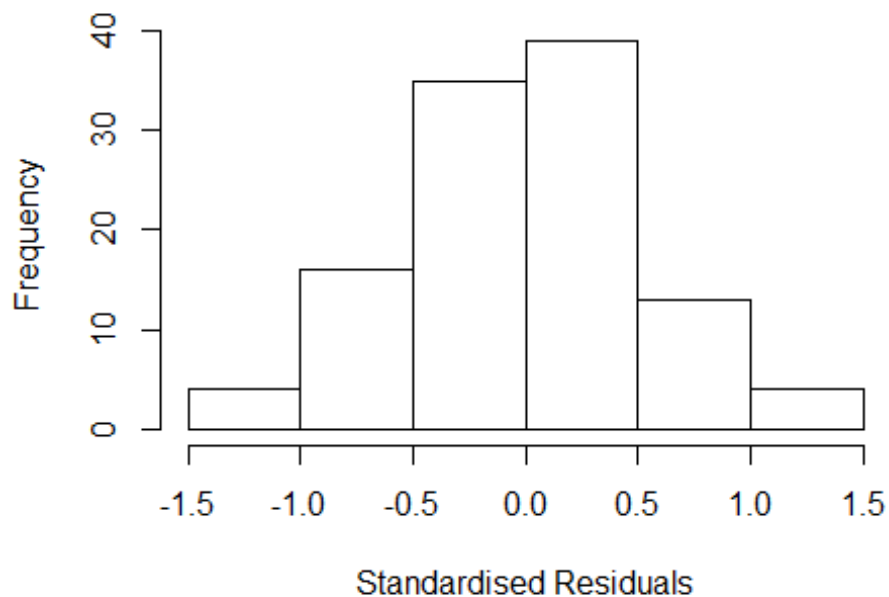
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.19222 -0.39391  0.00551  0.38341  1.33240
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    6.83e-01   1.08e-08  6.31e+07 < 2e-16 ***
## FirstAuthorFemale1 -4.89e-01   1.94e-01 -2.53e+00  0.0133 *
## LastAuthorFemale1  5.98e-02   1.63e-01  3.70e-01  0.7144
## UniqueAuthors2    -3.09e-02   1.42e-01 -2.20e-01  0.8285
## UniqueAuthors3     5.44e-03   1.75e-01  3.00e-02  0.9753
## UniqueAuthors4     3.35e-01   3.00e-01  1.12e+00  0.2678
## UniqueAuthors5     1.62e-01   3.94e-01  4.10e-01  0.6825
## Year1997          7.63e-01   1.74e-01  4.40e+00  3.1e-05 ***
## Year1998          2.92e-01   4.12e-01  7.10e-01  0.4809
## Year1999          2.05e-02   3.15e-01  7.00e-02  0.9481
## Year2000          3.47e-02   2.33e-01  1.50e-01  0.8821
## Year2001          6.56e-01   2.09e-01  3.14e+00  0.0023 **
## Year2002         -7.78e-02   5.63e-01 -1.40e-01  0.8905
## Year2003          4.01e-01   2.43e-01  1.65e+00  0.1031
## Year2004          6.78e-01   2.15e-01  3.16e+00  0.0022 **
## Year2005          5.15e-01   2.40e-01  2.15e+00  0.0347 *
## Year2006          6.38e-02   1.60e-01  4.00e-01  0.6909
## Year2007         -7.02e-02   1.63e-01 -4.30e-01  0.6668
## Year2008          1.74e-01   3.02e-01  5.80e-01  0.5649
## Year2009          1.41e-01   2.09e-01  6.70e-01  0.5042
## Year2010          9.51e-02   1.90e-01  5.00e-01  0.6179
## Year2011          3.78e-01   1.91e-01  1.98e+00  0.0506 .
## Year2012          3.75e-01   2.50e-01  1.50e+00  0.1360
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.602
## Multiple R-squared:  0.216, Adjusted R-squared:  0.0206
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 100 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.604  0.876  0.956  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      9.01e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500            50          2            1            1000      200

```

```
## trace.lev      mts compute.rd
##           0      1000         0
##           psi      subsampling      cov
##           "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"

## 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 1.585 1 1.259
## LastAuthorFemale 1.730 1 1.315
## Year 2.736 16 1.032

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
```

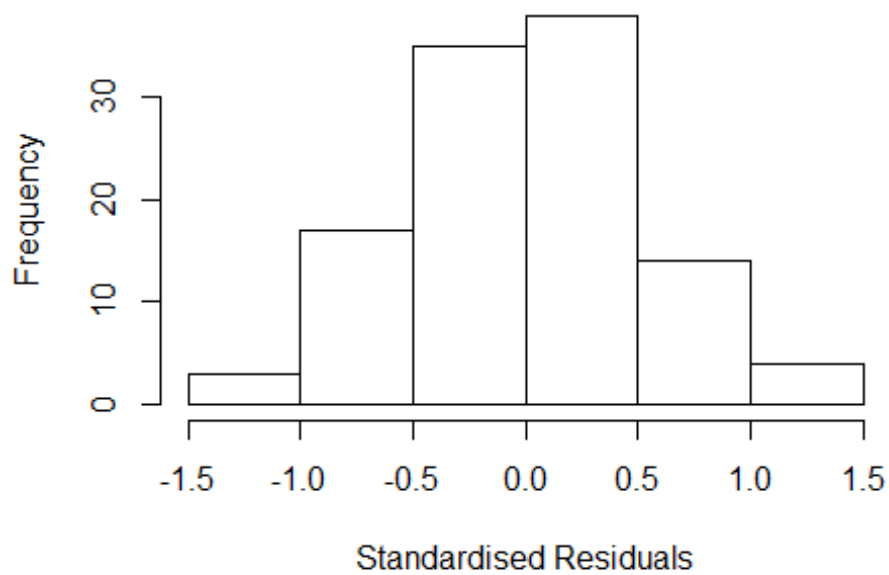
```

## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.11024 -0.35436  0.00295  0.35492  1.31952
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.6830     0.0000    Inf < 2e-16 ***
## FirstAuthorFemale1 -0.5133     0.1935   -2.65  0.0094 **
## LastAuthorFemale1  0.0470     0.1729    0.27  0.7866
## Year1997         0.7655     0.1479    5.18 1.3e-06 ***
## Year1998         0.2861     0.4160    0.69  0.4933
## Year1999         0.0238     0.3098    0.08  0.9390
## Year2000         0.0861     0.1697    0.51  0.6133
## Year2001         0.6918     0.1703    4.06  0.0001 ***
## Year2002        -0.0279     0.4986   -0.06  0.9555
## Year2003         0.3880     0.1883    2.06  0.0422 *
## Year2004         0.7233     0.1702    4.25 5.1e-05 ***
## Year2005         0.5231     0.2062    2.54  0.0129 *
## Year2006         0.0564     0.1711    0.33  0.7426
## Year2007        -0.0832     0.1537   -0.54  0.5897
## Year2008         0.2495     0.3032    0.82  0.4128
## Year2009         0.1377     0.2137    0.64  0.5210
## Year2010         0.1216     0.1570    0.77  0.4406
## Year2011         0.4272     0.1623    2.63  0.0099 **
## Year2012         0.3571     0.2315    1.54  0.1264
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.588
## Multiple R-squared:  0.204, Adjusted R-squared:  0.0487
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 102 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.594  0.864  0.963  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          9.01e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats

```

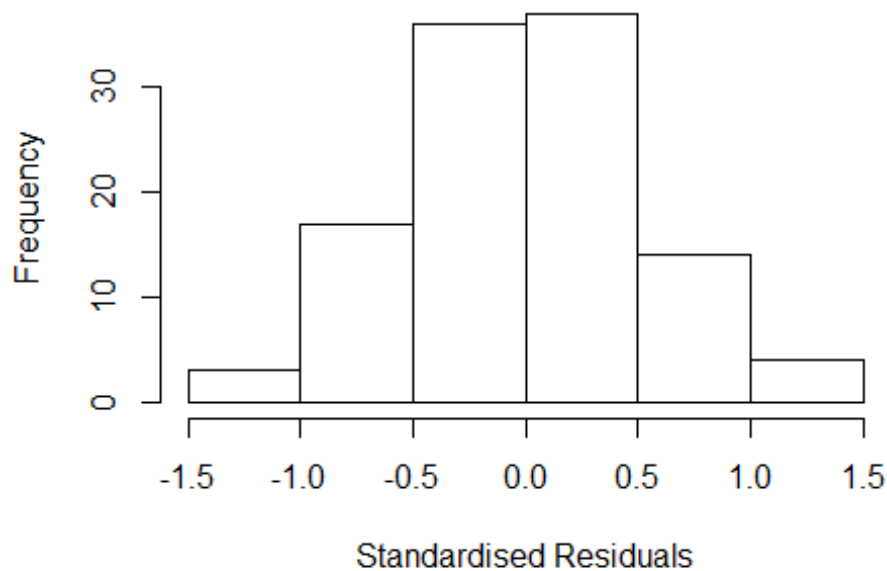
```
## "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead
```

Residuals from first and last author



```
## GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.59 1 1.261
## Year 1.59 16 1.015
```


Residuals from first author



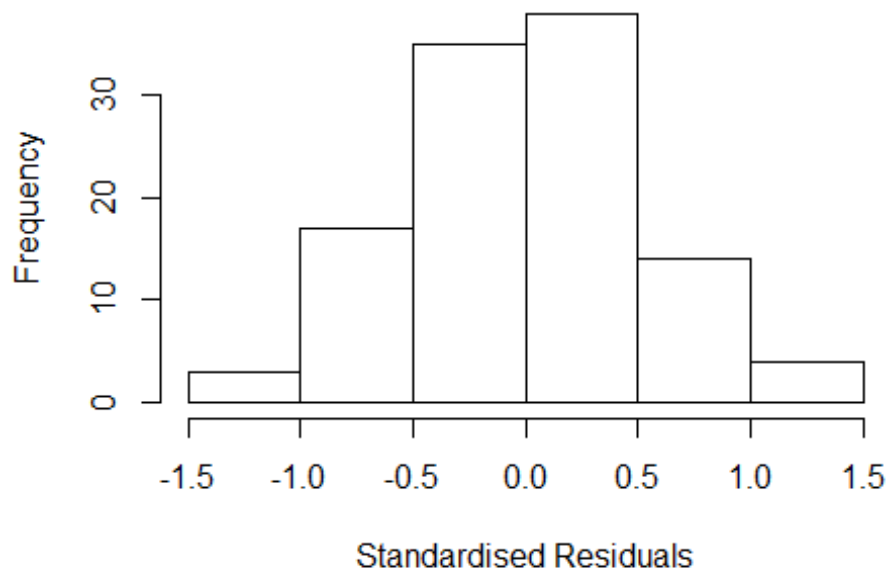
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.11e+00 -3.34e-01 -3.33e-16 3.56e-01 1.31e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6830 0.0000 Inf < 2e-16 ***
## FirstAuthorFemale1 -0.4939 0.2092 -2.36 0.0203 *
## Year1997 0.7655 0.1478 5.18 1.3e-06 ***
## Year1998 0.2809 0.4174 0.67 0.5026
## Year1999 0.0228 0.3087 0.07 0.9412
## Year2000 0.0797 0.1748 0.46 0.6495
## Year2001 0.6917 0.1701 4.07 1.0e-04 ***
## Year2002 -0.0251 0.4875 -0.05 0.9591
## Year2003 0.3880 0.1881 2.06 0.0419 *
## Year2004 0.7219 0.1706 4.23 5.4e-05 ***
## Year2005 0.5346 0.1906 2.80 0.0061 **
## Year2006 0.0635 0.1677 0.38 0.7060
```

```

## Year2007          -0.0766      0.1492   -0.51   0.6088
## Year2008          0.2548      0.2972    0.86   0.3935
## Year2009          0.1453      0.2107    0.69   0.4924
## Year2010          0.1308      0.1442    0.91   0.3667
## Year2011          0.4261      0.1622    2.63   0.0101 *
## Year2012          0.3563      0.2305    1.55   0.1255
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.599
## Multiple R-squared:  0.203, Adjusted R-squared:  0.0569
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 101 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.609  0.863  0.964  0.922  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      9.01e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.658 1          1.288
## Year            1.658 16          1.016

```

Residuals from last author



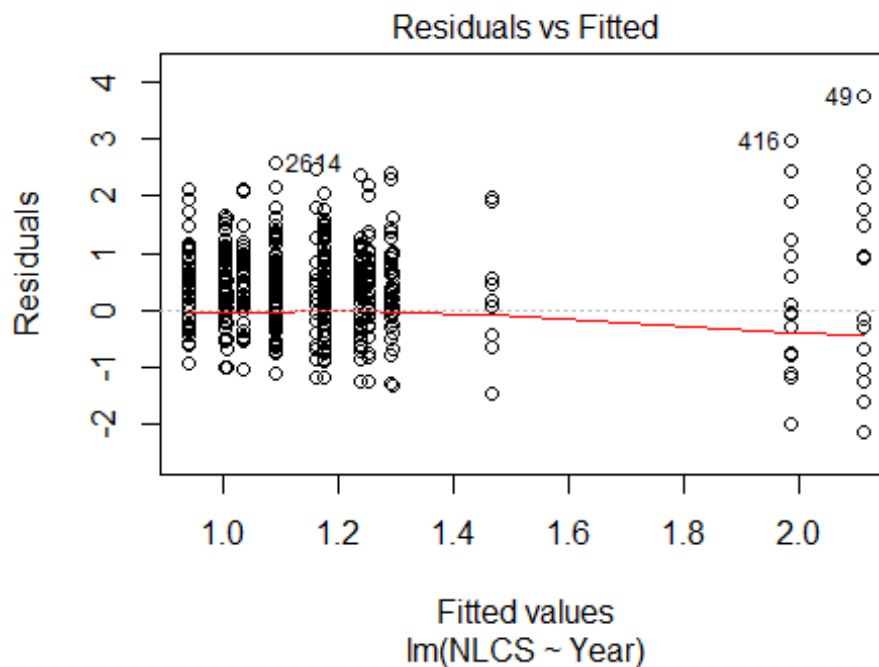
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.11e+00 -3.30e-01 -1.22e-15 3.78e-01 1.43e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 6.83e-01 5.34e-08 1.28e+07 < 2e-16 ***
## LastAuthorFemale1 -1.38e-01 1.77e-01 -7.80e-01 0.4372
## Year1997 7.65e-01 1.48e-01 5.18e+00 1.3e-06 ***
## Year1998 1.04e-01 5.93e-01 1.80e-01 0.8609
## Year1999 2.41e-02 3.10e-01 8.00e-02 0.9382
## Year2000 -7.57e-02 2.95e-01 -2.60e-01 0.7984
## Year2001 6.92e-01 1.70e-01 4.06e+00 0.0001 ***
## Year2002 -2.88e-02 5.00e-01 -6.00e-02 0.9542
## Year2003 3.88e-01 1.88e-01 2.06e+00 0.0423 *
## Year2004 7.24e-01 1.70e-01 4.26e+00 5.0e-05 ***
## Year2005 4.31e-01 2.38e-01 1.81e+00 0.0735 .
## Year2006 8.12e-02 1.64e-01 5.00e-01 0.6215
```

```

## Year2007          -5.81e-02   1.49e-01 -3.90e-01   0.6972
## Year2008          2.02e-01   3.30e-01  6.10e-01   0.5429
## Year2009          8.58e-03   1.96e-01  4.00e-02   0.9651
## Year2010          1.61e-01   1.55e-01  1.04e+00   0.3019
## Year2011          3.77e-01   1.67e-01  2.26e+00   0.0261 *
## Year2012          3.57e-01   2.32e-01  1.54e+00   0.1266
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.585
## Multiple R-squared:  0.161, Adjusted R-squared:  0.00776
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 102 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.527  0.872  0.959  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      9.01e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 111"
## [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
## 125 142 56 97 14 118 129 126 138 154 141 194 221 316 308
## 2011 2012
## 337 318
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 17 20 12 18 0 12 28 49 57 77 71 94 80 136 88
## 2011 2012

```

```
## 114 121
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 14 13 9 15 0 8 20 22 26 29 51 68 57 92 78
## 2011 2012
## 102 108
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 76, df = 15, p-value = 5e-10
```



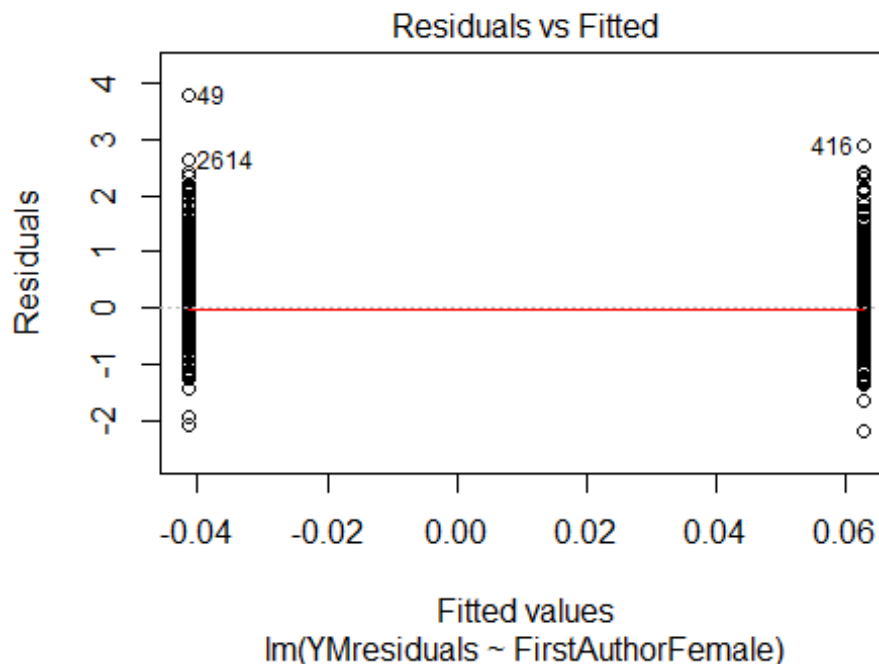
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.4, df = 1, p-value = 0.5

## [1] "Female first author team size 2018 geometric mean: 3.92465602673644"
## [1] "Male first author team size 2018 geometric mean: 3.49296539627053"

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

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.65254580478304"
## [1] "Male last author team size 2018 geometric mean: 3.71591294296248"
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 840, 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"

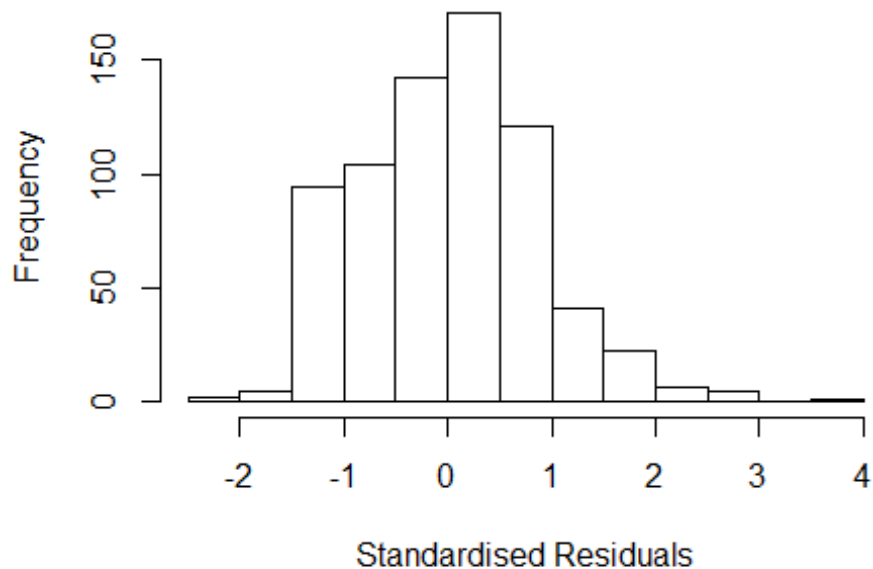
## Warning in lmrob.fit(x, y, control, init = init, mf = mf): M-step did NOT
## converge. Returning unconverged SM-estimate
```



```
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.600 1          1.265
```

| | | | |
|---------------------|-------|----|-------|
| ## LastAuthorFemale | 1.416 | 1 | 1.190 |
| ## Year | 1.430 | 15 | 1.012 |

Residuals from first and last author



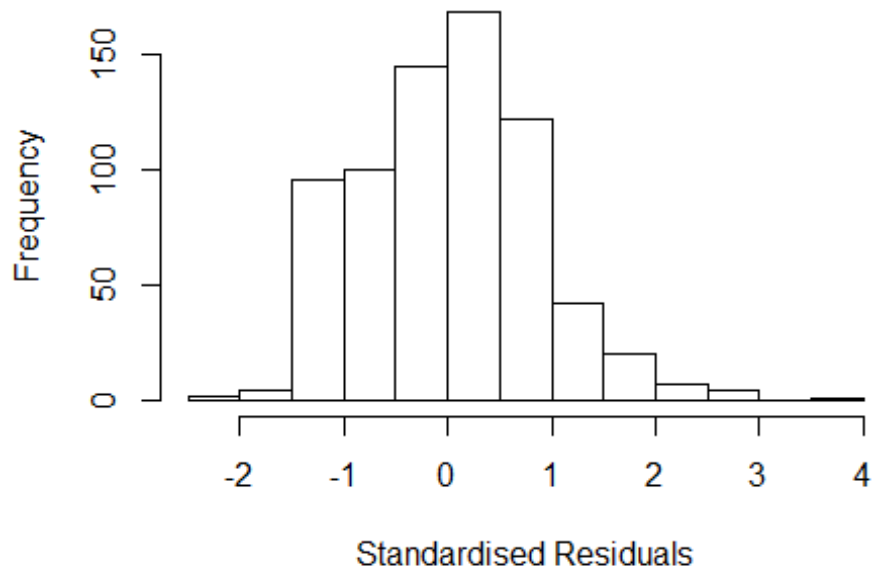
```
## [1] "List of 5 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 49      0029737381 5.850 1996      2700      1      3.710
## 416     0033594380 4.942 1999      2700      1      2.871
## 697     0037006247 3.626 2002      2700      1      2.526
## 786     0141677640 3.590 2003      2700      1      2.506
## 2614    84455173775 3.676 2011      2700      1      2.621
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2470 -0.5696  0.0175  0.5474  3.7098
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.1402     1.3555   1.58    0.11
## FirstAuthorFemale1  0.1067     0.0807   1.32    0.19
## LastAuthorFemale1 -0.0506     0.0799  -0.63    0.53
## Year1997          -1.1381     1.3918  -0.82    0.41
## Year1998           -0.6643     1.4425  -0.46    0.65
```

```

## Year1999          -0.1762      1.4683    -0.12      0.90
## Year2001          -1.1292      1.3870    -0.81      0.42
## Year2002          -1.1469      1.3786    -0.83      0.41
## Year2003          -1.0564      1.3860    -0.76      0.45
## Year2004          -1.0008      1.3734    -0.73      0.47
## Year2005          -0.7992      1.3735    -0.58      0.56
## Year2006          -0.9414      1.3671    -0.69      0.49
## Year2007          -0.8915      1.3694    -0.65      0.52
## Year2008          -1.1191      1.3692    -0.82      0.41
## Year2009          -1.1226      1.3710    -0.82      0.41
## Year2010          -1.0172      1.3705    -0.74      0.46
## Year2011          -1.0855      1.3668    -0.79      0.43
## Year2012          -1.3104      1.3669    -0.96      0.34
##
## Robust residual standard error: 0.837
## Multiple R-squared:  0.0705, Adjusted R-squared:  0.0477
## Convergence in 48 IRWLS iterations
##
## Robustness weights:
## 62 weights are ~= 1. The remaining 650 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0111 0.8700 0.9520 0.9110 0.9860 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.40e-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 1.271 1          1.127
## Year              1.271 15          1.008

```


Residuals from first author



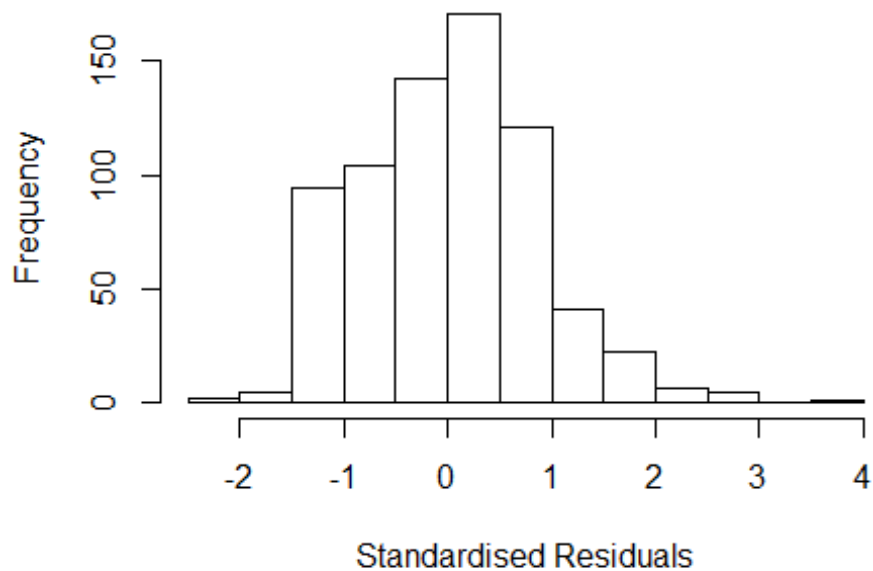
```
## [1] "List of 5 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 49      0029737381 5.850 1996      2700      1      3.710
## 416     0033594380 4.942 1999      2700      1      2.871
## 697     0037006247 3.626 2002      2700      1      2.526
## 786     0141677640 3.590 2003      2700      1      2.506
## 2614    84455173775 3.676 2011      2700      1      2.621
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2235 -0.5880  0.0133  0.5473  3.7125
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.138      1.356    1.58   0.12
## FirstAuthorFemale1  0.086      0.072    1.20   0.23
## Year1997        -1.135      1.393   -0.81   0.42
## Year1998        -0.662      1.444   -0.46   0.65
## Year1999        -0.192      1.469   -0.13   0.90
## Year2001        -1.128      1.388   -0.81   0.42
## Year2002        -1.153      1.380   -0.84   0.40
## Year2003        -1.065      1.388   -0.77   0.44
```

```

## Year2004          -1.003      1.375    -0.73     0.47
## Year2005          -0.802      1.375    -0.58     0.56
## Year2006          -0.942      1.368    -0.69     0.49
## Year2007          -0.899      1.371    -0.66     0.51
## Year2008          -1.124      1.371    -0.82     0.41
## Year2009          -1.125      1.372    -0.82     0.41
## Year2010          -1.020      1.372    -0.74     0.46
## Year2011          -1.089      1.368    -0.80     0.43
## Year2012          -1.314      1.368    -0.96     0.34
##
## Robust residual standard error: 0.837
## Multiple R-squared:  0.0691, Adjusted R-squared:  0.0477
## Convergence in 48 IRWLS iterations
##
## Robustness weights:
## 57 weights are ~= 1. The remaining 655 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0106 0.8710 0.9530 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      1.40e-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 1.118 1      1.057
## Year      1.118 15      1.004

```

Residuals from last author



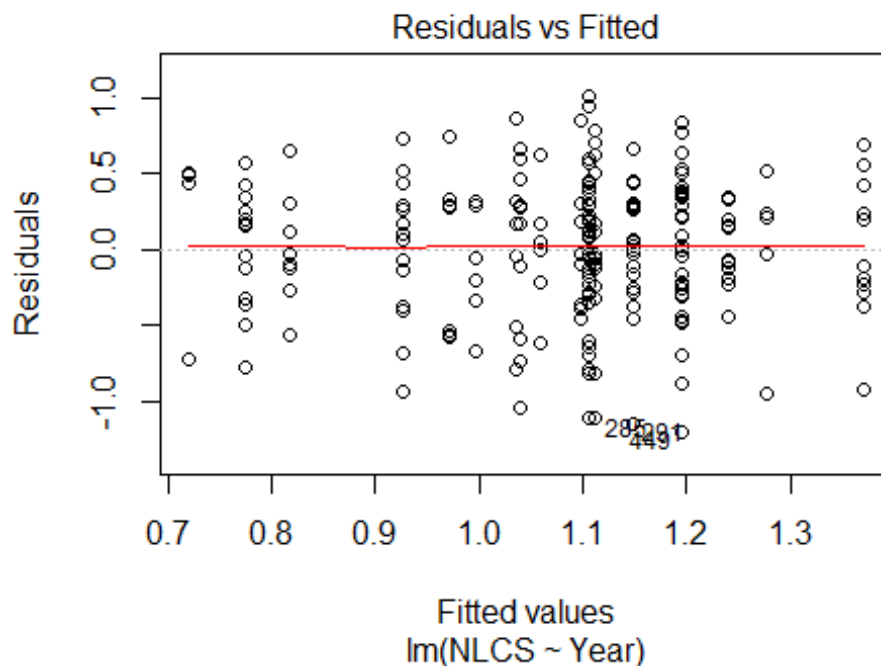
```
## [1] "List of 5 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 49      0029737381 5.850 1996      2700      1      3.710
## 416     0033594380 4.942 1999      2700      1      2.871
## 697     0037006247 3.626 2002      2700      1      2.526
## 786     0141677640 3.590 2003      2700      1      2.506
## 2614    84455173775 3.676 2011      2700      1      2.621
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q      Median        3Q       Max
## -2.100259 -0.548796 -0.000241  0.560661  3.749741
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.10026    1.22978    1.71   0.088 .
## LastAuthorFemale1 -0.00065    0.07059   -0.01   0.993
## Year1997       -1.06310    1.26230   -0.84   0.400
## Year1998       -0.62745    1.33138   -0.47   0.638
## Year1999       -0.15455    1.37043   -0.11   0.910
## Year2001       -1.06512    1.25880   -0.85   0.398
## Year2002       -1.08913    1.25209   -0.87   0.385
## Year2003       -0.98695    1.25546   -0.79   0.432
```

```

## Year2004          -0.94399      1.24489      -0.76      0.449
## Year2005          -0.72588      1.23968      -0.59      0.558
## Year2006          -0.87895      1.23574      -0.71      0.477
## Year2007          -0.82530      1.23686      -0.67      0.505
## Year2008          -1.05716      1.23849      -0.85      0.394
## Year2009          -1.04256      1.23449      -0.84      0.399
## Year2010          -0.93960      1.23483      -0.76      0.447
## Year2011          -1.01981      1.23463      -0.83      0.409
## Year2012          -1.24104      1.23362      -1.01      0.315
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.837
## Multiple R-squared:  0.0638, Adjusted R-squared:  0.0422
## Convergence in 47 IRWLS iterations
##
## Robustness weights:
## 67 weights are ~= 1. The remaining 645 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0074 0.8640 0.9520 0.9100 0.9840 0.9990
## Algorithmic parameters:
##           tuning.chi           bb           tuning.psi           refine.tol
##           1.55e+00           5.00e-01           4.69e+00           1.00e-07
##           rel.tol           solve.tol           eps.outlier           eps.x
##           1.00e-07           1.00e-07           1.40e-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: 712"
## [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
##   10   12   16   21   15   16   15   16   11   20   14   18   13   23   25
## 2011 2012
##   57   44
##

```

```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    7   10   11    5    7    8   12    6   14    8   13    8   15   19
## 2011 2012
##   38   32
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    7    8   10    4    6    8   10    5   12    6   12    8   14   14
## 2011 2012
##   33   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 = 13, df = 16, p-value = 0.6
```

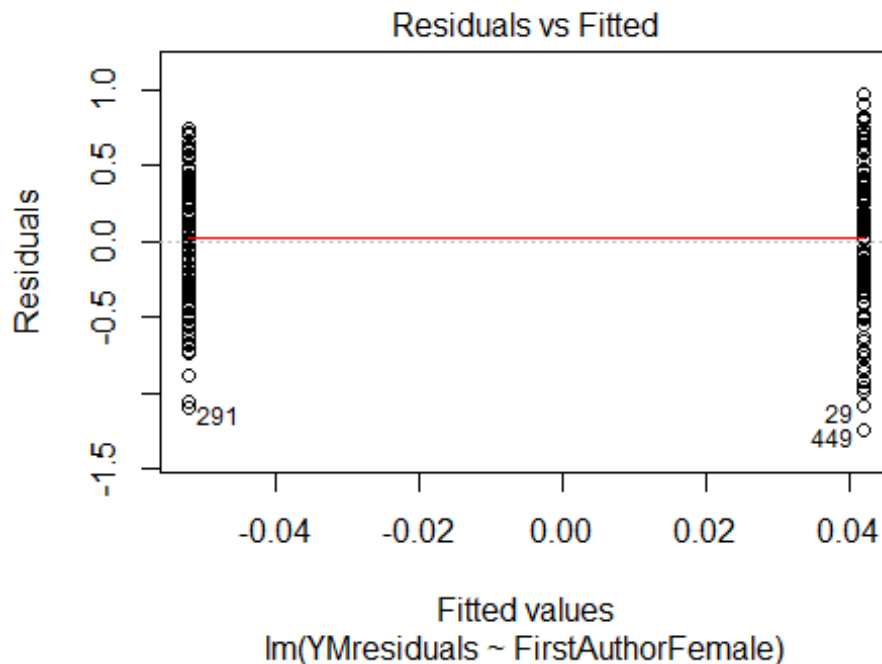


```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.42, df = 1, p-value = 0.5
## [1] "Female first author team size 2018 geometric mean: 4.98868342635275"
## [1] "Male first author team size 2018 geometric mean: 6.09508497056848"
```

```
## Warning in wilcox.test.default(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 = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.50566598012188"
## [1] "Male last author team size 2018 geometric mean: 5.12557525317945"

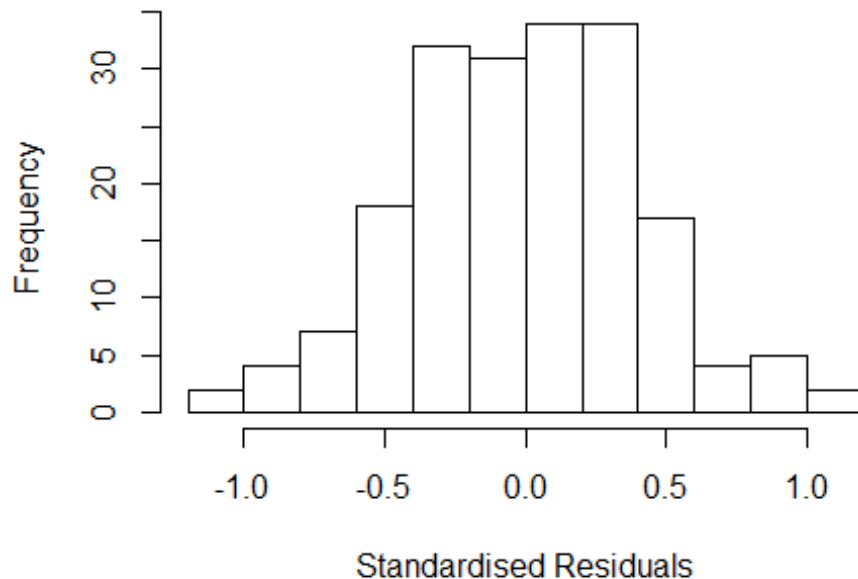
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## 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 1.926 1          1.388
```

| | | | |
|---------------------|-------|----|-------|
| ## LastAuthorFemale | 1.992 | 1 | 1.411 |
| ## UniqueAuthors | 4.089 | 4 | 1.192 |
| ## Year | 9.773 | 16 | 1.074 |

Residuals from first and last author and team size



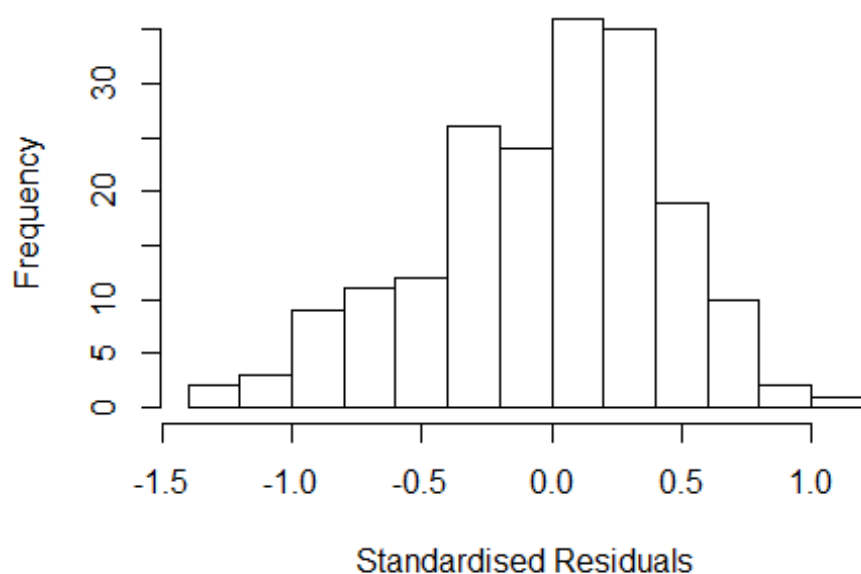
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.12855 -0.27438  0.00445  0.26894  1.18991
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.3984    0.1937    2.06  0.0413 *
## FirstAuthorFemale1 -0.0283    0.0744   -0.38  0.7037
## LastAuthorFemale1  0.1208    0.0798    1.51  0.1317
## UniqueAuthors2    0.5378    0.0957    5.62 7.8e-08 ***
## UniqueAuthors3    0.4959    0.1053    4.71 5.2e-06 ***
## UniqueAuthors4    0.4543    0.1462    3.11  0.0022 **
```

```

## UniqueAuthors5      0.6370      0.1090      5.84 2.6e-08 ***
## Year1997             0.2687      0.2388      1.13 0.2620
## Year1998             0.3802      0.2480      1.53 0.1272
## Year1999             0.4737      0.2729      1.74 0.0844 .
## Year2000             0.4022      0.2816      1.43 0.1550
## Year2001             0.3102      0.2195      1.41 0.1594
## Year2002             0.1044      0.2618      0.40 0.6906
## Year2003             0.3514      0.1980      1.78 0.0777 .
## Year2004             0.0805      0.2711      0.30 0.7667
## Year2005             0.0460      0.2095      0.22 0.8266
## Year2006             0.3747      0.2387      1.57 0.1184
## Year2007             0.0785      0.2326      0.34 0.7361
## Year2008             0.2149      0.2018      1.07 0.2884
## Year2009             0.2057      0.2399      0.86 0.3925
## Year2010             0.2758      0.2127      1.30 0.1965
## Year2011             0.2015      0.2006      1.00 0.3166
## Year2012             0.3106      0.2158      1.44 0.1519
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.422
## Multiple R-squared:  0.317, Adjusted R-squared:  0.227
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 175 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.407  0.892  0.957  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      5.26e-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 1.524 1      1.234
## LastAuthorFemale  2.517 1      1.587
## Year              3.281 16      1.038

```


Residuals from first and last author



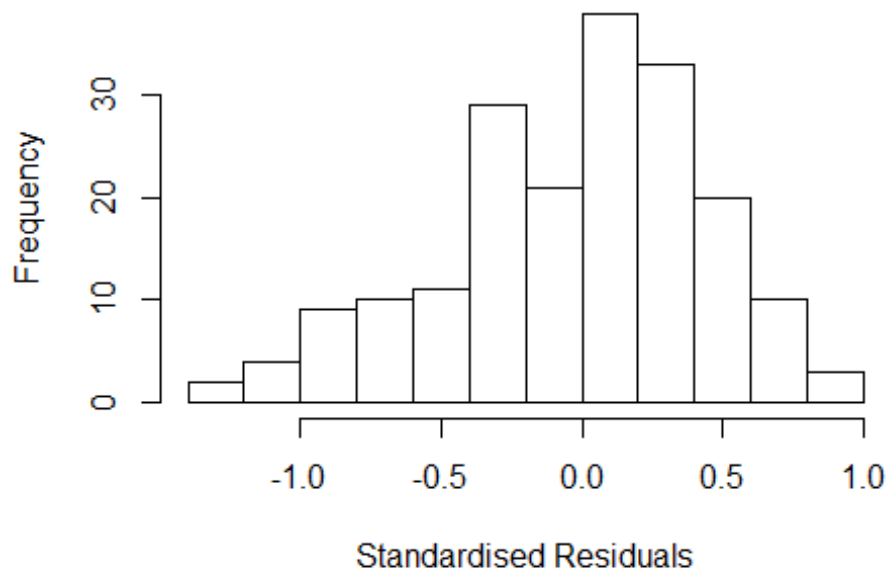
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3238 -0.2923 0.0409 0.3102 1.0012
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.768376 0.409581 1.88 0.062 .
## FirstAuthorFemale1 0.089162 0.079812 1.12 0.265
## LastAuthorFemale1 0.038947 0.105481 0.37 0.712
## Year1997 0.180630 0.461270 0.39 0.696
## Year1998 0.427317 0.459577 0.93 0.354
## Year1999 0.567527 0.454552 1.25 0.214
## Year2000 0.491721 0.517623 0.95 0.343
## Year2001 0.306279 0.461779 0.66 0.508
## Year2002 -0.000879 0.436626 0.00 0.998
## Year2003 0.399644 0.427640 0.93 0.351
## Year2004 0.104286 0.508749 0.20 0.838
## Year2005 0.043979 0.436395 0.10 0.920
```

```

## Year2006          0.324590    0.454747    0.71    0.476
## Year2007          -0.009370    0.435846   -0.02    0.983
## Year2008          0.175331    0.436834    0.40    0.689
## Year2009          0.266273    0.451539    0.59    0.556
## Year2010          0.228944    0.433398    0.53    0.598
## Year2011          0.256307    0.431695    0.59    0.553
## Year2012          0.413258    0.422709    0.98    0.330
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.454
## Multiple R-squared:  0.116, Adjusted R-squared:  0.0229
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 175 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.376  0.873   0.951   0.905   0.984   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      5.26e-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 1.468 1      1.212
## Year              1.468 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
## -1.2831 -0.2958 0.0427 0.3278 0.9783
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.76019 0.37883 2.01 0.046 *
## FirstAuthorFemale1 0.09997 0.07854 1.27 0.205
## Year1997 0.21517 0.41159 0.52 0.602
## Year1998 0.41307 0.43978 0.94 0.349
## Year1999 0.57610 0.42667 1.35 0.179
## Year2000 0.48422 0.49637 0.98 0.331
## Year2001 0.31027 0.43526 0.71 0.477
## Year2002 0.00946 0.40466 0.02 0.981
## Year2003 0.41231 0.39491 1.04 0.298
## Year2004 0.12311 0.47622 0.26 0.796
## Year2005 0.05982 0.40553 0.15 0.883
## Year2006 0.33586 0.42248 0.79 0.428
```

```

## Year2007          0.00883    0.40006    0.02    0.982
## Year2008          0.18062    0.40754    0.44    0.658
## Year2009          0.28307    0.41778    0.68    0.499
## Year2010          0.24227    0.40009    0.61    0.546
## Year2011          0.27657    0.39544    0.70    0.485
## Year2012          0.42290    0.39389    1.07    0.284
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.475
## Multiple R-squared:  0.111, Adjusted R-squared:  0.0231
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 176 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.446  0.882  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      5.26e-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 2.443  1          1.563
## Year              2.443 16          1.028
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.329 -0.293  0.014  0.315  1.056

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.7730    0.3953    1.96   0.052 .
## LastAuthorFemale1 0.0706    0.1028    0.69   0.493
## Year1997        0.2209    0.4435    0.50   0.619
## Year1998        0.4857    0.4380    1.11   0.269
## Year1999        0.6112    0.4289    1.43   0.156
## Year2000        0.5203    0.5017    1.04   0.301
## Year2001        0.3136    0.4548    0.69   0.491
## Year2002        0.0286    0.4178    0.07   0.946
## Year2003        0.4391    0.4070    1.08   0.282
## Year2004        0.1368    0.4987    0.27   0.784
## Year2005        0.0709    0.4199    0.17   0.866
## Year2006        0.3688    0.4423    0.83   0.406
## Year2007        0.0272    0.4172    0.07   0.948
## Year2008        0.2270    0.4162    0.55   0.586
## Year2009        0.3185    0.4315    0.74   0.461
## Year2010        0.2795    0.4124    0.68   0.499
## Year2011        0.2863    0.4146    0.69   0.491
## Year2012        0.4291    0.4077    1.05   0.294
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.451
## Multiple R-squared:  0.11, Adjusted R-squared:  0.0216
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 173 ones are summarized as
##   Min. 1st Qu. Median   Mean 3rd Qu.   Max.
##   0.364 0.870 0.951 0.902 0.983 0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      5.26e-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: 190"
## [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
##    2    4    3    3    3    7    5    9    7    9    7    6    7    7    5
## 2011 2012
##    3    2
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    1    1    1    3    1    5    3    6    5    4    4    6    2
## 2011 2012
##    3    2
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    1    1    0    2    0    4    2    5    3    2    4    5    1
## 2011 2012
##    2    2
## [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: 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 = 0, 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: 2.51984209978975"

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

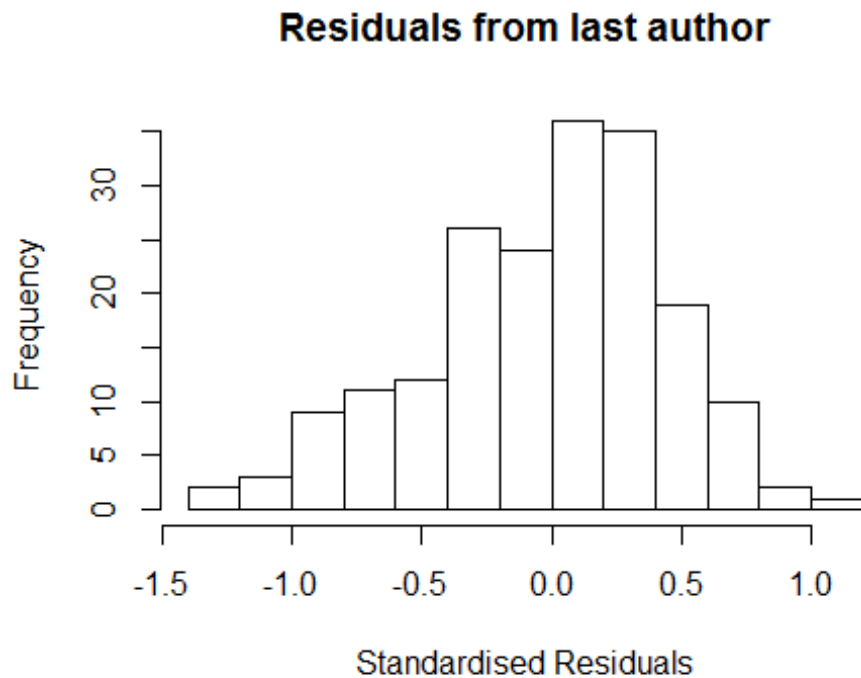
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"

```

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

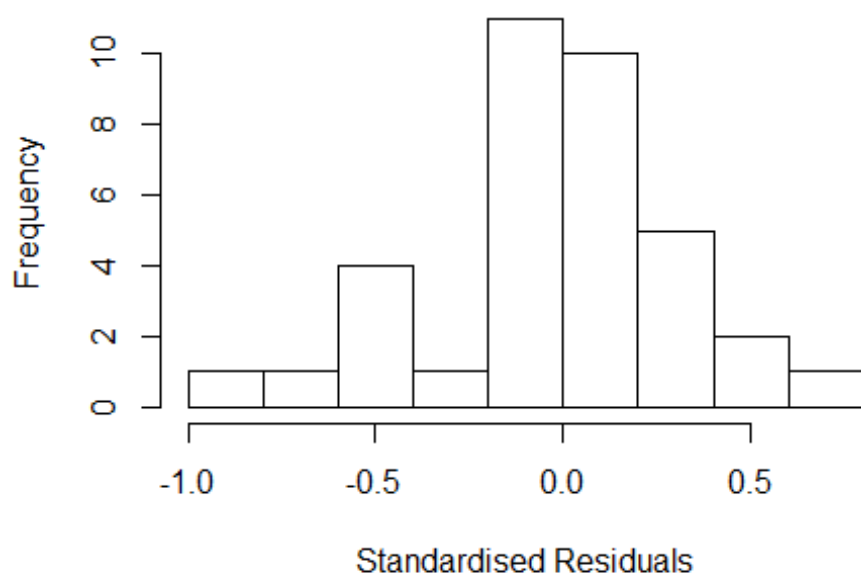
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

## 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 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
## -0.8847 -0.1328 0.0136 0.1776 0.6867
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 9.12e-01 6.06e-09 1.50e+08 <2e-16 ***
## FirstAuthorFemale1 -2.13e-01 3.09e-01 -6.90e-01 0.498
## LastAuthorFemale1 1.09e-01 1.89e-01 5.70e-01 0.572
## Year1997 3.24e-01 0.00e+00 Inf <2e-16 ***
## Year1998 3.31e-01 2.42e-01 1.37e+00 0.187
## Year1999 1.10e-02 0.00e+00 Inf <2e-16 ***
## Year2001 -2.13e-01 1.28e-01 -1.67e+00 0.112
## Year2003 -4.83e-01 1.78e-01 -2.72e+00 0.014 *
## Year2004 -1.14e-01 3.10e-01 -3.70e-01 0.717
## Year2005 1.61e-01 2.86e-01 5.60e-01 0.581
## Year2006 -3.46e-01 3.30e-01 -1.05e+00 0.308
## Year2007 9.64e-02 2.65e-01 3.60e-01 0.721
```



```

## Year2008          3.57e-01    2.01e-01    1.77e+00    0.092 .
## Year2009          1.80e-01    3.80e-01    4.70e-01    0.641
## Year2010          3.55e-01    3.09e-01    1.15e+00    0.264
## Year2011          2.65e-01    5.34e-01    5.00e-01    0.625
## Year2012          2.97e-01    1.82e-01    1.63e+00    0.119
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.393
## Multiple R-squared:  0.417, Adjusted R-squared:  -0.0745
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 29 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.591  0.878  0.976  0.922  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      2.78e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -9.758e+15  1      NaN
## Year              -9.758e+15 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
## -0.8021 -0.1348  0.0217  0.1749  0.6969
##

```

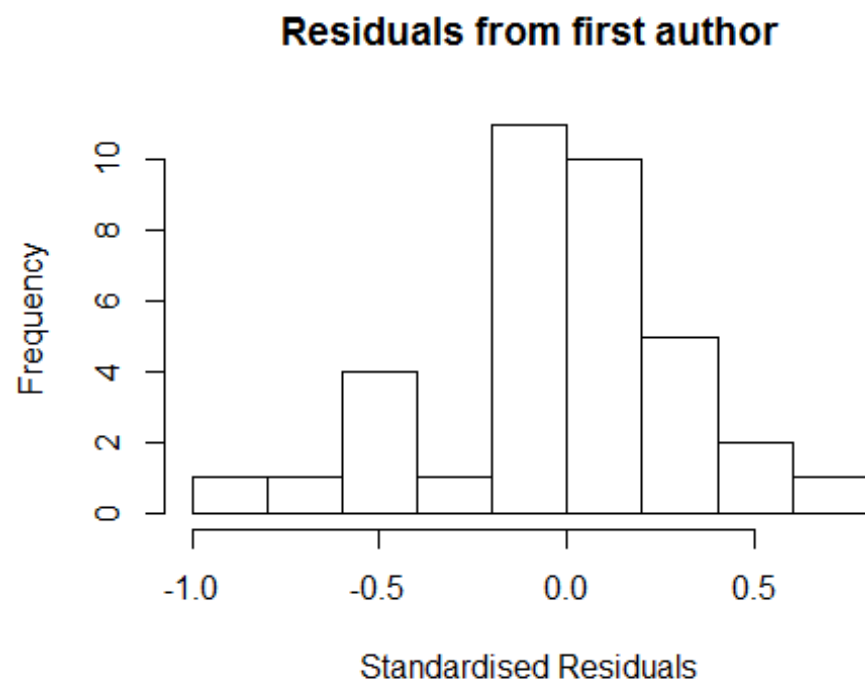
```

## Coefficients:
##               Estimate Std. Error  t value Pr(>|t|)
## (Intercept)      9.12e-01   1.63e-08  5.59e+07  <2e-16 ***
## FirstAuthorFemale1 -1.47e-01   2.19e-01 -6.70e-01   0.510
## Year1997          3.24e-01   1.10e-08  2.93e+07  <2e-16 ***
## Year1998          3.74e-01   2.19e-01  1.71e+00   0.103
## Year1999          1.10e-02   1.46e-08  7.53e+05  <2e-16 ***
## Year2001         -1.59e-01   1.26e-01 -1.26e+00   0.222
## Year2003         -4.84e-01   1.71e-01 -2.83e+00   0.010 *
## Year2004         -1.80e-01   2.21e-01 -8.20e-01   0.424
## Year2005          1.87e-01   2.82e-01  6.60e-01   0.515
## Year2006         -3.08e-01   2.98e-01 -1.04e+00   0.312
## Year2007          8.46e-02   2.36e-01  3.60e-01   0.724
## Year2008          3.69e-01   1.82e-01  2.03e+00   0.056 .
## Year2009          1.70e-01   3.34e-01  5.10e-01   0.617
## Year2010          2.89e-01   2.19e-01  1.32e+00   0.202
## Year2011          2.87e-01   4.89e-01  5.90e-01   0.564
## Year2012          3.19e-01   1.33e-01  2.40e+00   0.026 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.472
## Multiple R-squared:  0.389, Adjusted R-squared:  -0.0685
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 27 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.754  0.905  0.975  0.941  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           2.78e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01           5.00e-01
##   nResample    max.it    best.r.s    k.fast.s    k.max maxit.scale
##           500         50         2         1         1000         200
##   trace.lev    mts    compute.rd
##           0         1000         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

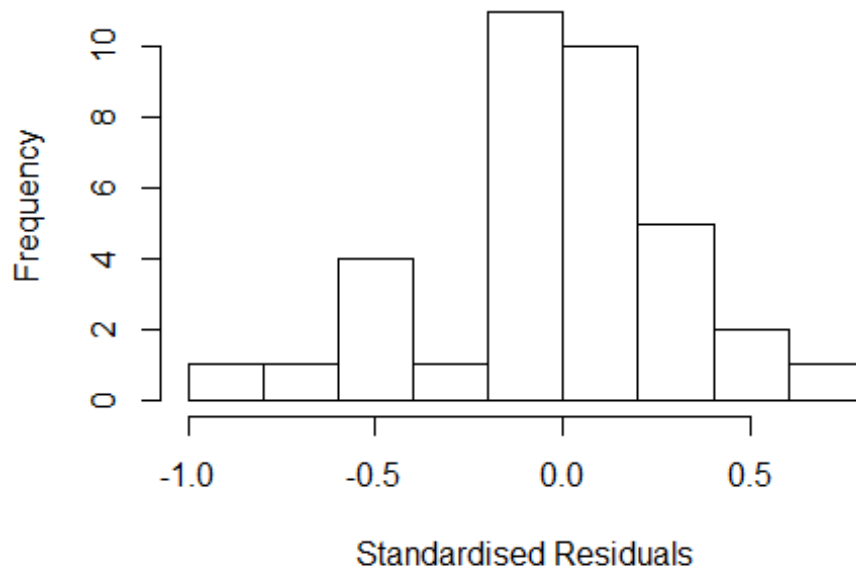
```

```
## 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 14          NaN
```

Residuals from last author



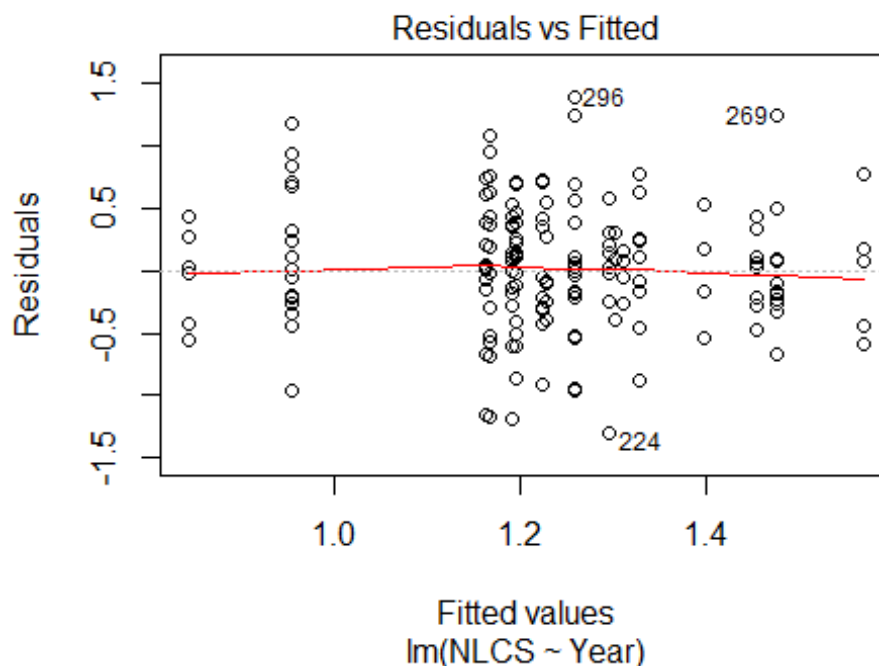
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -7.28e-01 -1.28e-01 2.22e-16 1.70e-01 7.75e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9120 0.0000 Inf < 2e-16 ***
## LastAuthorFemale1 0.0146 0.1397 0.10 0.9175
## Year1997 0.3240 0.0000 Inf < 2e-16 ***
## Year1998 0.2124 0.1397 1.52 0.1441
## Year1999 0.0110 0.0000 Inf < 2e-16 ***
## Year2001 -0.1663 0.1388 -1.20 0.2449
## Year2003 -0.4845 0.1680 -2.88 0.0092 **
## Year2004 -0.3275 0.0308 -10.63 1.1e-09 ***
## Year2005 0.0988 0.2015 0.49 0.6291
## Year2006 -0.4725 0.2378 -1.99 0.0608 .
## Year2007 -0.0698 0.1078 -0.65 0.5246
## Year2008 0.2891 0.1968 1.47 0.1574
```

```

## Year2009          0.0917      0.2884      0.32      0.7539
## Year2010          0.1420      0.0000      Inf      < 2e-16 ***
## Year2011          0.2057      0.3618      0.57      0.5760
## Year2012          0.2377      0.0752      3.16      0.0049 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.535
## Multiple R-squared:  0.368, Adjusted R-squared:  -0.106
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 24 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.818  0.939  0.954  0.947  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           2.78e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01           5.00e-01
##   nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500          50          2          1          1000          200
##   trace.lev      mts      compute.rd
##           0          1000          0
##           psi           subsampling           cov
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 36"
## [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
##   16   16   15   24   26   20   17   14   24   13   18   19   26   31   22
## 2011 2012
##   26   21
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    6    5   10    8    3    6    8   14    9   11   11   18   24   10
## 2011 2012
##   17   14
##

```

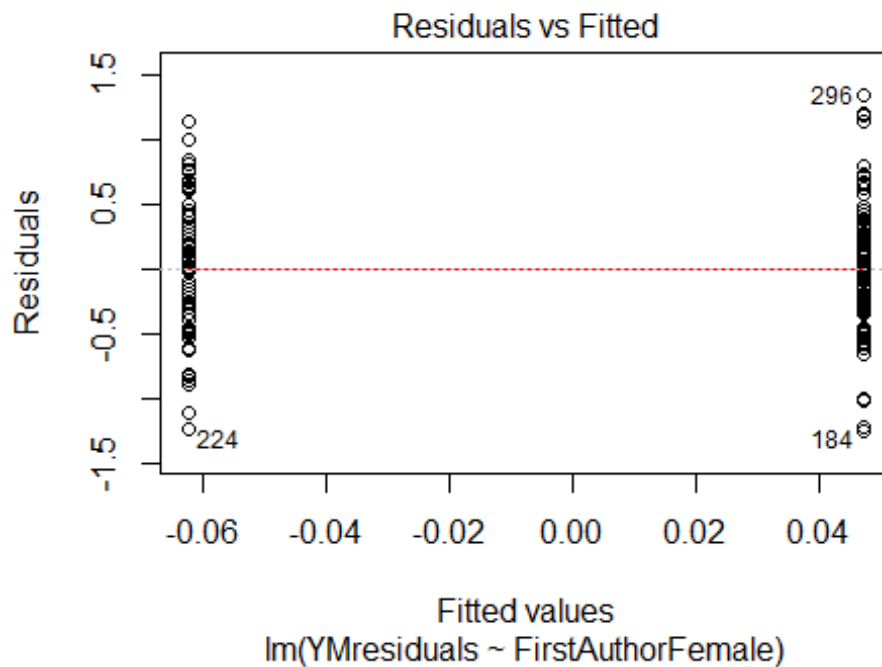
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    6    5   10    8    3    5    3    9    5    8    8   17   20   10
## 2011 2012
##   15   13
## [1] "Heteroscedasticity checks, confirming that there are problems with
##      these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 20, df = 16, p-value = 0.2
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.03, df = 1, p-value = 0.9
##
## [1] "Female first author team size 2018 geometric mean: 3.43754385517496"
## [1] "Male first author team size 2018 geometric mean: 4.91901897169873"
##
## Warning in wilcox.test.default(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] "Female last author team size 2018 geometric mean: 2.91295063024394"
## [1] "Male last author team size 2018 geometric mean: 5.17468047354489"

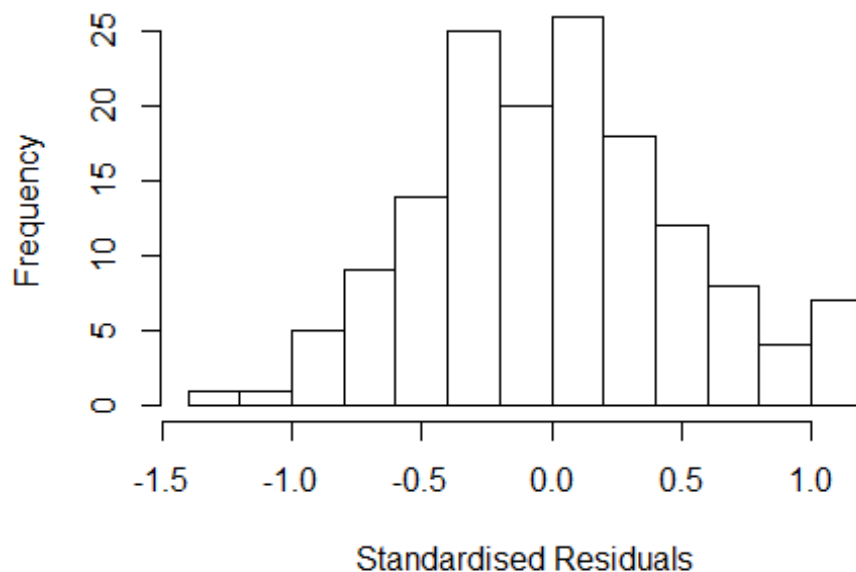
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5, p-value = 0.2
## 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 | 1.876 | 1 | 1.370 |
| LastAuthorFemale | 2.328 | 1 | 1.526 |
| UniqueAuthors | 7.984 | 4 | 1.297 |
| Year | 11.185 | 16 | 1.078 |

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
## -1.21107 -0.30921 0.00126 0.33375 1.16745
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.392 0.403 3.46 0.00074 ***
## FirstAuthorFemale1 -0.146 0.102 -1.43 0.15490
## LastAuthorFemale1 -0.150 0.142 -1.05 0.29556
## UniqueAuthors2 0.435 0.215 2.03 0.04479 *
## UniqueAuthors3 0.280 0.227 1.23 0.22054
## UniqueAuthors4 0.370 0.209 1.77 0.07864 .
## UniqueAuthors5 0.518 0.200 2.60 0.01055 *
## Year1997 -0.368 0.377 -0.98 0.33127
## Year1998 -0.309 0.313 -0.99 0.32555
## Year1999 -0.364 0.346 -1.05 0.29392
```

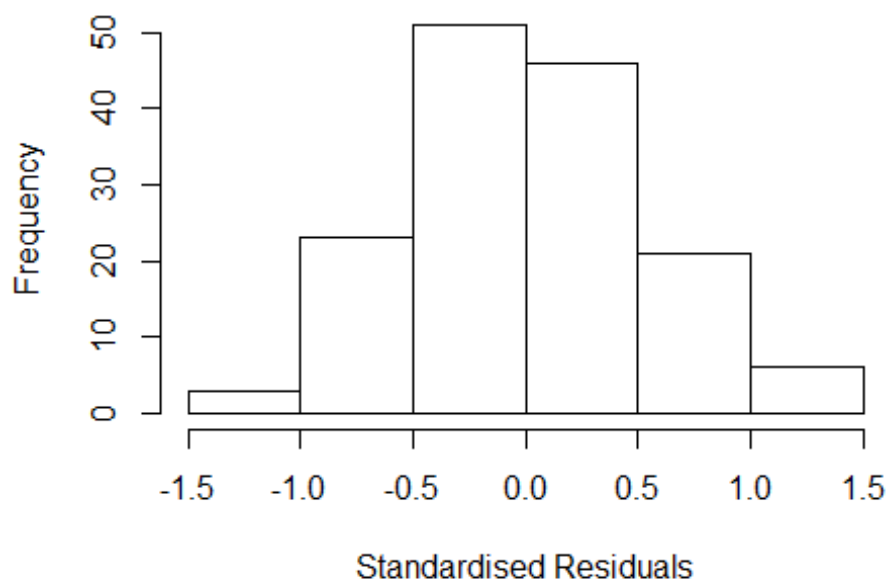


```

## Year2000          -0.245      0.343   -0.72   0.47487
## Year2001          -0.189      0.501   -0.38   0.70674
## Year2002          -0.394      0.341   -1.16   0.24920
## Year2003          -0.736      0.411   -1.79   0.07525 .
## Year2004          -0.468      0.346   -1.35   0.17863
## Year2005          -0.499      0.389   -1.28   0.20220
## Year2006          -0.500      0.366   -1.36   0.17467
## Year2007          -0.229      0.369   -0.62   0.53677
## Year2008          -0.429      0.376   -1.14   0.25672
## Year2009          -0.699      0.333   -2.10   0.03806 *
## Year2010          -0.350      0.351   -1.00   0.32098
## Year2011          -0.446      0.340   -1.31   0.19235
## Year2012          -0.546      0.414   -1.32   0.18974
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.515
## Multiple R-squared:  0.221, Adjusted R-squared:  0.0863
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 140 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.560  0.880  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      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 1.492 1      1.221
## LastAuthorFemale  1.462 1      1.209
## Year              1.653 16      1.016

```

Residuals from first and last author



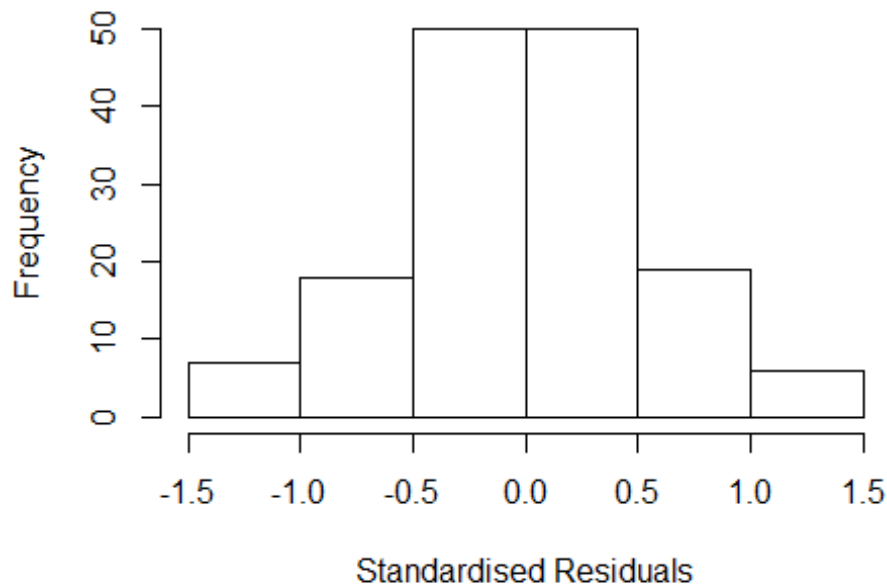
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2240 -0.3260 -0.0328 0.3333 1.3556
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.756 0.282 6.22 6.2e-09 ***
## FirstAuthorFemale1 -0.144 0.101 -1.42 0.157
## LastAuthorFemale1 -0.235 0.148 -1.59 0.115
## Year1997 -0.247 0.358 -0.69 0.491
## Year1998 -0.309 0.287 -1.08 0.282
## Year1999 -0.390 0.317 -1.23 0.222
## Year2000 -0.213 0.294 -0.72 0.470
## Year2001 -0.404 0.307 -1.32 0.191
## Year2002 -0.320 0.301 -1.06 0.290
## Year2003 -0.882 0.345 -2.56 0.012 *
## Year2004 -0.369 0.290 -1.27 0.205
## Year2005 -0.393 0.383 -1.03 0.307
```

```

## Year2006          -0.411      0.329   -1.25    0.213
## Year2007          -0.218      0.339   -0.64    0.521
## Year2008          -0.392      0.312   -1.26    0.211
## Year2009          -0.734      0.304   -2.41    0.017 *
## Year2010          -0.338      0.313   -1.08    0.283
## Year2011          -0.384      0.298   -1.29    0.199
## Year2012          -0.532      0.410   -1.30    0.197
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.527
## Multiple R-squared:  0.148, Adjusted R-squared:  0.0306
## 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.
## 0.488 0.871 0.957 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      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 1.465 1      1.211
## Year      1.465 16      1.012

```

Residuals from first author



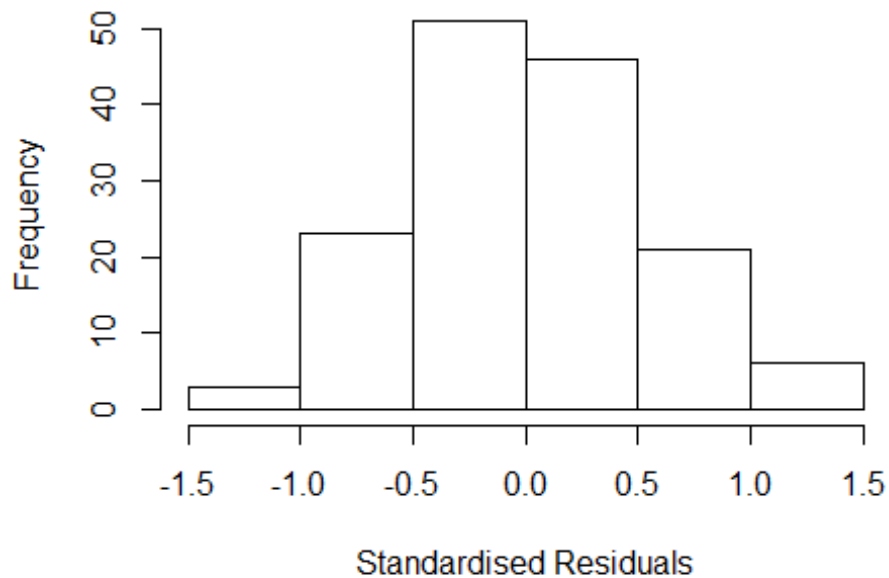
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.20766 -0.30380 0.00939 0.30114 1.32865
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.7429 0.2569 6.79 3.5e-10 ***
## FirstAuthorFemale1 -0.1859 0.0969 -1.92 0.0573 .
## Year1997 -0.2861 0.2987 -0.96 0.3400
## Year1998 -0.3176 0.2551 -1.24 0.2153
## Year1999 -0.4213 0.2916 -1.44 0.1509
## Year2000 -0.2143 0.2654 -0.81 0.4209
## Year2001 -0.3780 0.2775 -1.36 0.1755
## Year2002 -0.3284 0.2943 -1.12 0.2665
## Year2003 -0.9542 0.3096 -3.08 0.0025 **
## Year2004 -0.3595 0.2695 -1.33 0.1845
## Year2005 -0.3719 0.3833 -0.97 0.3338
## Year2006 -0.4741 0.3233 -1.47 0.1450
```

```

## Year2007          -0.2220      0.3061   -0.73   0.4696
## Year2008          -0.4225      0.2975   -1.42   0.1579
## Year2009          -0.7664      0.2850   -2.69   0.0081 **
## Year2010          -0.3559      0.2950   -1.21   0.2298
## Year2011          -0.3816      0.2691   -1.42   0.1585
## Year2012          -0.5352      0.4022   -1.33   0.1856
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.555
## Multiple R-squared:  0.118, Adjusted R-squared:  0.00493
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 133 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.545  0.876  0.960  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      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 1.286 1          1.134
## Year            1.286 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
## -1.17773 -0.31446 -0.00671 0.33989 1.43296
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.621 0.271 5.98 2e-08 ***
## LastAuthorFemale1 -0.273 0.140 -1.95 0.054 .
## Year1997 -0.147 0.383 -0.38 0.702
## Year1998 -0.255 0.284 -0.90 0.370
## Year1999 -0.318 0.321 -0.99 0.324
## Year2000 -0.129 0.291 -0.44 0.658
## Year2001 -0.315 0.324 -0.97 0.333
## Year2002 -0.262 0.290 -0.90 0.369
## Year2003 -0.734 0.333 -2.20 0.029 *
## Year2004 -0.279 0.288 -0.97 0.333
## Year2005 -0.328 0.384 -0.85 0.394
## Year2006 -0.383 0.324 -1.18 0.239
```

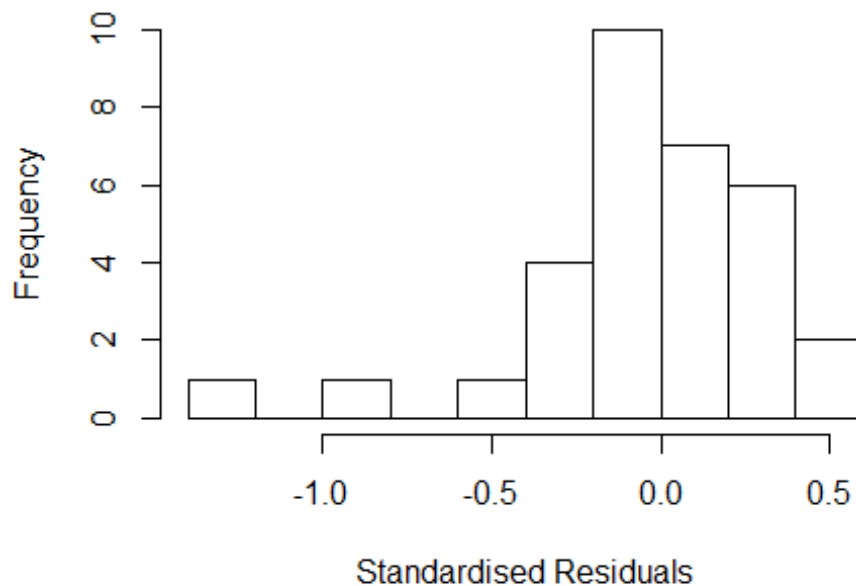
```

## Year2007          -0.110          0.333   -0.33    0.742
## Year2008          -0.311          0.311   -1.00    0.318
## Year2009          -0.639          0.309   -2.07    0.040 *
## Year2010          -0.241          0.313   -0.77    0.443
## Year2011          -0.309          0.298   -1.04    0.302
## Year2012          -0.443          0.375   -1.18    0.240
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.542
## Multiple R-squared:  0.136, Adjusted R-squared:  0.0252
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 136 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.464  0.874  0.953  0.912  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      6.67e-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 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
##    9    3    6    3    6    4    5    4    4    1    3    6    3    5    7
## 2011 2012
##    8    8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    2    3    1    3    4    2    3    1    1    3    1    2    3
## 2011 2012

```

```
##      3      4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      0      2      3      1      2      3      2      3      1      0      2      0      1      3
## 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: 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 2.214e+14 1      14881088
## LastAuthorFemale -4.812e+15 1           NaN
## Year              -1.057e+30 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))
```

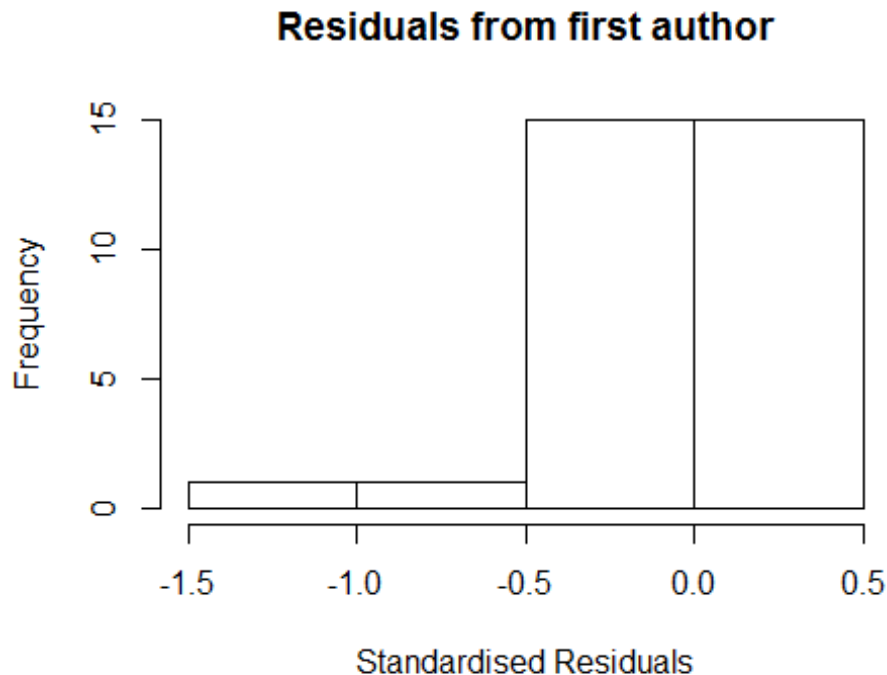


```

## \--> method = "MM"
## Residuals:
##      Min      1Q   Median      3Q      Max
## -1.21e+00 -1.51e-01  5.27e-16  1.94e-01  4.09e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.683      0.375   1.82   0.087 .
## FirstAuthorFemale1 -0.143      0.163  -0.88   0.391
## LastAuthorFemale1  0.035      0.259   0.14   0.894
## Year1998          -0.207      0.393  -0.53   0.605
## Year1999           0.422      0.361   1.17   0.260
## Year2000           0.350      0.470   0.74   0.467
## Year2001           0.341      0.406   0.84   0.413
## Year2002           0.311      0.442   0.70   0.491
## Year2003           0.783      0.361   2.17   0.045 *
## Year2004           0.762      0.592   1.29   0.216
## Year2005           0.599      0.352   1.70   0.109
## Year2007           0.305      0.419   0.73   0.478
## Year2009           0.725      0.523   1.39   0.184
## Year2010           0.313      0.395   0.79   0.439
## Year2011           0.344      0.411   0.84   0.415
## Year2012           0.672      0.641   1.05   0.310
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.415
## Multiple R-squared:  0.431, Adjusted R-squared:  -0.103
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
##  5 weights are ~= 1. The remaining 27 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.374  0.934  0.977   0.935  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      3.13e-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 -1.425e+15  1      NaN
## Year              -1.425e+15 13      NaN
```



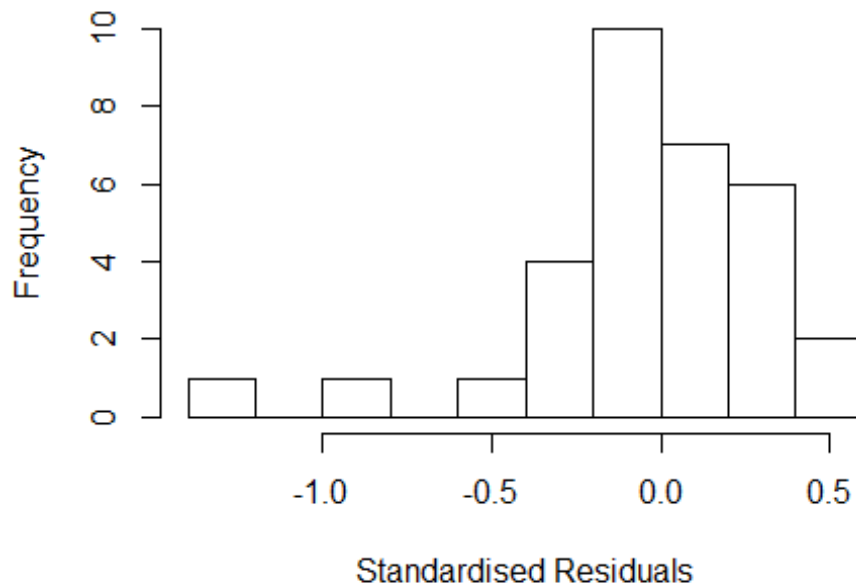
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.29e+00 -1.58e-01  2.78e-16  1.55e-01  4.02e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.676      0.357   1.89   0.075 .
## FirstAuthorFemale1 -0.129      0.152  -0.85   0.406
## Year1998         -0.207      0.395  -0.52   0.607
## Year1999          0.424      0.358   1.18   0.252
## Year2000          0.378      0.377   1.00   0.330
## Year2001          0.348      0.390   0.89   0.384
## Year2002          0.326      0.404   0.81   0.431
```

```

## Year2003          0.776      0.385      2.02      0.060 .
## Year2004          0.790      0.543      1.45      0.164
## Year2005          0.592      0.377      1.57      0.135
## Year2007          0.322      0.374      0.86      0.401
## Year2009          0.767      0.357      2.15      0.046 *
## Year2010          0.320      0.380      0.84      0.411
## Year2011          0.351      0.412      0.85      0.406
## Year2012          0.743      0.336      2.21      0.041 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.385
## Multiple R-squared:  0.47,   Adjusted R-squared:  0.0336
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 28 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.238  0.932  0.975   0.928   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      3.13e-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 -4.54e+14 1      NaN
## Year -4.54e+14 13      NaN

```

Residuals from last author



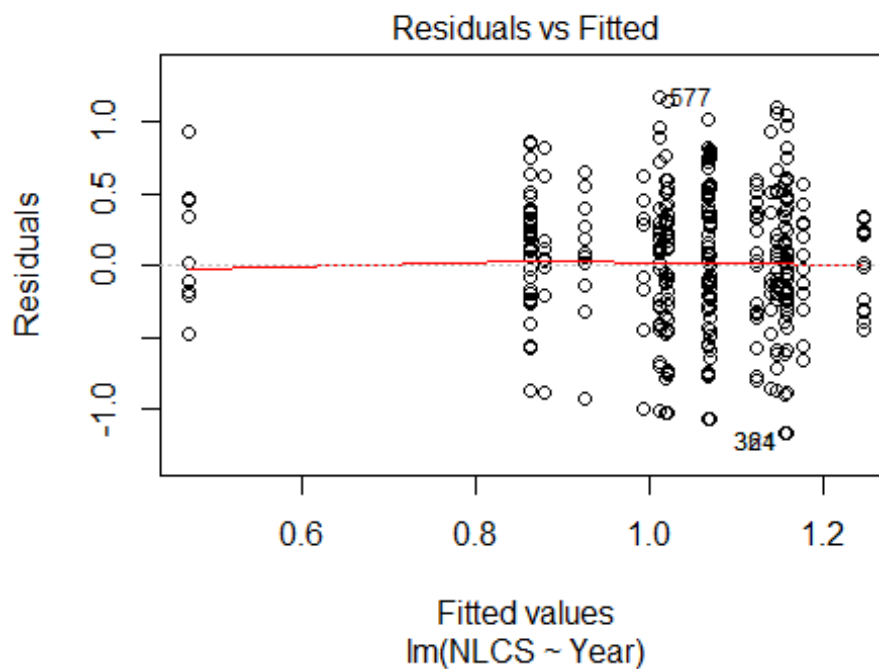
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.39e+00 -1.19e-01 1.67e-16 1.50e-01 4.53e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6110 0.2801 2.18 0.04347 *
## LastAuthorFemale1 -0.0394 0.1096 -0.36 0.72375
## Year1998 -0.2070 0.3015 -0.69 0.50155
## Year1999 0.4455 0.2838 1.57 0.13492
## Year2000 0.3524 0.2130 1.65 0.11634
## Year2001 0.4130 0.3209 1.29 0.21528
## Year2002 0.3677 0.3535 1.04 0.31282
## Year2003 0.7115 0.2905 2.45 0.02545 *
## Year2004 0.8595 0.4930 1.74 0.09934 .
## Year2005 0.5270 0.2801 1.88 0.07711 .
## Year2007 0.3422 0.2533 1.35 0.19439
## Year2009 0.8714 0.2130 4.09 0.00076 ***
```

```

## Year2010          0.3843      0.3125      1.23  0.23564
## Year2011          0.3351      0.3204      1.05  0.31031
## Year2012          0.7812      0.0793      9.85  1.9e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.379
## Multiple R-squared:  0.488, Adjusted R-squared:  0.066
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 26 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.149  0.929  0.978  0.917  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      3.13e-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: 32"
## [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
##   22   31   21   26   30   23   34   25   27   31   46   48   59   46   34
## 2011 2012
##   47   30
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   14   14   11   11   12    8   19   13   16   21   35   28   44   36   28
## 2011 2012
##   41   21
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010

```

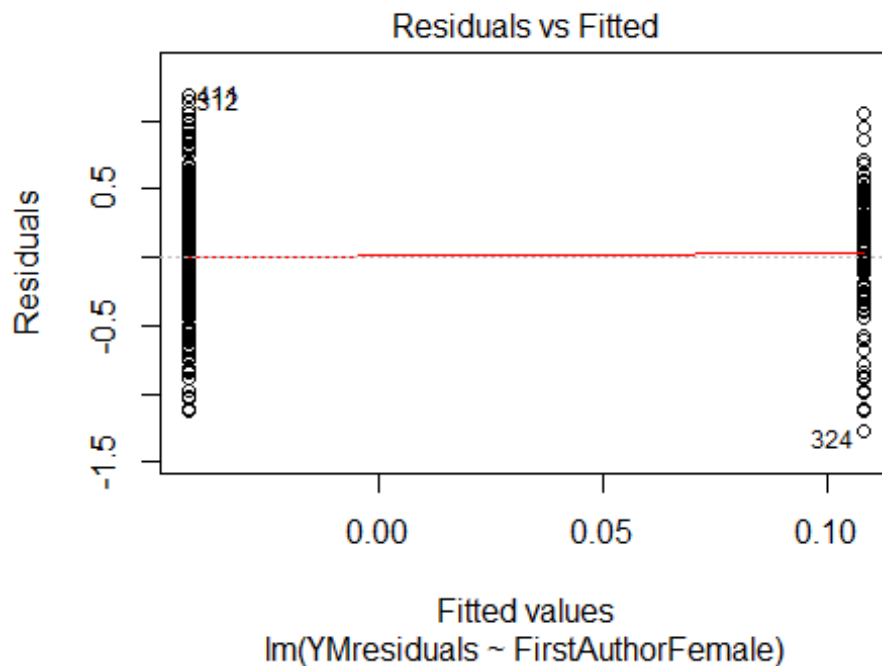
```
##      13      14      9      11      11      8      18      13      13      19      32      24      37      32      23
## 2011 2012
##      38      18
## [1] "Heteroscedasticity checks, confirming that there are problems with
##      these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 9, df = 16, p-value = 0.9
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.36, df = 1, p-value = 0.5
## [1] "Female first author team size 2018 geometric mean: 3.94248731045537"
## [1] "Male first author team size 2018 geometric mean: 5.09870965818875"
## Warning in wilcox.test.default(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 = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.41633344489691"
## [1] "Male last author team size 2018 geometric mean: 4.71978769132795"

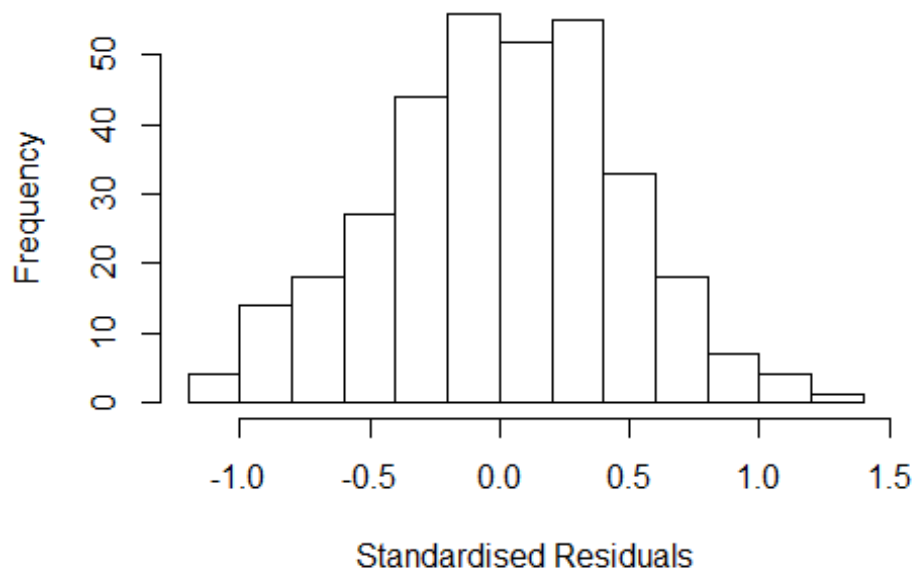
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## 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 | 1.294 | 1 | 1.138 |
| LastAuthorFemale | 1.373 | 1 | 1.172 |
| UniqueAuthors | 3.184 | 4 | 1.156 |
| Year | 5.062 | 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
## -1.1531 -0.3145 0.0153 0.2927 1.2305
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.42587 0.11713 3.64 0.00032 ***
## FirstAuthorFemale1 0.18065 0.06339 2.85 0.00467 **
## LastAuthorFemale1 -0.03286 0.08728 -0.38 0.70678
## UniqueAuthors2 0.00311 0.14260 0.02 0.98260
## UniqueAuthors3 -0.05285 0.12046 -0.44 0.66112
## UniqueAuthors4 0.08277 0.12237 0.68 0.49932
## UniqueAuthors5 0.25732 0.10960 2.35 0.01952 *
## Year1997 0.43031 0.16656 2.58 0.01024 *
## Year1998 0.52303 0.20433 2.56 0.01095 *
## Year1999 0.62092 0.18234 3.41 0.00075 ***
```

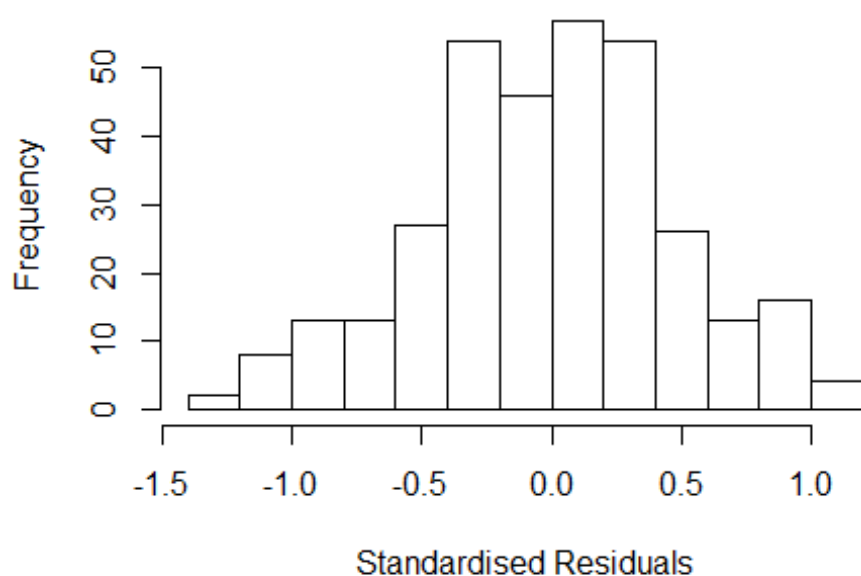


```

## Year2000      0.64114      0.15799      4.06 6.3e-05 ***
## Year2001      0.38959      0.26197      1.49 0.13798
## Year2002      0.58438      0.16031      3.65 0.00031 ***
## Year2003      0.66893      0.15510      4.31 2.2e-05 ***
## Year2004      0.49757      0.18804      2.65 0.00856 **
## Year2005      0.50891      0.18692      2.72 0.00684 **
## Year2006      0.54348      0.14872      3.65 0.00030 ***
## Year2007      0.49689      0.16207      3.07 0.00236 **
## Year2008      0.29417      0.15185      1.94 0.05362 .
## Year2009      0.44427      0.16706      2.66 0.00824 **
## Year2010      0.38761      0.15958      2.43 0.01571 *
## Year2011      0.44282      0.15214      2.91 0.00387 **
## Year2012      0.49432      0.16778      2.95 0.00346 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.457
## Multiple R-squared:  0.168, Adjusted R-squared:  0.109
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 303 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.448  0.872  0.951  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      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 1.332 1      1.154
## LastAuthorFemale 1.433 1      1.197
## Year      1.886 16      1.020

```

Residuals from first and last author



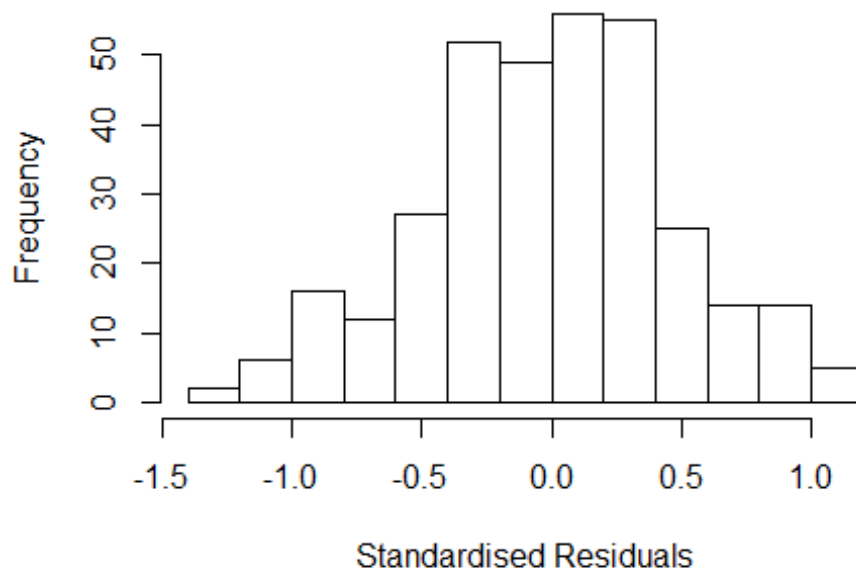
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2813 -0.3148 0.0101 0.3033 1.1294
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.4593 0.1238 3.71 0.00025 ***
## FirstAuthorFemale1 0.1796 0.0631 2.85 0.00468 **
## LastAuthorFemale1 -0.0639 0.0929 -0.69 0.49180
## Year1997 0.5045 0.1713 2.95 0.00347 **
## Year1998 0.5439 0.1736 3.13 0.00190 **
## Year1999 0.6191 0.1858 3.33 0.00097 ***
## Year2000 0.7657 0.1578 4.85 1.9e-06 ***
## Year2001 0.4956 0.2483 2.00 0.04675 *
## Year2002 0.6633 0.1561 4.25 2.8e-05 ***
## Year2003 0.7329 0.1481 4.95 1.2e-06 ***
## Year2004 0.6379 0.2020 3.16 0.00174 **
## Year2005 0.6592 0.1718 3.84 0.00015 ***
```

```

## Year2006          0.6424      0.1530      4.20  3.5e-05 ***
## Year2007          0.5968      0.1579      3.78  0.00019 ***
## Year2008          0.3769      0.1555      2.42  0.01593 *
## Year2009          0.5525      0.1612      3.43  0.00069 ***
## Year2010          0.5424      0.1578      3.44  0.00067 ***
## Year2011          0.5443      0.1554      3.50  0.00053 ***
## Year2012          0.6437      0.1592      4.04  6.6e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.464
## Multiple R-squared:  0.113, Adjusted R-squared:  0.0618
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 310 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.427  0.875   0.951   0.904   0.985   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      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 1.348 1      1.161
## Year              1.348 16      1.009

```

Residuals from first author



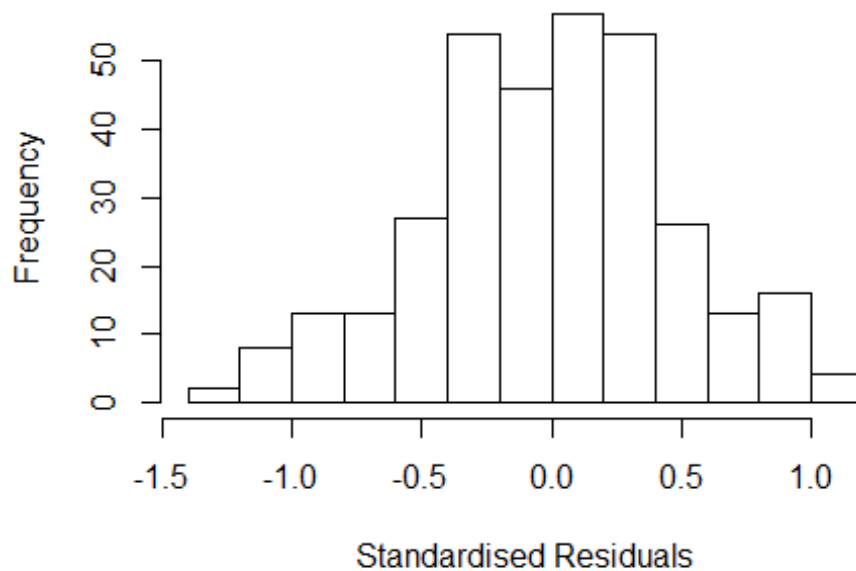
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.275 -0.305 0.010 0.320 1.128
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.4543 0.1220 3.72 0.00023 ***
## FirstAuthorFemale1 0.1740 0.0637 2.73 0.00662 **
## Year1997 0.5098 0.1700 3.00 0.00292 **
## Year1998 0.5225 0.1694 3.08 0.00222 **
## Year1999 0.6181 0.1848 3.34 0.00093 ***
## Year2000 0.7672 0.1586 4.84 2.1e-06 ***
## Year2001 0.5008 0.2406 2.08 0.03817 *
## Year2002 0.6654 0.1566 4.25 2.8e-05 ***
## Year2003 0.7254 0.1442 5.03 8.3e-07 ***
## Year2004 0.6468 0.2001 3.23 0.00136 **
## Year2005 0.6660 0.1708 3.90 0.00012 ***
## Year2006 0.6457 0.1516 4.26 2.7e-05 ***
```

```

## Year2007          0.5971      0.1565      3.82  0.00016 ***
## Year2008          0.3725      0.1537      2.42  0.01594 *
## Year2009          0.5448      0.1596      3.41  0.00073 ***
## Year2010          0.5308      0.1550      3.42  0.00070 ***
## Year2011          0.5331      0.1533      3.48  0.00058 ***
## Year2012          0.6373      0.1594      4.00  8.0e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.463
## Multiple R-squared:  0.112, Adjusted R-squared:  0.0644
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 309 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.428  0.869  0.950  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      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 1.449 1      1.204
## Year              1.449 16      1.012

```

Residuals from last author



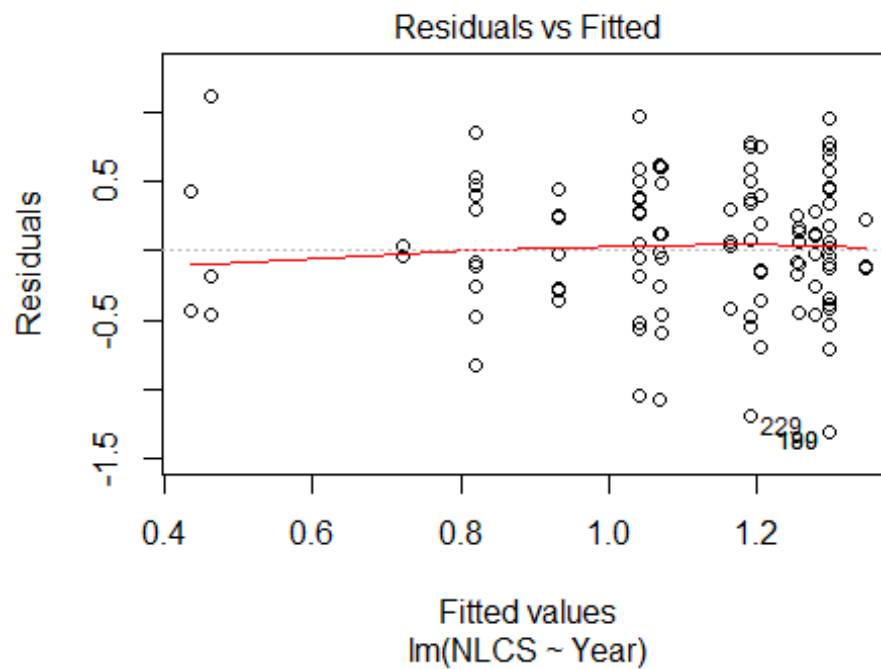
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1741 -0.2967 0.0239 0.3244 1.1668
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.4907 0.1255 3.91 0.00011 ***
## LastAuthorFemale1 -0.0387 0.0950 -0.41 0.68424
## Year1997 0.4713 0.1723 2.74 0.00659 **
## Year1998 0.5895 0.1663 3.54 0.00045 ***
## Year1999 0.6412 0.1971 3.25 0.00127 **
## Year2000 0.7509 0.1597 4.70 3.9e-06 ***
## Year2001 0.5473 0.2270 2.41 0.01650 *
## Year2002 0.6834 0.1619 4.22 3.2e-05 ***
## Year2003 0.7668 0.1532 5.01 9.3e-07 ***
## Year2004 0.6159 0.1987 3.10 0.00211 **
## Year2005 0.6857 0.1815 3.78 0.00019 ***
## Year2006 0.6425 0.1579 4.07 5.9e-05 ***
```

```

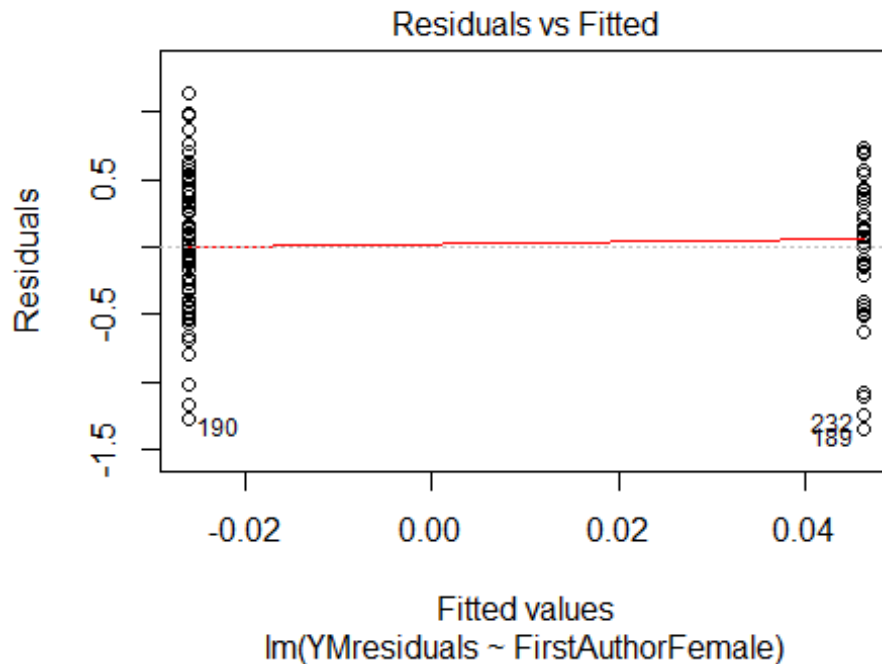
## Year2007          0.6344      0.1607      3.95  9.8e-05 ***
## Year2008          0.3982      0.1571      2.54  0.01173 *
## Year2009          0.5805      0.1636      3.55  0.00045 ***
## Year2010          0.5225      0.1633      3.20  0.00151 **
## Year2011          0.5740      0.1603      3.58  0.00040 ***
## Year2012          0.6638      0.1593      4.17  4.0e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.475
## Multiple R-squared:  0.0875, Adjusted R-squared:  0.0383
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 311 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.521  0.866  0.952  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      3.00e-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 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
##   12   7   6   8   12   6   8   6   12   17   12   14   25   13   19
## 2011 2012
##   14   13
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   2    4    2    3    7    3    6    4    7    9    8    8   14    6   14
## 2011 2012

```

```
## 11 11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1 3 2 3 4 2 6 4 4 6 6 8 10 4 8
## 2011 2012
## 6 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 = 25, df = 16, p-value = 0.07
```



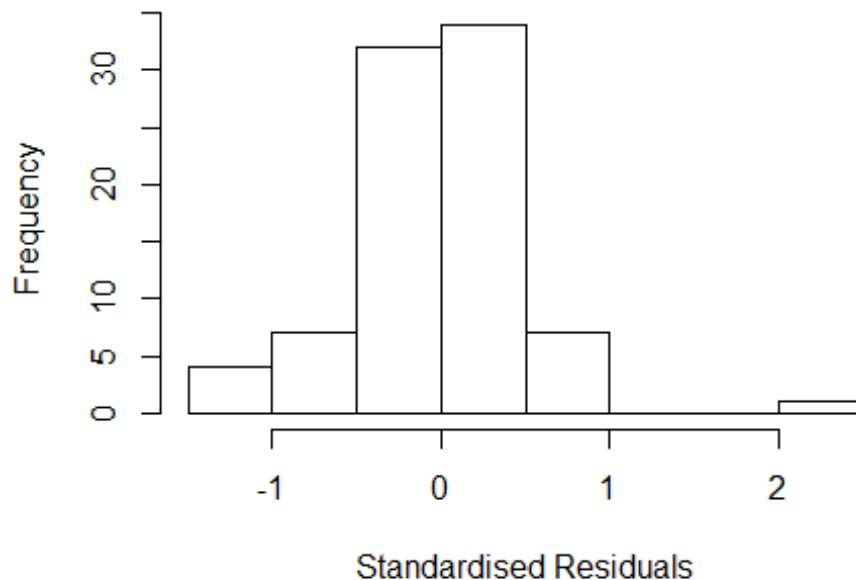
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.26, df = 1, p-value = 0.6
```

```
## [1] "Female first author team size 2018 geometric mean: 4"
## [1] "Male first author team size 2018 geometric mean: 3.87298334620742"
##
## 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: 3"
## [1] "Male last author team size 2018 geometric mean: 4.47213595499958"
##
## 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] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
```

| | GVIF | Df | GVIF ^{1/(2*Df)} |
|-------------------|--------|----|--------------------------|
| FirstAuthorFemale | 10.3 | 1 | 3.209 |
| LastAuthorFemale | 30.0 | 1 | 5.477 |
| UniqueAuthors | 800.2 | 4 | 2.306 |
| Year | 3117.0 | 16 | 1.286 |

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
## -1.37e+00 -2.13e-01 -4.55e-15 2.50e-01 2.17e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.50770 0.45524 1.12 0.26905
## FirstAuthorFemale1 -0.00694 0.16315 -0.04 0.96623
## LastAuthorFemale1 0.07264 0.41973 0.17 0.86317
## UniqueAuthors2 -0.34873 0.50358 -0.69 0.49120
## UniqueAuthors3 0.15307 0.42793 0.36 0.72179
## UniqueAuthors4 0.34591 0.43168 0.80 0.42602
## UniqueAuthors5 0.35930 0.45524 0.79 0.43297
## Year1997 -0.75719 0.13789 -5.49 7.9e-07 ***
## Year1998 0.31610 0.29567 1.07 0.28916
## Year1999 0.46375 0.30073 1.54 0.12815
```

```

## Year2000          0.39114      0.22372      1.75  0.08535 .
## Year2001          0.54297      0.14726      3.69  0.00048 ***
## Year2002          0.51485      0.18095      2.85  0.00601 **
## Year2003          0.43123      0.17837      2.42  0.01857 *
## Year2004          0.03018      0.10680      0.28  0.77843
## Year2005          0.73440      0.27924      2.63  0.01075 *
## Year2006          0.50869      0.18408      2.76  0.00752 **
## Year2007          0.62784      0.23854      2.63  0.01070 *
## Year2008          0.26834      0.23605      1.14  0.26000
## Year2009          0.66395      0.20906      3.18  0.00233 **
## Year2010          0.47465      0.23870      1.99  0.05117 .
## Year2011          0.25000      0.14031      1.78  0.07969 .
## Year2012          0.50567      0.62938      0.80  0.42479
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.374
## Multiple R-squared:  0.392, Adjusted R-squared:  0.176
## Convergence in 43 IRWLS iterations
##
## Robustness weights:
## observation 3 is an outlier with |weight| = 0 ( < 0.0012);
## 4 weights are ~= 1. The remaining 80 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.155  0.883   0.962   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          1.18e-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 -6.307e+13  1          NaN
## LastAuthorFemale -1.553e+13  1          NaN
## Year -2.895e+16 16          NaN
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields      residuals

```

```

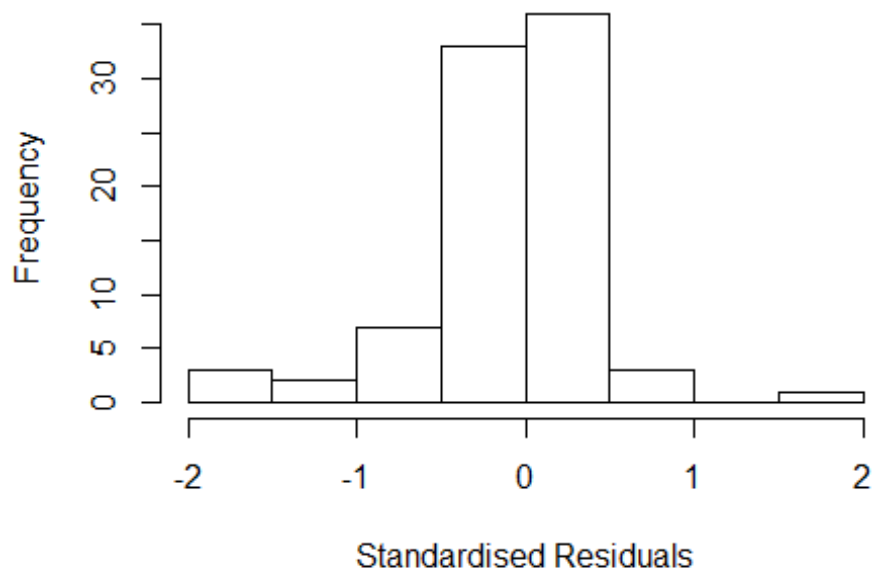
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.65910 -0.20735 -0.00638  0.19269  1.56562
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    8.67e-01   1.46e-08  5.93e+07 < 2e-16 ***
## FirstAuthorFemale1 7.37e-02   1.19e-01  6.20e-01  0.53673
## LastAuthorFemale1 -2.55e-01   1.49e-01 -1.71e+00  0.09260 .
## Year1997        -8.61e-01   5.76e-02 -1.49e+01 < 2e-16 ***
## Year1998       -1.81e-01   5.93e-02 -3.06e+00  0.00324 **
## Year1999        3.12e-01   1.65e-01  1.89e+00  0.06274 .
## Year2000        3.06e-01   1.32e-01  2.32e+00  0.02353 *
## Year2001        5.03e-01   1.68e-01  3.00e+00  0.00382 **
## Year2002        3.68e-01   9.26e-02  3.97e+00  0.00018 ***
## Year2003        3.05e-01   1.35e-01  2.26e+00  0.02710 *
## Year2004       -2.41e-02   1.14e-01 -2.10e-01  0.83323
## Year2005        7.05e-01   2.83e-01  2.49e+00  0.01522 *
## Year2006        4.02e-01   1.92e-01  2.09e+00  0.04008 *
## Year2007        4.33e-01   9.46e-02  4.58e+00  2.1e-05 ***
## Year2008        2.84e-01   2.30e-01  1.23e+00  0.22159
## Year2009        5.28e-01   2.12e-01  2.49e+00  0.01528 *
## Year2010        5.98e-01   2.39e-01  2.50e+00  0.01500 *
## Year2011        1.87e-01   1.42e-01  1.32e+00  0.19068
## Year2012        7.18e-01   1.65e-01  4.35e+00  4.8e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.35
## Multiple R-squared:  0.434, Adjusted R-squared:  0.279
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## observation 83 is an outlier with |weight| = 0 ( < 0.0012);
## 9 weights are ~ = 1. The remaining 75 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0045 0.8820 0.9580 0.8630 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      1.18e-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"

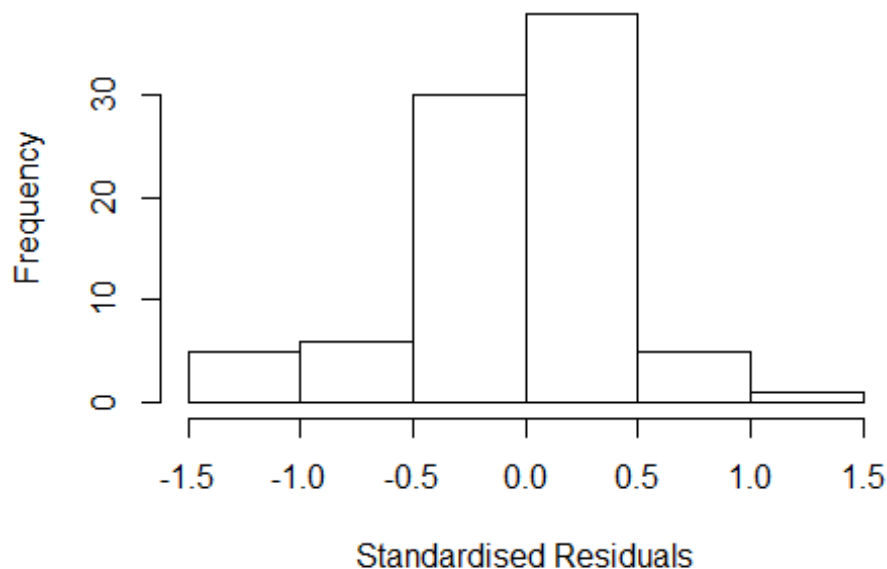
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead
```

Residuals from first and last author



```
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 9.872 1          3.142
## Year              9.872 16          1.074
```

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
## -1.4142 -0.2149 0.0385 0.2553 1.4502
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8670 0.0000 Inf < 2e-16 ***
## FirstAuthorFemale1 -0.0141 0.1406 -0.10 0.92063
## Year1997 -0.7452 0.3363 -2.22 0.03009 *
## Year1998 -0.1375 0.0773 -1.78 0.08004 .
## Year1999 0.4007 0.1795 2.23 0.02897 *
## Year2000 0.2664 0.1499 1.78 0.08003 .
## Year2001 0.5465 0.1384 3.95 0.00019 ***
## Year2002 0.4057 0.0980 4.14 1e-04 ***
## Year2003 0.3156 0.1677 1.88 0.06423 .
## Year2004 -0.0231 0.1124 -0.21 0.83767
## Year2005 0.6729 0.2267 2.97 0.00415 **
## Year2006 0.4236 0.1803 2.35 0.02177 *
```

```

## Year2007          0.4228      0.0868      4.87      7e-06 ***
## Year2008          0.2356      0.2238      1.05      0.29627
## Year2009          0.4919      0.2243      2.19      0.03174 *
## Year2010          0.4706      0.2400      1.96      0.05402 .
## Year2011          0.1808      0.1606      1.13      0.26424
## Year2012          0.5472      0.4326      1.26      0.21029
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.296, Adjusted R-squared:  0.118
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 80 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.167  0.891  0.960  0.887  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 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 16.56 1      4.069
## Year      16.56 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.61130 -0.20186 -0.00489  0.23700  1.56711

```

```

##
## Coefficients:
##              Estimate Std. Error   t value Pr(>|t|)
## (Intercept)    8.67e-01   3.85e-08  2.25e+07 < 2e-16 ***
## LastAuthorFemale1 -2.28e-01   1.36e-01 -1.68e+00  0.09802 .
## Year1997        -8.62e-01   4.73e-02 -1.82e+01 < 2e-16 ***
## Year1998        -1.44e-01   2.73e-02 -5.30e+00  1.4e-06 ***
## Year1999         3.86e-01   1.09e-01  3.55e+00  0.00072 ***
## Year2000         3.19e-01   1.27e-01  2.51e+00  0.01446 *
## Year2001         5.39e-01   1.26e-01  4.28e+00  6.1e-05 ***
## Year2002         4.02e-01   8.88e-02  4.53e+00  2.5e-05 ***
## Year2003         3.55e-01   9.02e-02  3.94e+00  0.00020 ***
## Year2004        -2.41e-02   1.14e-01 -2.10e-01  0.83299
## Year2005         7.17e-01   2.76e-01  2.60e+00  0.01152 *
## Year2006         4.33e-01   1.90e-01  2.28e+00  0.02591 *
## Year2007         4.50e-01   9.28e-02  4.85e+00  7.7e-06 ***
## Year2008         2.98e-01   2.45e-01  1.22e+00  0.22822
## Year2009         5.65e-01   2.14e-01  2.65e+00  0.01012 *
## Year2010         6.00e-01   2.79e-01  2.15e+00  0.03511 *
## Year2011         2.27e-01   1.42e-01  1.60e+00  0.11386
## Year2012         7.44e-01   1.76e-01  4.22e+00  7.4e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.349
## Multiple R-squared:  0.426, Adjusted R-squared:  0.28
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 2 observations c(81,83) are outliers with |weight| <= 0.00069 ( <
0.0012);
## 7 weights are ~= 1. The remaining 76 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0062 0.8880 0.9560 0.8770 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          1.18e-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 2707"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2002 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    1    2    1    1    1    2    1    1    1    4    5   13    8   10    6
##
## 1996 1997 1998 1999 2000 2002 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    0    2    0    1    0    1    0    1    1    4    5   13    6    8    5
##
## 1996 1997 1998 1999 2000 2002 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    0    2    0    1    0    0    0    1    1    4    5   13    6    8    5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 7"
## [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 = 0.5
## 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: 4.61043629205845"
## [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

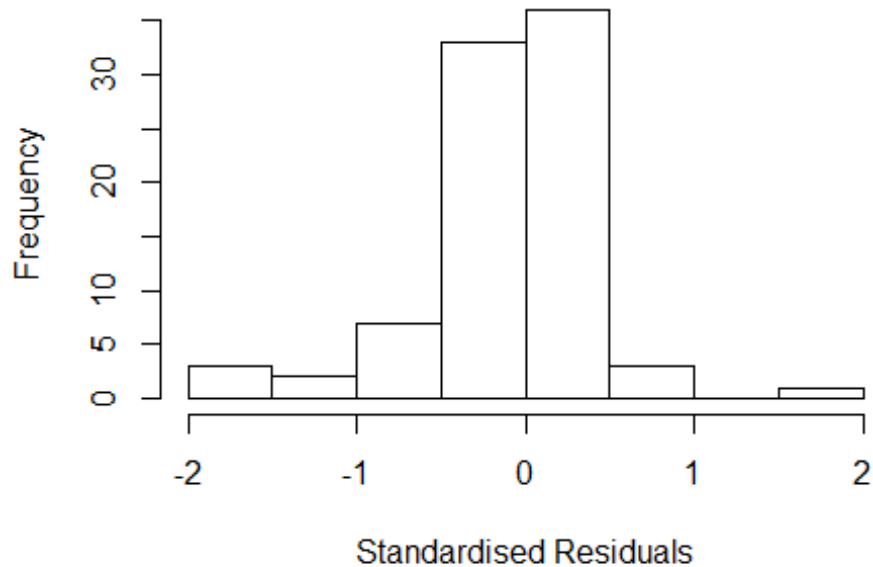
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"

## Warning in lmrob.S(x, y, control = control, mf = mf): 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 9          NaN

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "S"
## Residuals:
##   Min      1Q  Median      3Q      Max
## -1.118  0.000  0.000  0.000  0.632
##
## Exact fit detected
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.191      0.000      NA      NA
## FirstAuthorFemale1 -0.081      0.000      NA      NA
## LastAuthorFemale1  0.638      0.000      NA      NA
## Year1999          -0.845      0.000      NA      NA
## Year2005          -2.007      0.000      NA      NA
```

```

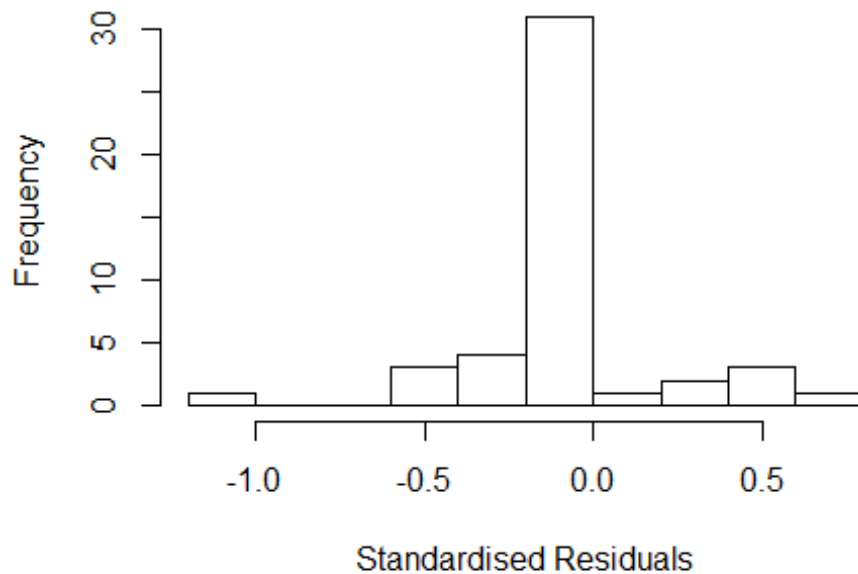
## Year2006          -2.191      0.000      NA      NA
## Year2007          -2.191      0.000      NA      NA
## Year2008          -2.191      0.000      NA      NA
## Year2009          -2.191      0.000      NA      NA
## Year2010          -1.602      0.000      NA      NA
## Year2011          -1.799      0.000      NA      NA
## Year2012          -2.191      0.000      NA      NA
##
## Robustness weights:
## 15 observations c(2,6,12,26,28,29,30,33,34,35,36,37,39,41,42)
## are outliers with |weight| = 0 ( < 0.0022);
## 31 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          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 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 9          NaN

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "S"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.23    0.00    0.00    0.00    1.55
##
## Exact fit detected
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.191      0.000      NA      NA
## FirstAuthorFemale1 0.949      0.000      NA      NA
## Year1999         -1.875      0.000      NA      NA
## Year2005         -1.369      0.000      NA      NA
## Year2006         -2.191      0.000      NA      NA
## Year2007         -1.631      0.000      NA      NA
```

```

## Year2008          -2.191      0.000      NA      NA
## Year2009          -2.191      0.000      NA      NA
## Year2010          -0.964      0.000      NA      NA
## Year2011          -2.191      0.000      NA      NA
## Year2012          -2.191      0.000      NA      NA
##
## Robustness weights:
## 16 observations c(2,7,8,9,12,26,28,29,30,31,33,34,38,39,41,42)
## are outliers with |weight| = 0 ( < 0.0022);
## 30 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          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"
## [1] "Sample size for the above analysis: 46"
## [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
##      9   12   15    6   10    8    5   10   10   14   12   16   18   19   14
## 2011 2012
##      14   15
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2    1    2    0    3    0    1    7    5    7    2    6   11    6    6
## 2011 2012
##      4    7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2    1    2    0    2    0    1    2    3    5    1    6    8    6    5
## 2011 2012
##      2    5

```

```

## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.58935930364667"
## [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 = 5.5, p-value = 0.9
## 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: 4.9567822111796"

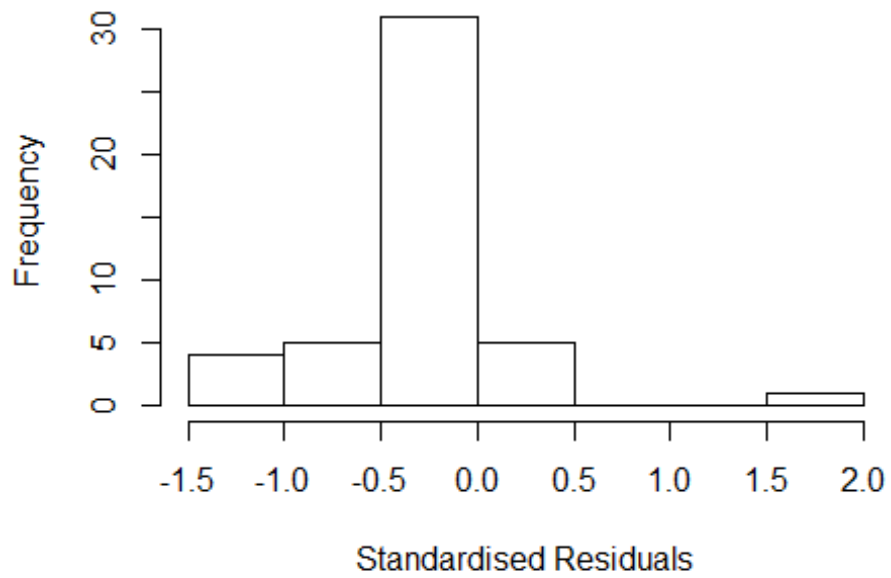
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## 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] "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

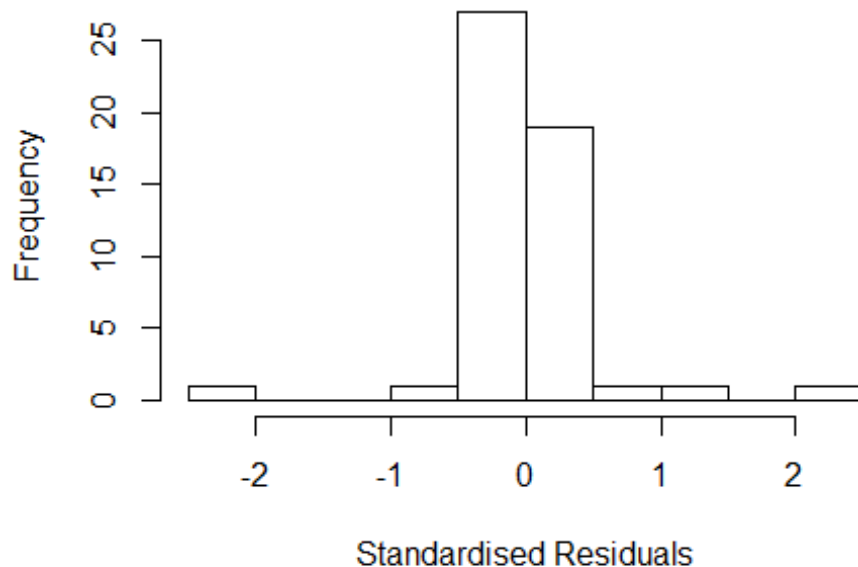
```

Residuals from first author



| | | | | |
|----|-------------------|------------|----|-----------------|
| ## | | GVIF | Df | GVIF^(1/(2*Df)) |
| ## | FirstAuthorFemale | -3.241e+14 | 1 | NaN |
| ## | LastAuthorFemale | 7.908e+13 | 1 | 8892651 |
| ## | UniqueAuthors | -1.663e+31 | 4 | NaN |
| ## | Year | -3.446e+45 | 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
## -2.136 -0.192 0.000 0.221 2.109
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3330 0.0000 Inf < 2e-16 ***
## FirstAuthorFemale1 -0.4604 0.1419 -3.24 0.00289 **
## LastAuthorFemale1 0.3552 0.1439 2.47 0.01946 *
## UniqueAuthors2 0.9120 0.1612 5.66 3.7e-06 ***
## UniqueAuthors3 0.9065 0.2080 4.36 0.00014 ***
## UniqueAuthors4 1.0352 0.2032 5.10 1.8e-05 ***
## UniqueAuthors5 1.0954 0.2039 5.37 8.1e-06 ***
## Year1997 0.0534 0.1809 0.30 0.77001
## Year1998 -0.5558 0.3695 -1.50 0.14293
## Year2000 -0.6145 0.2089 -2.94 0.00623 **
```

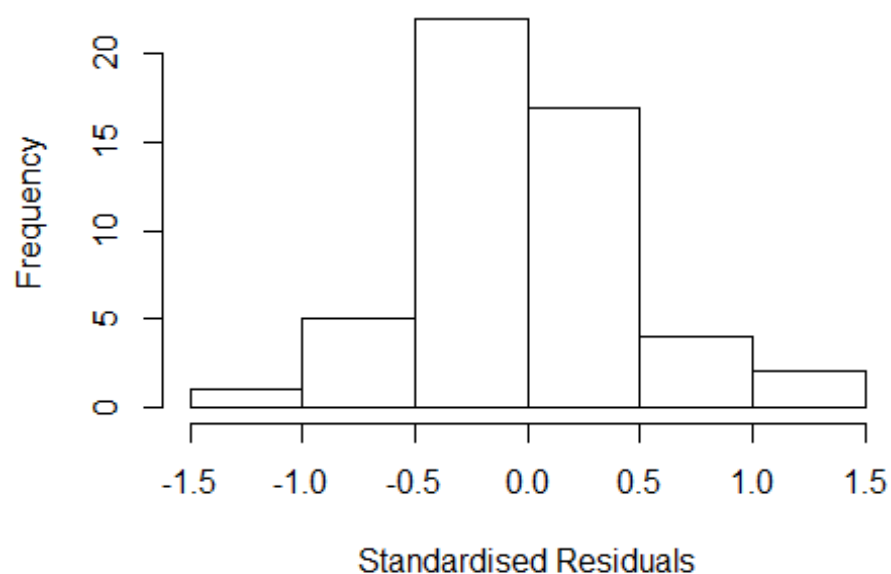


```

## Year2002          -0.2478      0.2353   -1.05   0.30068
## Year2003          -1.3656      0.4092   -3.34   0.00227 **
## Year2004          -1.2044      0.1359   -8.86   7.0e-10 ***
## Year2005          -0.1104      0.0965   -1.14   0.26154
## Year2006          -1.5788      0.1743   -9.06   4.4e-10 ***
## Year2007          -1.0739      0.2687   -4.00   0.00039 ***
## Year2008          -1.0752      0.2223   -4.84   3.7e-05 ***
## Year2009          -1.1076      0.1446   -7.66   1.5e-08 ***
## Year2010          -1.1302      0.2474   -4.57   7.8e-05 ***
## Year2011          -1.9305      0.1533  -12.59   1.7e-13 ***
## Year2012          -1.0311      0.2092   -4.93   2.8e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.725, Adjusted R-squared:  0.541
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 2 observations c(2,22) are outliers with |weight| = 0 ( < 0.002);
## 10 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.204  0.930  0.970  0.933  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          1.96e-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"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.049e+01 1          3.239
## LastAuthorFemale -8.043e+14 1          NaN
## Year              -4.306e+15 14          NaN

```

Residuals from first and last author



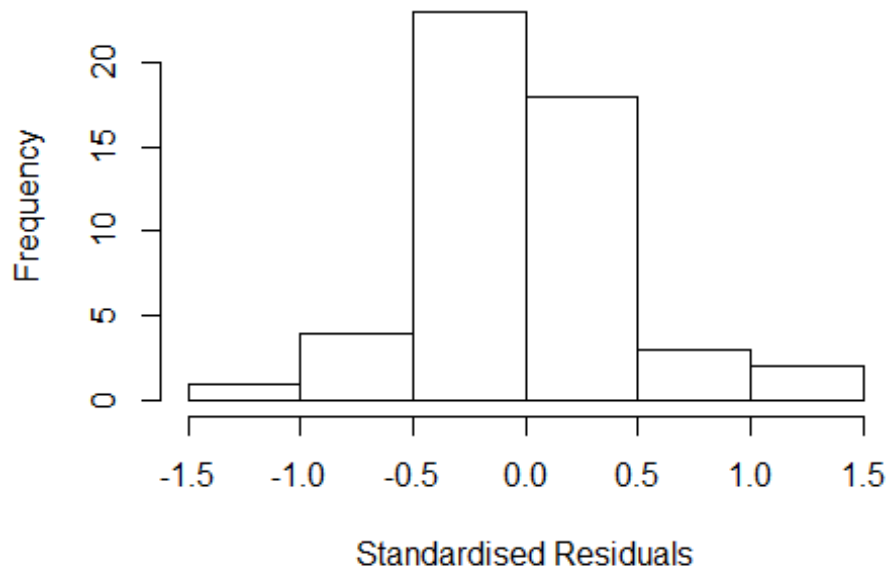
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.02e+00 -2.76e-01 -6.66e-16 2.48e-01 1.21e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8153 0.4656 1.75 0.089 .
## FirstAuthorFemale1 -0.2361 0.1982 -1.19 0.242
## LastAuthorFemale1 0.1664 0.1940 0.86 0.397
## Year1997 1.2588 0.4999 2.52 0.017 *
## Year1998 0.3824 1.3184 0.29 0.774
## Year2000 0.7003 0.4787 1.46 0.153
## Year2002 1.1464 0.5155 2.22 0.033 *
## Year2003 -0.1602 0.6003 -0.27 0.791
## Year2004 -0.0109 0.6036 -0.02 0.986
## Year2005 0.8670 0.5316 1.63 0.112
## Year2006 -0.2502 0.4999 -0.50 0.620
## Year2007 0.4364 0.5596 0.78 0.441
```

```

## Year2008          0.2042      0.5477      0.37      0.712
## Year2009          0.2053      0.4782      0.43      0.670
## Year2010          0.2541      0.5722      0.44      0.660
## Year2011         -0.4774      0.5037     -0.95      0.350
## Year2012          0.4569      0.4820      0.95      0.350
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.515
## Multiple R-squared:  0.398, Adjusted R-squared:  0.114
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 47 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.560  0.916  0.972  0.926  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      1.96e-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"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.456 1          1.859
## Year              3.456 14          1.045

```

Residuals from first author



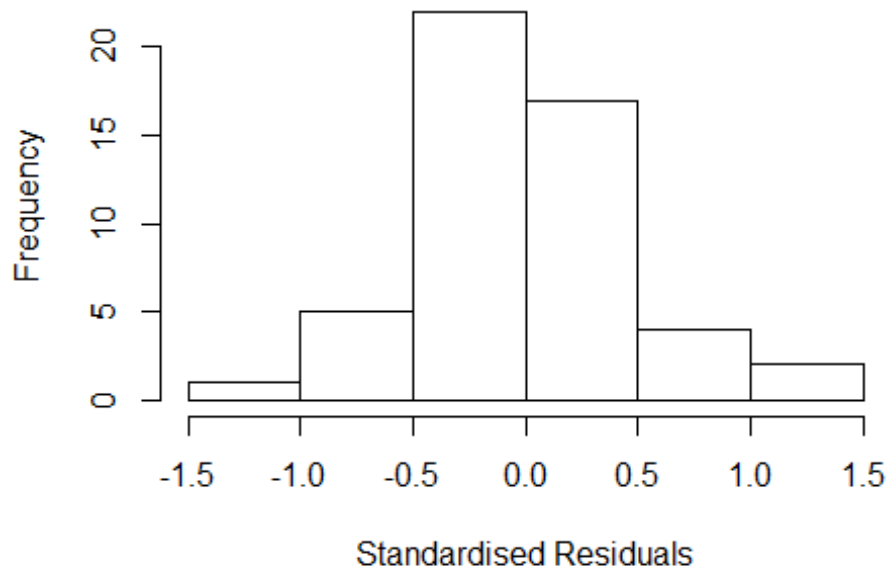
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.02e+00 -3.04e-01 -8.88e-16 2.53e-01 1.24e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8985 0.3558 2.53 0.0162 *
## FirstAuthorFemale1 -0.2080 0.2062 -1.01 0.3201
## Year1997 1.1475 0.4110 2.79 0.0084 **
## Year1998 0.4375 1.3077 0.33 0.7399
## Year2000 0.6030 0.3760 1.60 0.1177
## Year2002 1.2015 0.4110 2.92 0.0060 **
## Year2003 -0.2715 0.5327 -0.51 0.6135
## Year2004 -0.0484 0.5744 -0.08 0.9334
## Year2005 0.7660 0.4414 1.74 0.0915 .
## Year2006 -0.3615 0.4110 -0.88 0.3851
## Year2007 0.3322 0.4647 0.71 0.4794
## Year2008 0.1599 0.4719 0.34 0.7368
```

```

## Year2009          0.1600      0.3829      0.42      0.6786
## Year2010          0.1852      0.4914      0.38      0.7085
## Year2011         -0.5055      0.4323     -1.17      0.2502
## Year2012          0.3959      0.3841      1.03      0.3098
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.51
## Multiple R-squared:  0.391, Adjusted R-squared:  0.129
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 47 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.534  0.928  0.967  0.924  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      1.96e-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"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale -6.736e+14 1      NaN
## Year            -6.736e+14 14      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 -2.94e-01 2.22e-16 2.78e-01 1.21e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.83449 0.44857 1.86 0.071 .
## LastAuthorFemale1 0.12802 0.21209 0.60 0.550
## Year1997 1.00351 0.44857 2.24 0.032 *
## Year1998 0.16549 1.38474 0.12 0.906
## Year2000 0.56301 0.44866 1.25 0.218
## Year2002 0.92949 0.44658 2.08 0.045 *
## Year2003 -0.41549 0.56428 -0.74 0.466
## Year2004 -0.09669 0.55936 -0.17 0.864
## Year2005 0.69957 0.47340 1.48 0.148
## Year2006 -0.50549 0.44857 -1.13 0.267
## Year2007 0.29337 0.57167 0.51 0.611
## Year2008 -0.03835 0.45942 -0.08 0.934
```

```

## Year2009      0.08814      0.47579      0.19      0.854
## Year2010      0.00178      0.48317      0.00      0.997
## Year2011     -0.71350      0.44335     -1.61      0.117
## Year2012      0.35037      0.46256      0.76      0.454
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.501
## Multiple R-squared:  0.39,   Adjusted R-squared:  0.128
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 44 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.539  0.920  0.954  0.914  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      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 2709"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1998 1999 2001 2008 2011 2012
##    1    1    1    1    1    1
##
## 1998 1999 2001 2008 2011 2012
##    1    0    0    0    1    1
##
## 1998 1999 2001 2008 2011 2012
##    0    0    0    0    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: 2"
## [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 2006 2007 2008 2009 2010 2011
##    3    2    2    2    2    3    4    4    2    1    2    4    1    5    2
## 2012
##    3
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2006 2007 2008 2009 2010 2011
##    0    0    0    0    0    2    0    2    1    1    1    3    1    2    2
## 2012
##    3
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2006 2007 2008 2009 2010 2011
##    0    0    0    0    0    2    0    1    1    1    0    2    1    2    1
## 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: 5"
##
## 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: 5"
## [1] "Male last author team size 2018 geometric mean: 2"
##
## Wilcoxon rank sum test
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1, p-value = 1
## alternative hypothesis: true location shift is not equal to 0

```



```

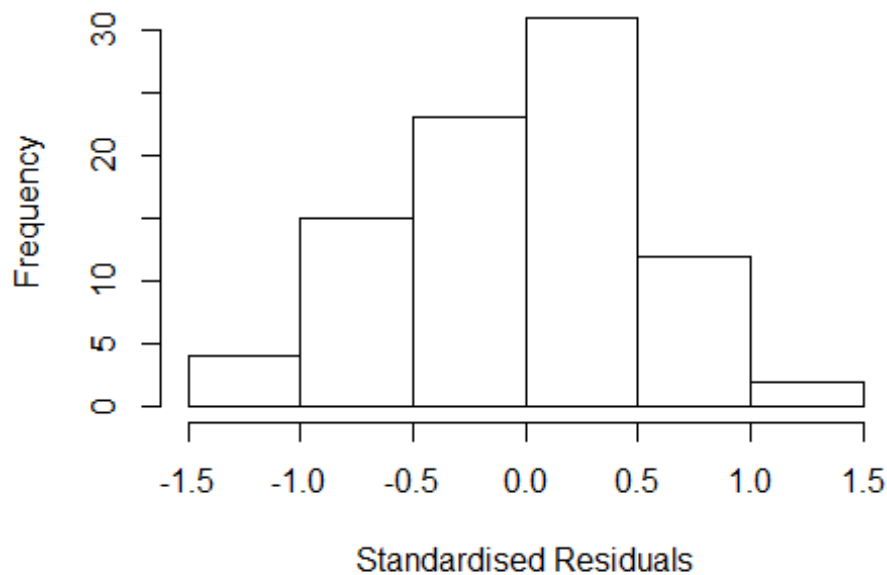
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
## [1] "Regression 3: First author gender, Year as factors"
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 13"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2711"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7    8    9    8   12    9    8    8   23   16   17   22   18   19   20
## 2011 2012
##   15   22
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    0    2    1    4    2    4    3   11    7   10   12   10    9   16
## 2011 2012
##   12   16
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    0    0    1    2    1    4    3    7    4    7   11    6    9   13
## 2011 2012
##    7   12
## [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: 4.71769398031653"
##
## Wilcoxon rank sum test
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0, p-value = 0.5
## 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: 4.71769398031653"
##
## Wilcoxon rank sum test
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0, 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 | 2.574e+13 | 1 | 5.074e+06 |
| ## LastAuthorFemale | 2.048e+00 | 1 | 1.431e+00 |
| ## UniqueAuthors | 2.384e+14 | 4 | 6.268e+01 |
| ## Year | 1.252e+15 | 13 | 3.808e+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
## -1.37441 -0.35614 0.00347 0.39732 1.10394
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.22170 0.32924 3.71 0.00042 ***
## FirstAuthorFemale1 -0.09318 0.17768 -0.52 0.60170
```

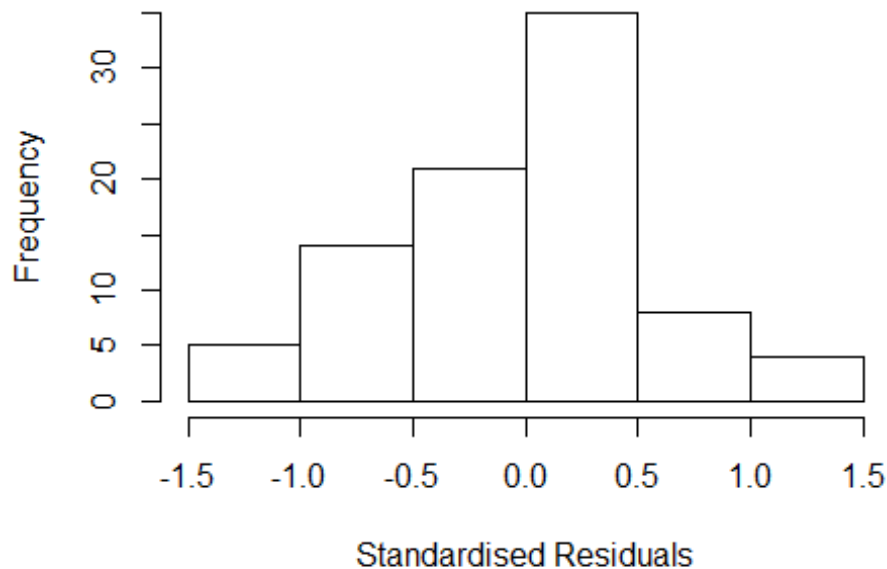
```

## LastAuthorFemale1    0.38996    0.17734    2.20    0.03134 *
## UniqueAuthors2      -0.46990    0.25600   -1.84    0.07086 .
## UniqueAuthors3      -0.40751    0.28850   -1.41    0.16242
## UniqueAuthors4       0.04449    0.24209    0.18    0.85476
## UniqueAuthors5       0.17883    0.30451    0.59    0.55900
## Year2000             -0.97133    0.29608   -3.28    0.00165 **
## Year2001             -0.53552    0.24429   -2.19    0.03185 *
## Year2002             -0.02265    0.24883   -0.09    0.92775
## Year2003             -0.01956    0.53729   -0.04    0.97106
## Year2004              0.00844    0.33974    0.02    0.98025
## Year2005             -0.21215    0.25572   -0.83    0.40969
## Year2006             -0.14734    0.13868   -1.06    0.29186
## Year2007             -0.19666    0.31175   -0.63    0.53032
## Year2008             -0.02612    0.43416   -0.06    0.95221
## Year2009              0.06757    0.24880    0.27    0.78678
## Year2010             -0.12117    0.22186   -0.55    0.58678
## Year2011             -0.21047    0.37290   -0.56    0.57436
## Year2012             -0.14693    0.25767   -0.57    0.57044
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.566
## Multiple R-squared:  0.245, Adjusted R-squared:  0.031
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 75 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.536  0.879  0.940  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      1.15e-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"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -2.474e+25  1          NaN

```

| | | | |
|---------------------|------------|----|-------|
| ## LastAuthorFemale | 3.179e+00 | 1 | 1.783 |
| ## Year | -5.942e+25 | 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
## -1.1324 -0.4232 0.0893 0.3755 1.3444
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3464 0.1834 7.34 2.8e-10 ***
## FirstAuthorFemale1 -0.1734 0.1834 -0.95 0.34761
## LastAuthorFemale1 0.3721 0.1901 1.96 0.05428 .
## Year2000 -1.0248 0.2889 -3.55 0.00069 ***
## Year2001 -0.4814 0.1834 -2.62 0.01060 *
## Year2002 -0.1917 0.1501 -1.28 0.20590
## Year2003 0.0793 0.5304 0.15 0.88158
## Year2004 -0.3051 0.2793 -1.09 0.27827
## Year2005 -0.2443 0.2523 -0.97 0.33603
```

```

## Year2006          -0.3398      0.1721    -1.97   0.05231 .
## Year2007          -0.4512      0.2576    -1.75   0.08423 .
## Year2008          -0.2164      0.3136    -0.69   0.49238
## Year2009          -0.2140      0.2909    -0.74   0.46440
## Year2010          -0.3747      0.2564    -1.46   0.14828
## Year2011          -0.3968      0.3434    -1.16   0.25173
## Year2012          -0.2873      0.2636    -1.09   0.27960
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.659
## Multiple R-squared:  0.0929, Adjusted R-squared:  -0.0988
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~1. The remaining 81 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.657  0.882  0.963  0.928  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      1.15e-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"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.287e+15 1      3.587e+07
## Year              1.287e+15 13      3.812e+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

```

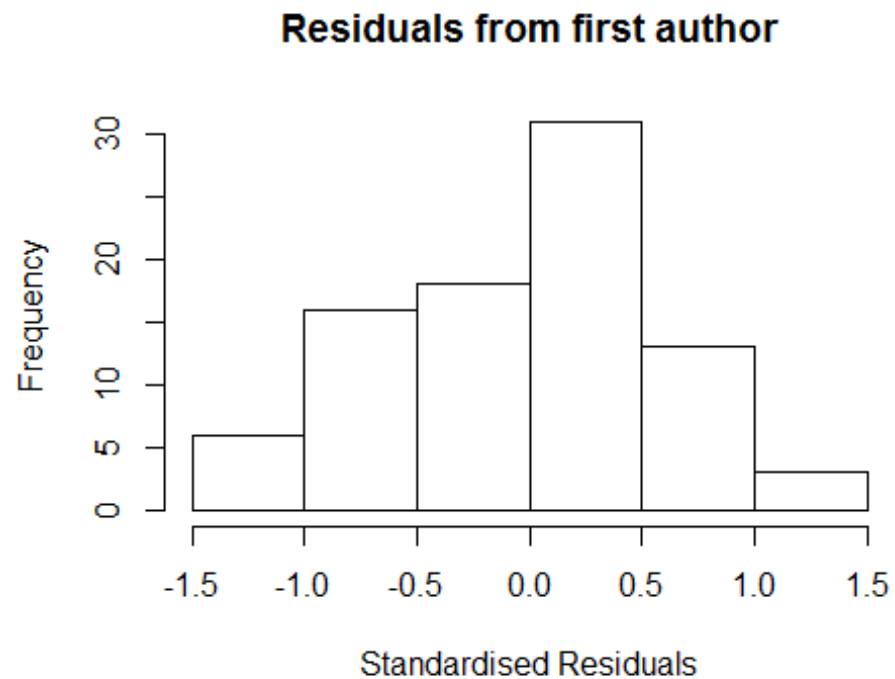
```

## -1.1993 -0.5010 0.0956 0.3976 1.3045
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2443    0.1744   7.13 6.3e-10 ***
## FirstAuthorFemale1 -0.0713    0.1744  -0.41  0.684
## Year2000          -0.7877    0.3639  -2.16  0.034 *
## Year2001          -0.3793    0.1744  -2.17  0.033 *
## Year2002          -0.1147    0.1556  -0.74  0.463
## Year2003           0.1942    0.4017   0.48  0.630
## Year2004          -0.2352    0.2621  -0.90  0.373
## Year2005          -0.0809    0.2868  -0.28  0.779
## Year2006          -0.2826    0.1409  -2.01  0.049 *
## Year2007          -0.2516    0.2222  -1.13  0.261
## Year2008          -0.0793    0.3423  -0.23  0.817
## Year2009          -0.0451    0.2780  -0.16  0.872
## Year2010          -0.2272    0.2585  -0.88  0.382
## Year2011          -0.2548    0.3277  -0.78  0.439
## Year2012          -0.1555    0.2599  -0.60  0.552
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.672
## Multiple R-squared:  0.0498, Adjusted R-squared:  -0.135
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 78 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.686 0.893 0.954 0.925 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      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"

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

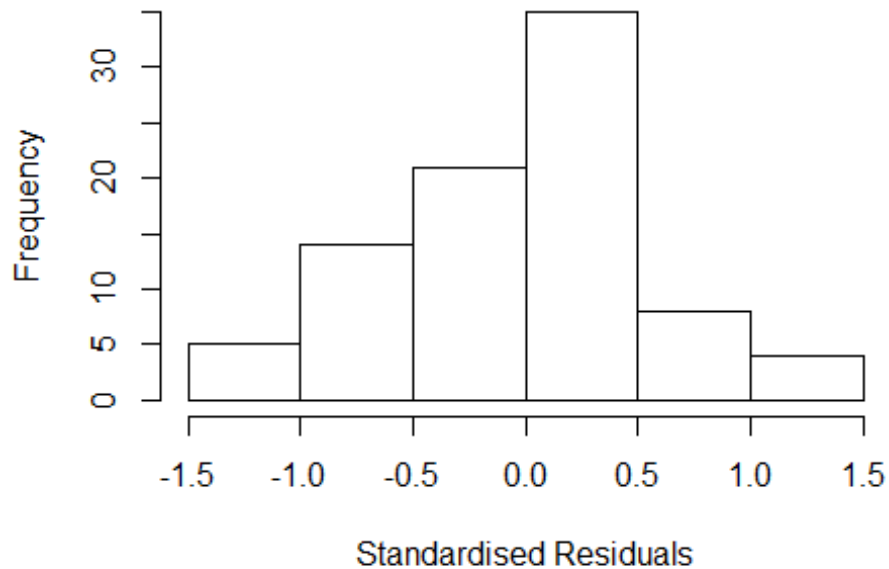
```

```
## 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
## -1.0817 -0.4344 0.0986 0.3664 1.4023
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1730 0.0000 Inf < 2e-16 ***
## LastAuthorFemale1 0.3095 0.1569 1.97 0.052 .
## Year2000 -0.9068 0.2096 -4.33 4.8e-05 ***
## Year2001 -0.3080 0.0000 -Inf < 2e-16 ***
## Year2002 -0.0606 0.0898 -0.68 0.502
## Year2003 0.0957 0.5128 0.19 0.852
## Year2004 -0.1828 0.2273 -0.80 0.424
## Year2005 -0.0976 0.1723 -0.57 0.573
## Year2006 -0.2939 0.1584 -1.86 0.068 .
## Year2007 -0.2745 0.1666 -1.65 0.104
## Year2008 -0.0913 0.3101 -0.29 0.769
## Year2009 -0.0939 0.2739 -0.34 0.733
```

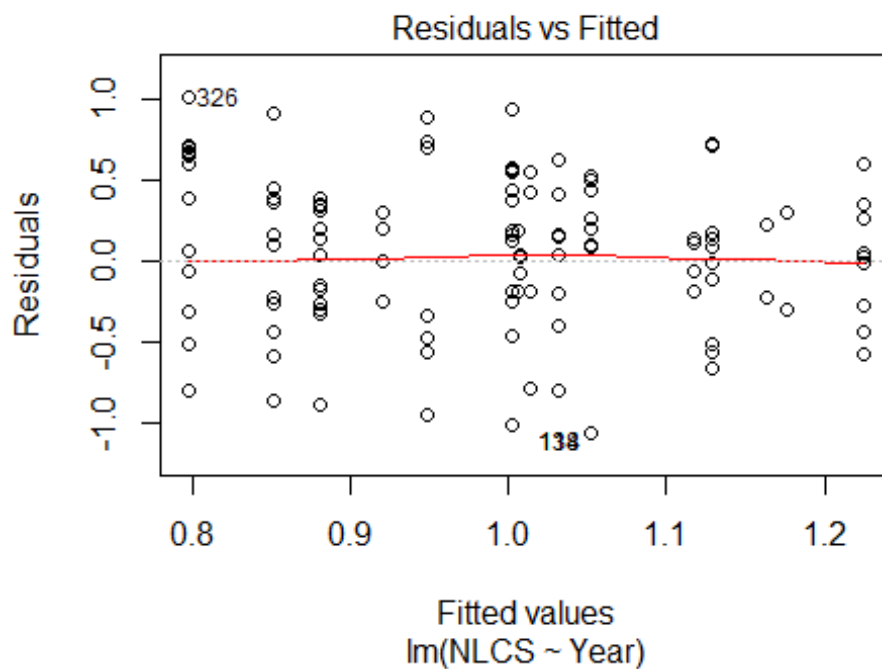


```

## Year2010          -0.2231      0.1934    -1.15      0.252
## Year2011          -0.2813      0.3610    -0.78      0.438
## Year2012          -0.1745      0.2358    -0.74      0.462
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.644
## Multiple R-squared:  0.0804, Adjusted R-squared:  -0.0984
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 79 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.615  0.868  0.962  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          1.15e-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 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
##   11    7   11    7    9    8    9   11   11   14   10   18   20   20   17
## 2011 2012
##   33   26
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    2    2    4    4    0    2    5    7    9    8   11   14    9   10
## 2011 2012
##   17   19
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010

```

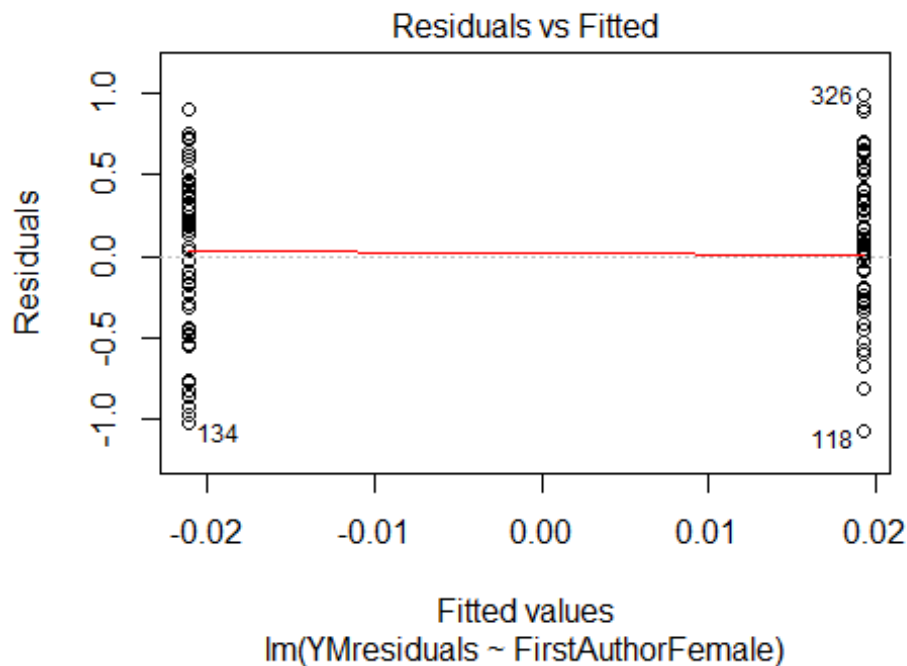
```
##      0      0      2      1      2      0      2      2      6      8      6      11      9      7      8
## 2011 2012
## 15 16
## [1] "Heteroscedasticity checks, confirming that there are problems with
## these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 22, df = 15, p-value = 0.1
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.4, df = 1, p-value = 0.5
##
## [1] "Female first author team size 2018 geometric mean: 5.24327781817326"
## [1] "Male first author team size 2018 geometric mean: 5.50156321196687"
##
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
```

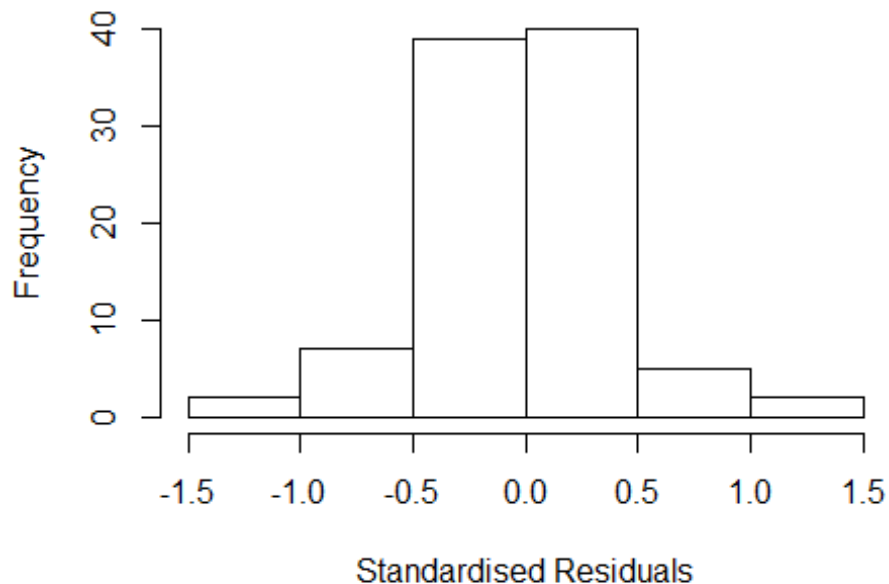
```
## W = 24, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.9895148616825"
## [1] "Male last author team size 2018 geometric mean: 6.632343027447"

## Warning in wilcox.test.default(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 = 0.2
## 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  7.549  1          2.748
## LastAuthorFemale  4.249  1          2.061
## UniqueAuthors    38.798  4          1.580
## Year             157.734 13          1.215
```

Residuals from first and last author and team size



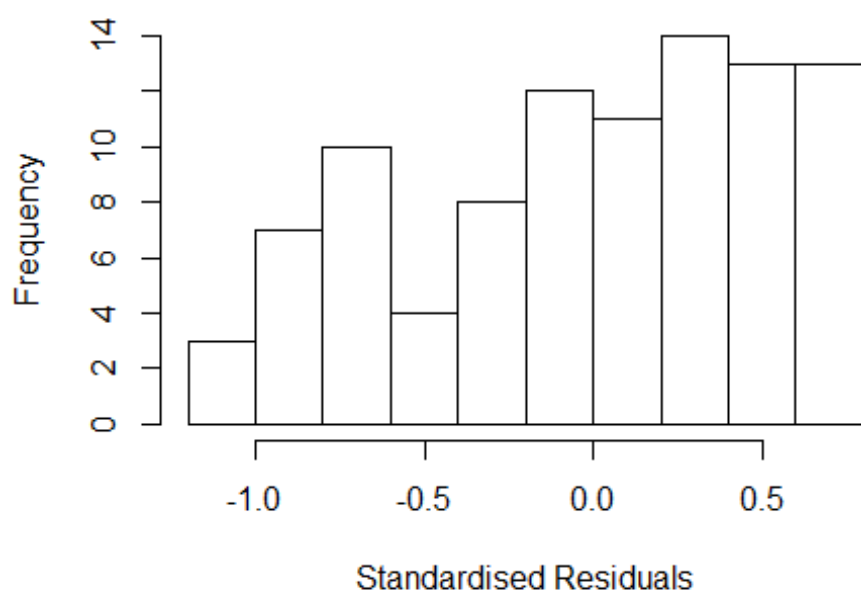
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.34e+00 -1.94e-01 1.50e-15 2.06e-01 1.42e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.1356 0.2270 0.60 0.552
## FirstAuthorFemale1 -0.0351 0.1076 -0.33 0.745
## LastAuthorFemale1 0.3223 0.1006 3.20 0.002 **
## UniqueAuthors2 0.0689 0.1677 0.41 0.682
## UniqueAuthors3 0.8861 0.3062 2.89 0.005 **
## UniqueAuthors4 0.9678 0.1847 5.24 1.4e-06 ***
## UniqueAuthors5 1.1491 0.1864 6.16 3.3e-08 ***
## Year1999 -0.1333 0.1612 -0.83 0.411
## Year2000 -0.3017 0.6237 -0.48 0.630
## Year2002 -0.2436 0.2121 -1.15 0.254
```

```

## Year2003          -0.3066      0.2717    -1.13      0.263
## Year2004           0.0603      0.3365      0.18      0.858
## Year2005           0.0596      0.1688      0.35      0.725
## Year2006          -0.1908      0.2510     -0.76      0.449
## Year2007          -0.0289      0.1795     -0.16      0.873
## Year2008          -0.2498      0.1687     -1.48      0.143
## Year2009           0.0171      0.1941      0.09      0.930
## Year2010          -0.0822      0.2705     -0.30      0.762
## Year2011          -0.0581      0.1964     -0.30      0.768
## Year2012          -0.0644      0.1793     -0.36      0.721
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.339
## Multiple R-squared:  0.634, Adjusted R-squared:  0.542
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 87 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0392 0.8780 0.9590 0.8780 0.9830 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      1.05e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.873 1      1.369
## LastAuthorFemale 1.795 1      1.340
## Year              3.131 13      1.045

```

Residuals from first and last author



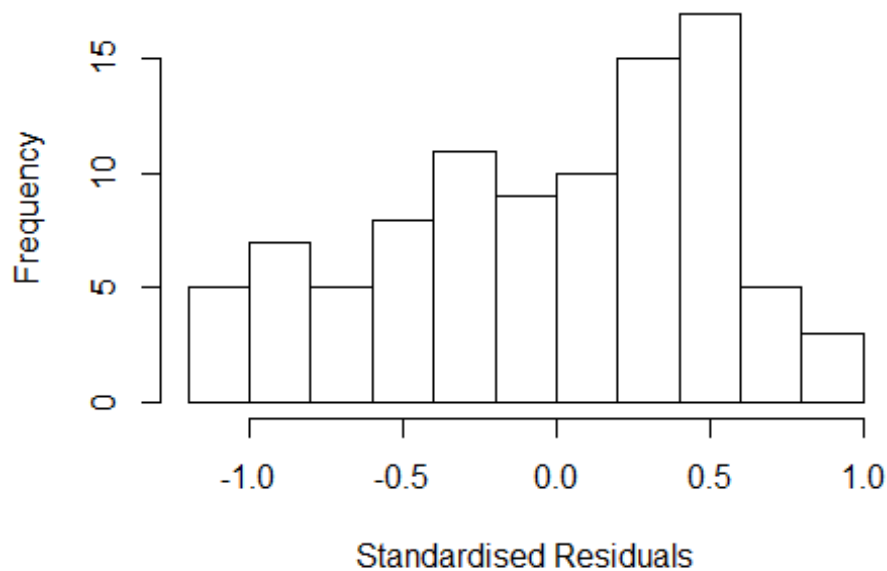
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0573 -0.4080 0.0881 0.4175 0.7798
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1875 0.2251 5.28 1.1e-06 ***
## FirstAuthorFemale1 -0.0221 0.1347 -0.16 0.870
## LastAuthorFemale1 0.2350 0.1363 1.72 0.089 .
## Year1999 -0.2305 0.2251 -1.02 0.309
## Year2000 -0.3425 0.5995 -0.57 0.569
## Year2002 -0.1595 0.2612 -0.61 0.543
## Year2003 -0.2225 0.3015 -0.74 0.463
## Year2004 -0.1065 0.4214 -0.25 0.801
## Year2005 -0.1302 0.3194 -0.41 0.685
## Year2006 -0.1773 0.2872 -0.62 0.539
## Year2007 -0.4252 0.2768 -1.54 0.128
## Year2008 -0.4105 0.2597 -1.58 0.118
```

```

## Year2009          0.0682      0.2612      0.26      0.795
## Year2010          -0.1891     0.2971     -0.64     0.526
## Year2011          -0.2188     0.2793     -0.78     0.436
## Year2012          -0.3433     0.2920     -1.18     0.243
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.574
## Multiple R-squared:  0.0864, Adjusted R-squared:  -0.087
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 89 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.715  0.886  0.947  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      1.05e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.885 1          1.373
## Year              1.885 13          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
## -1.144 -0.459  0.080  0.415  0.905
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1526     0.2531   4.55 1.9e-05 ***
## FirstAuthorFemale1  0.0479     0.1368   0.35  0.73
## Year1999         -0.2654     0.2531  -1.05  0.30
## Year2000         -0.3425     0.5652  -0.61  0.55
## Year2002         -0.1944     0.2857  -0.68  0.50
## Year2003         -0.2574     0.3229  -0.80  0.43
## Year2004         -0.0522     0.4152  -0.13  0.90
## Year2005         -0.0569     0.3406  -0.17  0.87
## Year2006         -0.0876     0.3142  -0.28  0.78
## Year2007         -0.3324     0.2938  -1.13  0.26
## Year2008         -0.3341     0.2814  -1.19  0.24
## Year2009          0.1064     0.2807   0.38  0.71
```



```

## Year2010          -0.0636      0.3124   -0.20      0.84
## Year2011          -0.1325      0.2989   -0.44      0.66
## Year2012          -0.3000      0.3210   -0.93      0.35
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.576
## Multiple R-squared:  0.0593, Adjusted R-squared:  -0.105
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 90 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.673  0.896  0.942  0.921  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      1.05e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.804 1          1.343
## Year              1.804 13          1.023
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.0591 -0.4119  0.0842  0.4301  0.7761
##
## Coefficients:
##      Estimate Std. Error t value Pr(>|t|)

```

```

## (Intercept)          1.1765      0.2233      5.27  1.1e-06 ***
## LastAuthorFemale1    0.2282      0.1371      1.66    0.10 .
## Year1999             -0.2415      0.2233     -1.08    0.28
## Year2000             -0.3425      0.5813     -0.59    0.56
## Year2002             -0.1705      0.2596     -0.66    0.51
## Year2003             -0.2335      0.2999     -0.78    0.44
## Year2004             -0.0977      0.4177     -0.23    0.82
## Year2005             -0.1323      0.3253     -0.41    0.69
## Year2006             -0.1749      0.2944     -0.59    0.55
## Year2007             -0.4258      0.2835     -1.50    0.14
## Year2008             -0.4108      0.2674     -1.54    0.13
## Year2009              0.0620      0.2653      0.23    0.82
## Year2010             -0.1881      0.3017     -0.62    0.53
## Year2011             -0.2170      0.2853     -0.76    0.45
## Year2012             -0.3455      0.2968     -1.16    0.25
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.588
## Multiple R-squared:  0.086, Adjusted R-squared:  -0.0739
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 89 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.726  0.892  0.947   0.925  0.984  0.999
## Algorithmic parameters:
##           tuning.chi           bb           tuning.psi           refine.tol
##           1.55e+00           5.00e-01           4.69e+00           1.00e-07
##           rel.tol           solve.tol           eps.outlier           eps.x
##           1.00e-07           1.00e-07           1.05e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01           5.00e-01
##   nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500          50          2          1          1000          200
##   trace.lev      mts      compute.rd
##           0          1000          0
##           psi           subsampling           cov
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 95"
## [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]"

```

```

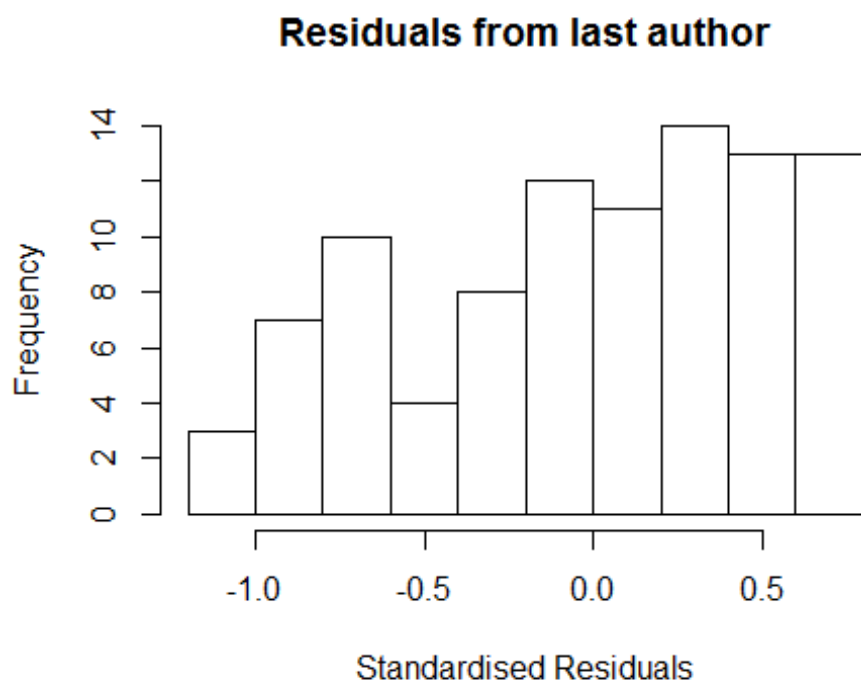
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    3    5    2    7    3    3    3    3    8    3   10   16   13   21
## 2011 2012
##   17   11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    0    5    1    0    0    2    0    1    7    1    4    9    6    9
## 2011 2012
##    7    4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    0    4    1    0    0    2    0    1    7    1    2    5    4    9
## 2011 2012
##    7    4
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.72791927319135"
## [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 = 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: 3.46410161513775"
## [1] "Male last author team size 2018 geometric mean: 3.68403149864039"

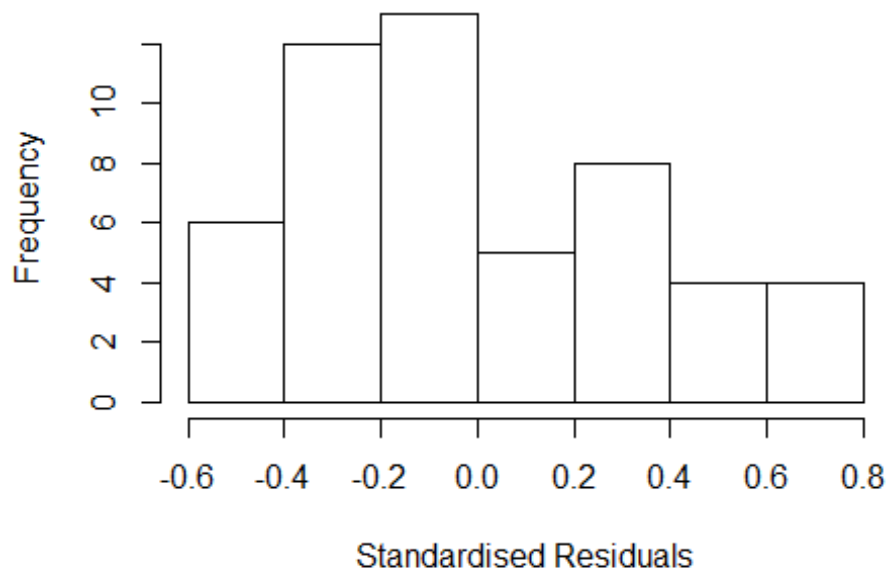
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7, p-value = 0.9
## 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 6.954e+00 1      2.637
## LastAuthorFemale -2.122e+15 1      NaN
## UniqueAuthors    -1.738e+31 4      NaN
## Year              -1.891e+32 12     NaN
```

Residuals from first and last author and team size



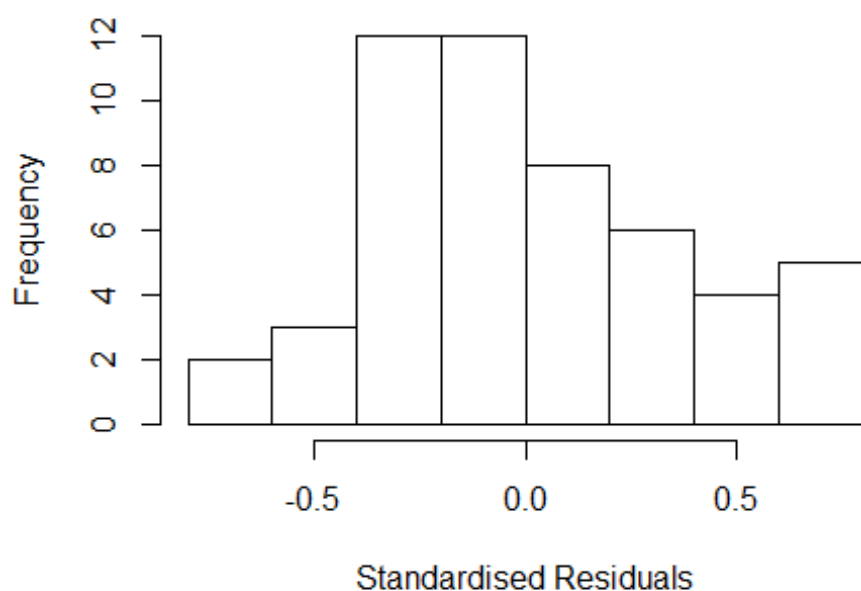
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.579 -0.262 -0.021 0.302 0.734
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.3227 0.2154 1.50 0.144
## FirstAuthorFemale1 0.2279 0.1321 1.73 0.094 .
## LastAuthorFemale1 -0.0344 0.1785 -0.19 0.848
## UniqueAuthors2 0.3475 0.2841 1.22 0.230
## UniqueAuthors3 0.3315 0.3710 0.89 0.378
## UniqueAuthors4 0.0981 0.2739 0.36 0.722
## UniqueAuthors5 0.2771 0.3109 0.89 0.379
## Year1998 0.6234 0.3709 1.68 0.102
## Year1999 0.4619 0.4176 1.11 0.277
## Year2002 0.0408 0.3474 0.12 0.907
```

```

## Year2004          0.6317      0.4049      1.56      0.128
## Year2005         -0.0617      0.3791     -0.16      0.872
## Year2006          0.3353      0.2154      1.56      0.129
## Year2007          0.2566      0.6156      0.42      0.680
## Year2008          0.1810      0.2903      0.62      0.537
## Year2009          0.2810      0.4425      0.64      0.530
## Year2010          0.5893      0.3955      1.49      0.146
## Year2011          0.4534      0.3908      1.16      0.254
## Year2012          0.3923      0.3965      0.99      0.330
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.496
## Multiple R-squared:  0.426, Adjusted R-squared:  0.113
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 44 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.810  0.906  0.963   0.946   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      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 7.282e+01 1      8.533
## LastAuthorFemale -8.165e+14 1      NaN
## Year -7.856e+17 12      NaN

```

Residuals from first and last author



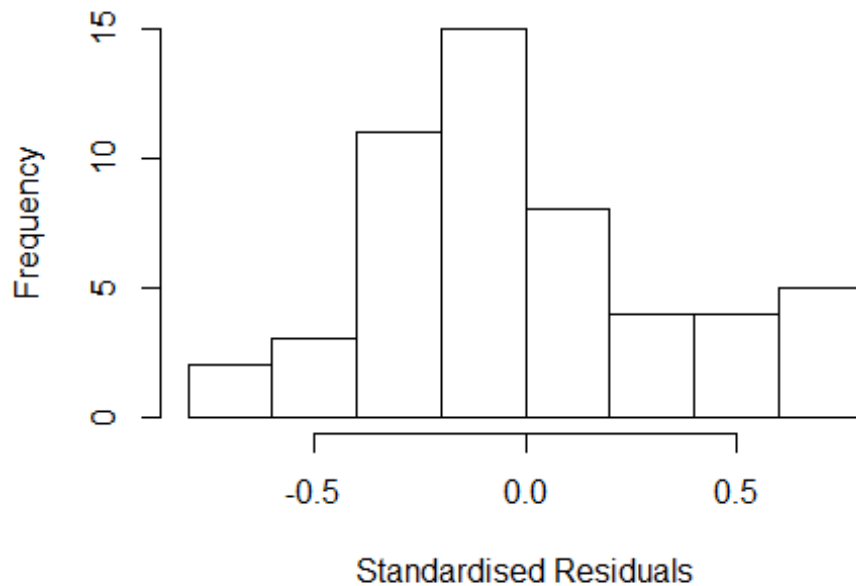
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.73081 -0.26710 -0.00173 0.22584 0.73319
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.3895 0.2031 1.92 0.06289 .
## FirstAuthorFemale1 0.2471 0.1179 2.10 0.04307 *
## LastAuthorFemale1 -0.0797 0.1498 -0.53 0.59785
## Year1998 0.7550 0.3378 2.23 0.03155 *
## Year1999 0.7265 0.2031 3.58 0.00099 ***
## Year2002 0.1519 0.2078 0.73 0.46928
## Year2004 0.8871 0.2154 4.12 0.00021 ***
## Year2005 0.1396 0.2413 0.58 0.56652
## Year2006 0.2685 0.2031 1.32 0.19439
## Year2007 0.3413 1.2046 0.28 0.77852
## Year2008 0.2564 0.2289 1.12 0.26998
## Year2009 0.4410 0.2988 1.48 0.14849
```

```

## Year2010          0.8169      0.2617      3.12  0.00348 **
## Year2011          0.6707      0.2372      2.83  0.00754 **
## Year2012          0.6246      0.2575      2.43  0.02029 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.444
## Multiple R-squared:  0.402, Adjusted R-squared:  0.175
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 44 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.767  0.874  0.958  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      1.92e-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 -797.4 1      NaN
## Year              -797.4 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
## -0.717 -0.285 -0.020  0.206  0.731
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.350      0.189   1.85  0.07229 .
## FirstAuthorFemale1 0.233      0.124   1.88  0.06748 .
## Year1998          0.766      0.338   2.26  0.02940 *
## Year1999          0.766      0.189   4.05  0.00024 ***
## Year2002          0.199      0.195   1.02  0.31482
## Year2004          0.847      0.189   4.48  6.6e-05 ***
## Year2005          0.176      0.240   0.73  0.46865
## Year2006          0.308      0.189   1.63  0.11100
## Year2007          0.348      1.072   0.33  0.74689
## Year2008          0.302      0.217   1.39  0.17240
## Year2009          0.404      0.282   1.43  0.16108
## Year2010          0.817      0.255   3.20  0.00278 **
```

```

## Year2011          0.683      0.228      3.00  0.00478 **
## Year2012          0.635      0.258      2.46  0.01859 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.434
## Multiple R-squared:  0.398, Adjusted R-squared:  0.192
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 44 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.759  0.889  0.954  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      1.92e-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 2.992e+14  1      1.73e+07
## Year              2.992e+14 12      4.01e+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
## -8.24e-01 -2.99e-01 -3.33e-16  2.75e-01  8.36e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.4340      0.2614   1.66   0.1051

```

```

## LastAuthorFemale1 -0.0196    0.1735   -0.11    0.9105
## Year1998          0.7975    0.3603    2.21    0.0330 *
## Year1999          0.6820    0.2614    2.61    0.0129 *
## Year2002          0.2310    0.2768    0.83    0.4092
## Year2004          0.7827    0.3037    2.58    0.0140 *
## Year2005          0.2310    0.2915    0.79    0.4331
## Year2006          0.2240    0.2614    0.86    0.3968
## Year2007          0.3903    2.2654    0.17    0.8641
## Year2008          0.3137    0.2946    1.06    0.2937
## Year2009          0.4141    0.3661    1.13    0.2651
## Year2010          0.8936    0.3163    2.83    0.0075 **
## Year2011          0.6963    0.2963    2.35    0.0241 *
## Year2012          0.7402    0.2873    2.58    0.0140 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.445
## Multiple R-squared:  0.351, Adjusted R-squared:  0.129
## Convergence in 46 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 44 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.705  0.913  0.953   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] "Sample size for the above analysis: 52"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2714"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2001 2002 2004 2006 2007 2008 2009 2010 2011 2012

```

```

##      3      4      4      1      4      2      2      5      6      4      4      7      7      3
##
## 1996 1997 1998 1999 2001 2002 2004 2006 2007 2008 2009 2010 2011 2012
##      2      2      3      1      0      1      1      5      6      4      3      6      5      2
##
## 1996 1997 1998 1999 2001 2002 2004 2006 2007 2008 2009 2010 2011 2012
##      2      2      3      1      0      0      0      4      6      4      3      6      5      2
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.04282321707986"
## [1] "Male first 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 = 4, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.16179145028782"
## [1] "Male last author team size 2018 geometric mean: 5.53264746889037"

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

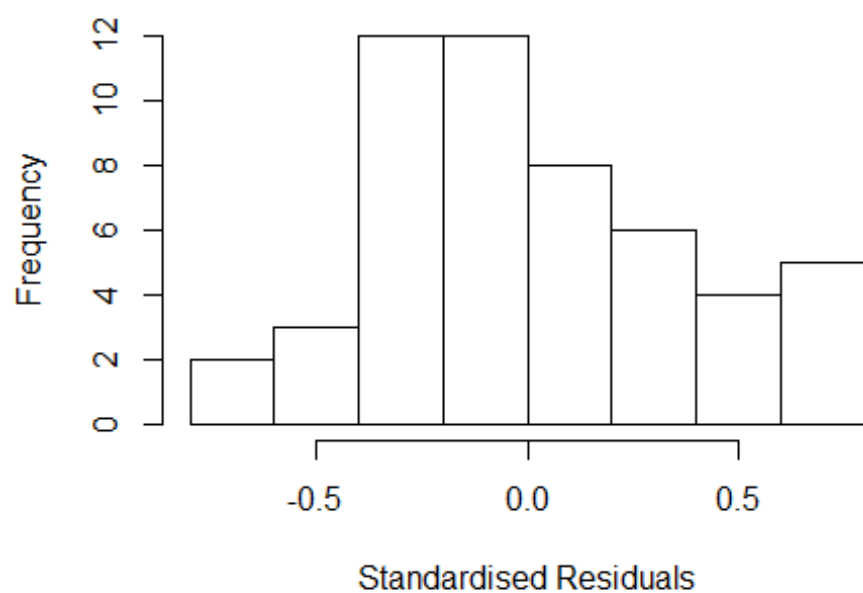
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 6, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"

## 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

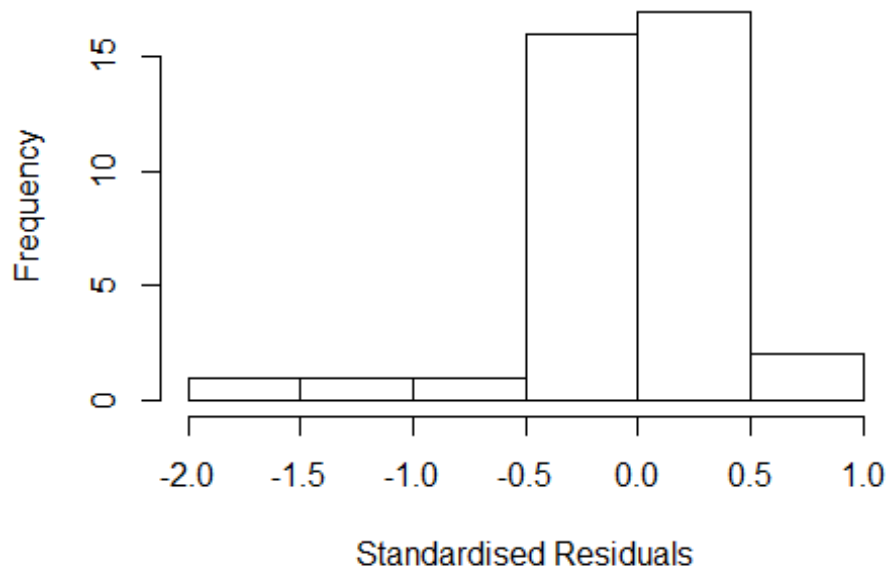
```

Residuals from last author



| | | | | |
|----|-------------------|------|----|-----------------|
| ## | | GVIF | Df | GVIF^(1/(2*Df)) |
| ## | FirstAuthorFemale | NaN | 1 | NaN |
| ## | LastAuthorFemale | NaN | 1 | NaN |
| ## | UniqueAuthors | NaN | 4 | NaN |
| ## | Year | NaN | 10 | NaN |

Residuals from first and last author and team size



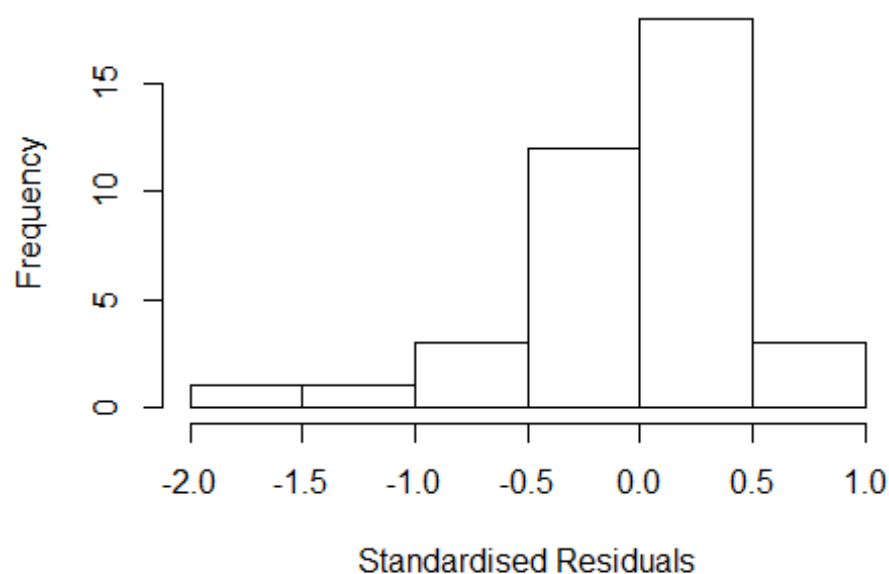
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.8497 -0.2581 0.0108 0.2232 0.6694
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.858 0.479 1.79 0.0877 .
## FirstAuthorFemale1 0.389 0.121 3.22 0.0041 **
## LastAuthorFemale1 -0.271 0.000 -Inf <2e-16 ***
## UniqueAuthors2 0.383 0.298 1.29 0.2121
## UniqueAuthors3 -0.015 0.330 -0.05 0.9642
## UniqueAuthors4 -0.114 0.342 -0.33 0.7411
## UniqueAuthors5 0.227 0.305 0.75 0.4645
## Year1997 0.962 0.476 2.02 0.0561 .
## Year1998 0.766 0.821 0.93 0.3613
## Year1999 -0.399 0.454 -0.88 0.3892
```

```

## Year2006          0.429      0.474      0.90      0.3758
## Year2007          0.518      0.501      1.03      0.3129
## Year2008          0.871      0.517      1.69      0.1066
## Year2009          1.300      0.436      2.98      0.0072 **
## Year2010          0.988      0.580      1.70      0.1034
## Year2011          0.244      0.532      0.46      0.6515
## Year2012          1.389      0.437      3.18      0.0045 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.47
## Multiple R-squared:  0.657, Adjusted R-squared:  0.395
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 33 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0873 0.9110 0.9750 0.9170 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      2.63e-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 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 36.32 1      6.027
## LastAuthorFemale  12.56 1      3.544
## Year              84.89 10     1.249

```

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
## -1.9783 -0.2471 0.0287 0.1529 0.6667
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8005 0.3628 2.21 0.0368 *
## FirstAuthorFemale1 0.3853 0.1551 2.48 0.0200 *
## LastAuthorFemale1 -0.2787 0.0885 -3.15 0.0042 **
## Year1997 1.2162 0.3941 3.09 0.0049 **
## Year1998 0.8258 0.7216 1.14 0.2632
## Year1999 -0.4565 0.3628 -1.26 0.2199
## Year2006 0.6593 0.4176 1.58 0.1270
## Year2007 0.7334 0.4201 1.75 0.0931 .
## Year2008 0.9381 0.4661 2.01 0.0550 .
## Year2009 1.5204 0.4607 3.30 0.0029 **
## Year2010 1.0712 0.4267 2.51 0.0189 *
## Year2011 0.3967 0.4310 0.92 0.3661
```

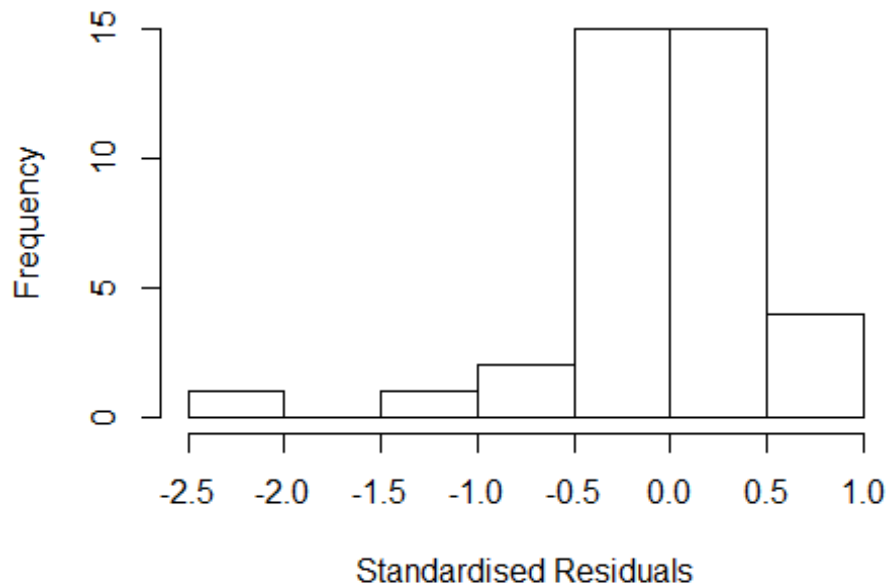


```

## Year2012          1.3872      0.3922      3.54      0.0016 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.467
## Multiple R-squared:  0.634, Adjusted R-squared:  0.458
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 31 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0335 0.8950 0.9580 0.9030 0.9950 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          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 6.583 1      2.566
## Year              6.583 10      1.099

```

Residuals from first author



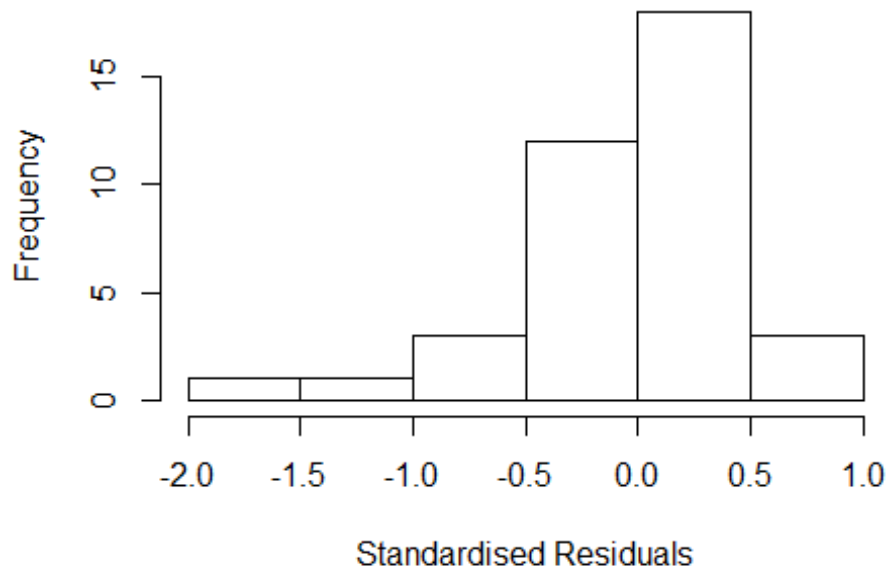
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -2.2416 -0.3077 0.0187 0.2496 0.7105
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.801 0.345 2.32 0.02847 *
## FirstAuthorFemale1 0.374 0.135 2.78 0.01000 *
## Year1997 1.082 0.360 3.00 0.00584 **
## Year1998 0.782 0.654 1.20 0.24235
## Year1999 -0.457 0.345 -1.32 0.19739
## Year2006 0.524 0.407 1.29 0.20908
## Year2007 0.641 0.392 1.64 0.11403
## Year2008 0.952 0.438 2.17 0.03906 *
## Year2009 1.443 0.388 3.72 0.00098 ***
## Year2010 1.067 0.408 2.61 0.01472 *
## Year2011 0.334 0.410 0.81 0.42344
## Year2012 1.253 0.376 3.34 0.00256 **
```

```

## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.543
## Multiple R-squared:  0.595, Adjusted R-squared:  0.423
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 3 weights are ~= 1. The remaining 35 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0502 0.9410 0.9690 0.9270 0.9910 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           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 8.055 1 2.838
## Year            8.055 10 1.110

```

Residuals from last author



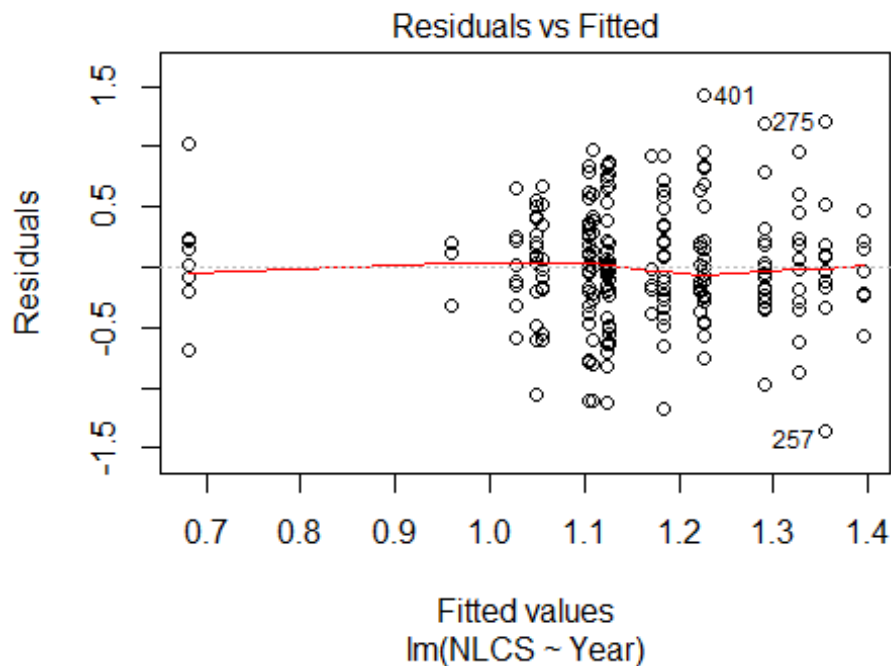
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.7193 -0.2345 0.0124 0.2239 0.8343
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.800 0.357 2.24 0.03372 *
## LastAuthorFemale1 -0.314 0.209 -1.51 0.14429
## Year1997 1.427 0.441 3.24 0.00329 **
## Year1998 0.810 0.699 1.16 0.25668
## Year1999 -0.456 0.357 -1.28 0.21238
## Year2006 0.874 0.400 2.19 0.03798 *
## Year2007 1.026 0.405 2.53 0.01766 *
## Year2008 1.243 0.379 3.28 0.00297 **
## Year2009 1.798 0.392 4.58 0.00010 ***
## Year2010 1.233 0.534 2.31 0.02919 *
## Year2011 0.670 0.442 1.52 0.14179
## Year2012 1.598 0.418 3.82 0.00074 ***
```

```

## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.487
## Multiple R-squared:  0.544, Adjusted R-squared:  0.351
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 32 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.187  0.911  0.977  0.903  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           2.63e-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 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
##    15   20   21   18   25   26   23   16   30   19   24   38   35   26   42
## 2011 2012
##    47   46
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     3    6    8    9   13   14    8   10   16    9   12   25   18   16   24
## 2011 2012
##    27   26
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     3    4    5    5   12   13    6    8   11    4   11   22   13   11   20
## 2011 2012
##    24   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 = 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

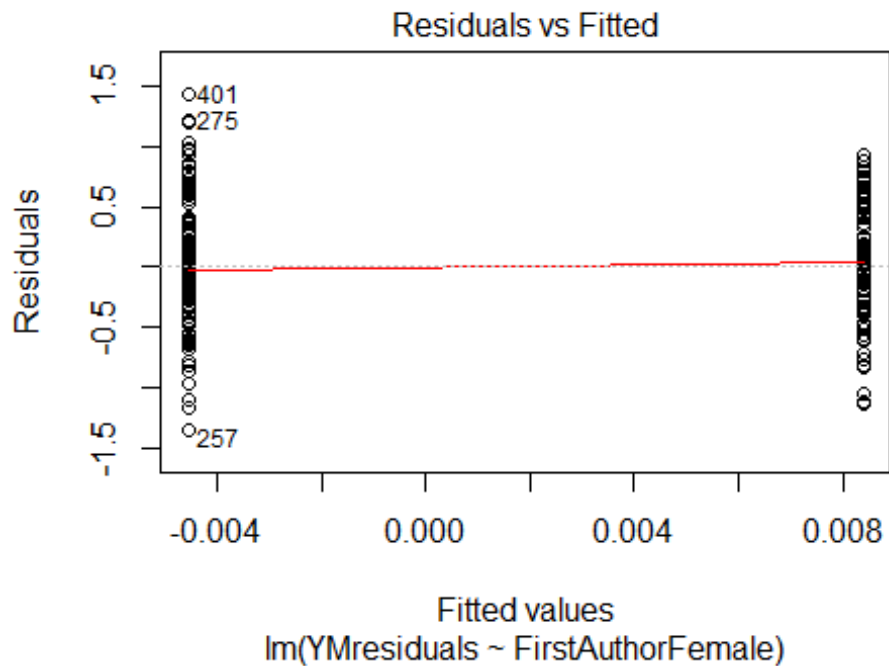
## [1] "Female first author team size 2018 geometric mean: 7.69438074227847"
## [1] "Male first author team size 2018 geometric mean: 7.57081313039624"

## Warning in wilcox.test.default(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 = 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: 8.15004426887511"

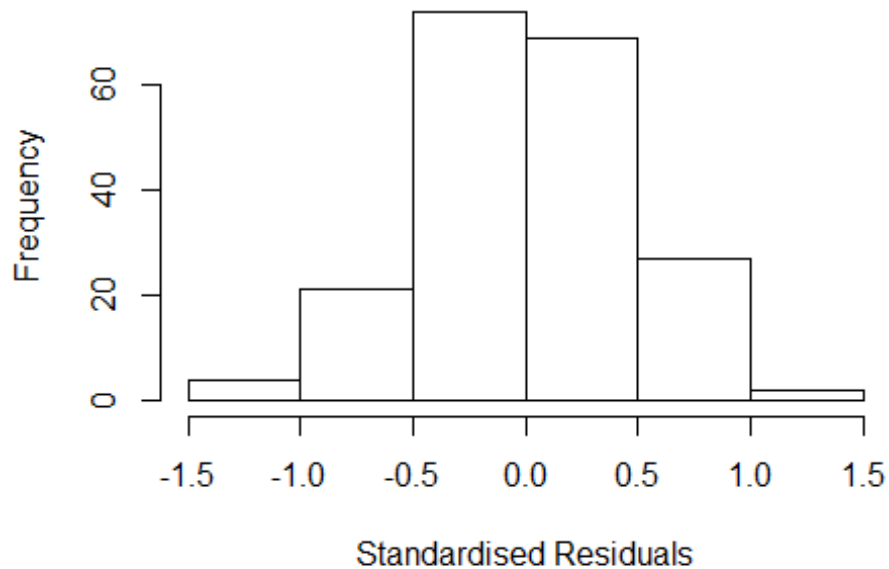
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## 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 | 1.844 | 1 | 1.358 |
| LastAuthorFemale | 1.492 | 1 | 1.221 |
| UniqueAuthors | 8.430 | 4 | 1.305 |
| Year | 17.063 | 16 | 1.093 |

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
## -1.28230 -0.28615 -0.00804 0.26119 1.46218
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.28504 0.21477 5.98 1.2e-08 ***
## FirstAuthorFemale1 0.08720 0.08210 1.06 0.290
## LastAuthorFemale1 0.09205 0.10802 0.85 0.395
## UniqueAuthors2 -0.17055 0.16562 -1.03 0.305
## UniqueAuthors3 -0.43656 0.17168 -2.54 0.012 *
## UniqueAuthors4 -0.31091 0.15820 -1.97 0.051 .
## UniqueAuthors5 -0.03955 0.13041 -0.30 0.762
## Year1997 -0.23597 0.19926 -1.18 0.238
## Year1998 0.06218 0.20062 0.31 0.757
## Year1999 -0.32886 0.30519 -1.08 0.283
```

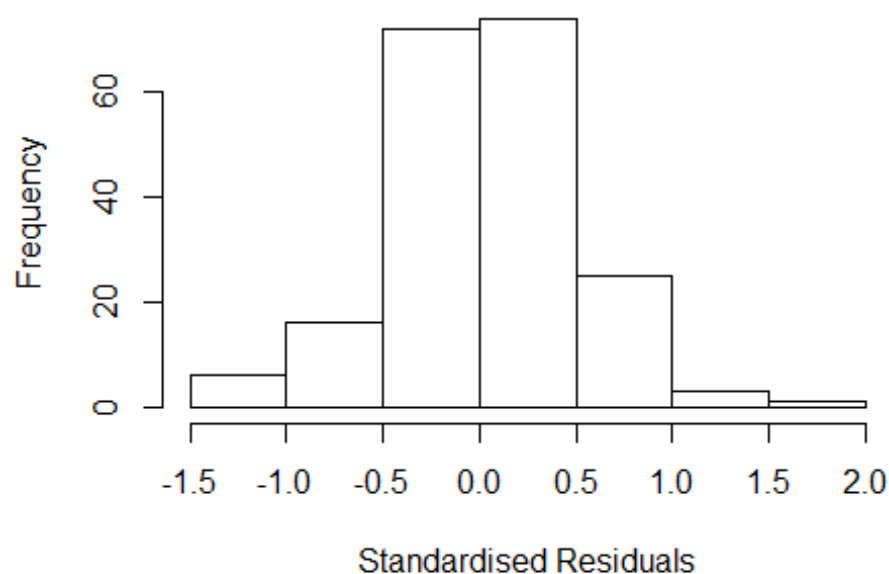


```

## Year2000      0.13180      0.22535      0.58      0.559
## Year2001      0.02532      0.24279      0.10      0.917
## Year2002     -0.18834      0.22183     -0.85      0.397
## Year2003     -0.09236      0.21532     -0.43      0.669
## Year2004     -0.17063      0.23016     -0.74      0.459
## Year2005     -0.01948      0.23849     -0.08      0.935
## Year2006      0.08767      0.21104      0.42      0.678
## Year2007     -0.13313      0.22081     -0.60      0.547
## Year2008      0.00446      0.19855      0.02      0.982
## Year2009     -0.06846      0.25157     -0.27      0.786
## Year2010     -0.05567      0.22319     -0.25      0.803
## Year2011     -0.06265      0.20289     -0.31      0.758
## Year2012     -0.05039      0.20907     -0.24      0.810
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.435
## Multiple R-squared:  0.13,   Adjusted R-squared:  0.0198
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 173 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.235  0.864  0.953   0.895   0.985   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      5.08e-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 1.615 1      1.271
## LastAuthorFemale  1.389 1      1.179
## Year              2.199 16      1.025

```

Residuals from first and last author



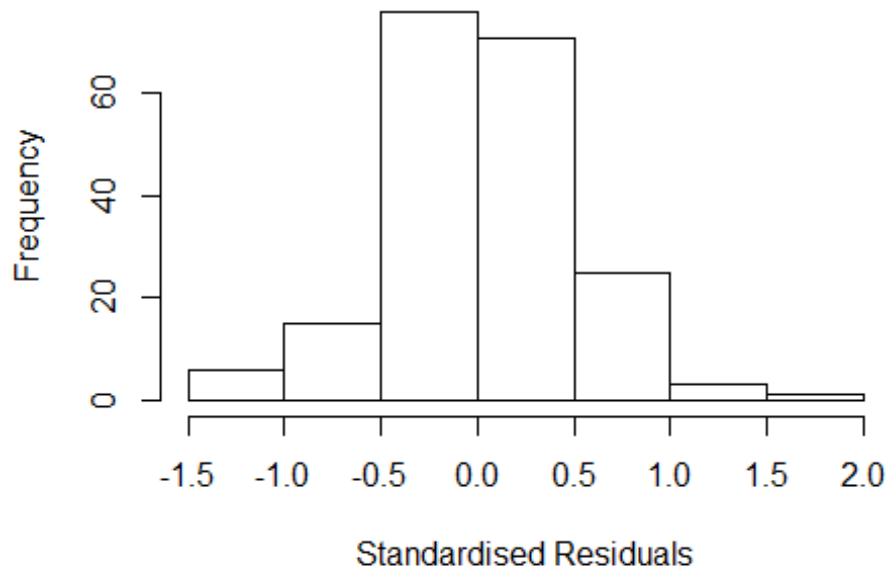
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2906 -0.2844 0.0116 0.2546 1.5336
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.94723 0.13422 7.06 3.7e-11 ***
## FirstAuthorFemale1 0.04809 0.08098 0.59 0.553
## LastAuthorFemale1 0.08897 0.10693 0.83 0.407
## Year1997 0.01660 0.15037 0.11 0.912
## Year1998 0.24055 0.18666 1.29 0.199
## Year1999 -0.17331 0.29856 -0.58 0.562
## Year2000 0.37054 0.20217 1.83 0.068 .
## Year2001 0.24959 0.19664 1.27 0.206
## Year2002 0.08195 0.18932 0.43 0.666
## Year2003 0.07295 0.18610 0.39 0.696
## Year2004 0.09223 0.19197 0.48 0.631
## Year2005 0.19552 0.23804 0.82 0.413
```

```

## Year2006          0.34340    0.15839    2.17    0.031 *
## Year2007          0.12363    0.17968    0.69    0.492
## Year2008          0.24749    0.15349    1.61    0.109
## Year2009         -0.00383    0.27290   -0.01    0.989
## Year2010          0.17119    0.20004    0.86    0.393
## Year2011          0.20316    0.16348    1.24    0.216
## Year2012          0.20335    0.16563    1.23    0.221
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.425
## Multiple R-squared:  0.0628, Adjusted R-squared:  -0.0319
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 181 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.166  0.859  0.952  0.886  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      5.08e-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 1.595 1      1.263
## Year              1.595 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
## -1.29743 -0.28747 0.00674 0.26507 1.53707
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9436 0.1335 7.07 3.4e-11 ***
## FirstAuthorFemale1 0.0584 0.0810 0.72 0.472
## Year1997 0.0151 0.1509 0.10 0.920
## Year1998 0.2601 0.1829 1.42 0.157
## Year1999 -0.1737 0.2978 -0.58 0.560
## Year2000 0.3694 0.2013 1.83 0.068 .
## Year2001 0.2513 0.1962 1.28 0.202
## Year2002 0.0967 0.1834 0.53 0.599
## Year2003 0.0847 0.1835 0.46 0.645
## Year2004 0.1225 0.1826 0.67 0.503
## Year2005 0.1946 0.2372 0.82 0.413
## Year2006 0.3538 0.1549 2.28 0.024 *
```

```

## Year2007          0.1404      0.1794      0.78      0.435
## Year2008          0.2586      0.1532      1.69      0.093 .
## Year2009          0.0312      0.2549      0.12      0.903
## Year2010          0.1713      0.1997      0.86      0.392
## Year2011          0.2144      0.1641      1.31      0.193
## Year2012          0.2164      0.1661      1.30      0.194
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.421
## Multiple R-squared:  0.0596, Adjusted R-squared:  -0.0298
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 180 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.155  0.846  0.949  0.884  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      5.08e-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 1.376 1      1.173
## Year      1.376 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
## -1.31021 -0.29509  0.00479  0.25727  1.52927

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9643    0.1349    7.15 2.1e-11 ***
## LastAuthorFemale1 0.1002    0.1073    0.93  0.352
## Year1997        0.0234    0.1500    0.16  0.876
## Year1998        0.2400    0.1846    1.30  0.195
## Year1999       -0.1756    0.3054   -0.57  0.566
## Year2000        0.3744    0.2077    1.80  0.073 .
## Year2001        0.2491    0.1987    1.25  0.212
## Year2002        0.0844    0.1937    0.44  0.664
## Year2003        0.0779    0.1875    0.42  0.678
## Year2004        0.0848    0.1949    0.43  0.664
## Year2005        0.1960    0.2441    0.80  0.423
## Year2006        0.3459    0.1625    2.13  0.035 *
## Year2007        0.1126    0.1783    0.63  0.529
## Year2008        0.2526    0.1572    1.61  0.110
## Year2009       -0.0222    0.2744   -0.08  0.936
## Year2010        0.1584    0.2017    0.79  0.433
## Year2011        0.2041    0.1681    1.21  0.226
## Year2012        0.2066    0.1699    1.22  0.226
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.421
## Multiple R-squared:  0.0615, Adjusted R-squared:  -0.0277
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 176 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.160  0.842  0.949  0.881  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      5.08e-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: 197"
## [1] ""

```

```

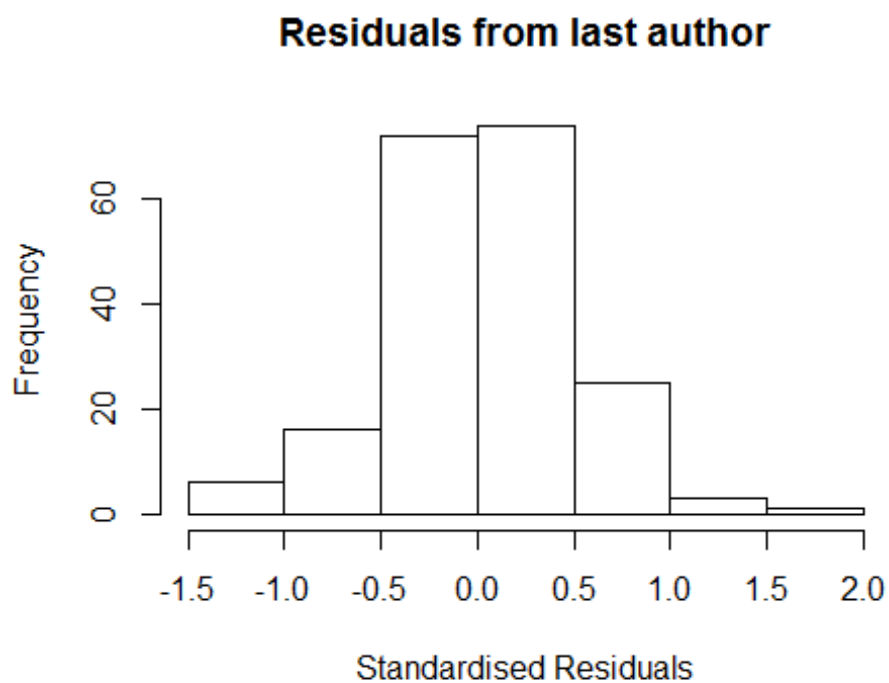
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2716"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    8   15    6   13   17    8    8    7   11   17   14   12    8   13
## 2011 2012
##   10   17
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    4    1    4    0    2    5    6    6    8   12    7    8    6    9
## 2011 2012
##   10   13
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    3    0    3    0    2    5    3    3    8   12    5    8    4    8
## 2011 2012
##    8   12
## [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: 7.45391021762111"

## Warning in wilcox.test.default(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 = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.03658897187566"
## [1] "Male last author team size 2018 geometric mean: 8.11905933503662"

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

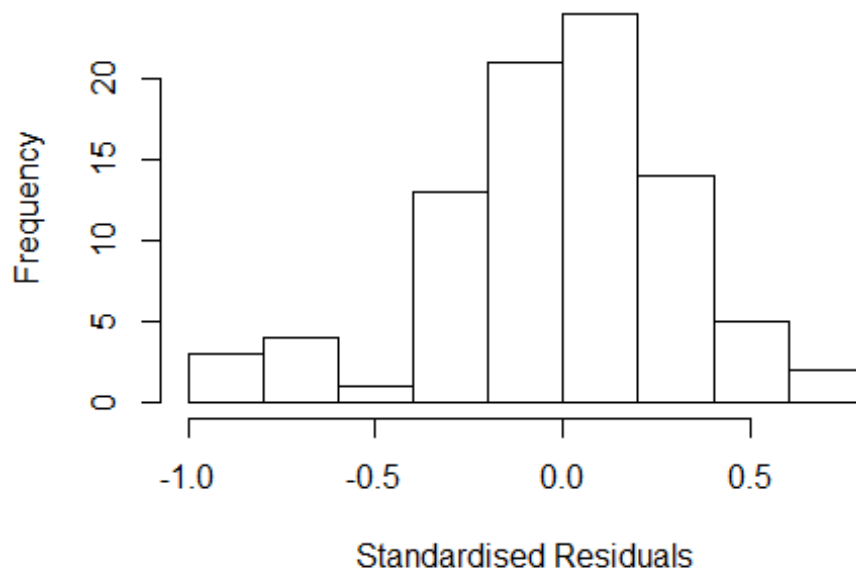
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1, p-value = 0.07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
## Year as factors"
##
```

| | GVIF | Df | $GVIF^{1/(2 \cdot Df)}$ |
|-------------------|---------|----|-------------------------|
| FirstAuthorFemale | 10.37 | 1 | 3.220 |
| LastAuthorFemale | 3.59 | 1 | 1.895 |
| UniqueAuthors | 663.04 | 4 | 2.253 |
| Year | 2594.89 | 14 | 1.324 |

Residuals from first and last author and team size



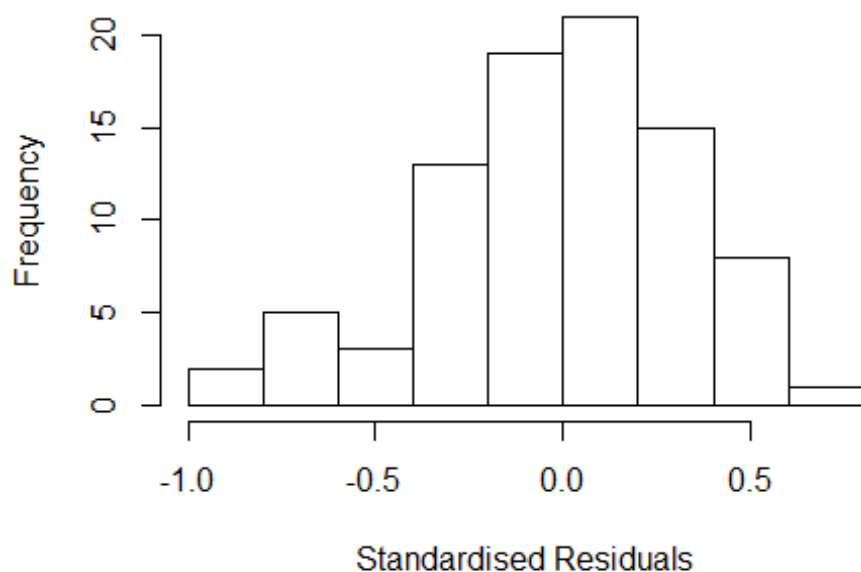
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.88502 -0.17971  0.00334  0.19670  0.69785
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.19397    0.37246   3.21  0.0021 **
## FirstAuthorFemale1 -0.19003    0.13413  -1.42  0.1613
## LastAuthorFemale1 -0.16525    0.09392  -1.76  0.0831 .
## UniqueAuthors2     0.33419    0.26263   1.27  0.2077
## UniqueAuthors3     0.44855    0.30915   1.45  0.1515
## UniqueAuthors4     0.21328    0.35736   0.60  0.5527
## UniqueAuthors5     0.34222    0.49750   0.69  0.4939
## Year1997         -0.40983    0.44404  -0.92  0.3594
## Year1999         -0.08009    0.11302  -0.71  0.4810
## Year2001         -0.56772    0.09265  -6.13  5.6e-08 ***
```

```

## Year2002          0.07879    0.48977    0.16    0.8727
## Year2003         -0.75958    0.43006   -1.77    0.0820 .
## Year2004         -0.30002    0.20789   -1.44    0.1537
## Year2005         -0.34727    0.14767   -2.35    0.0217 *
## Year2006         -0.28786    0.16519   -1.74    0.0861 .
## Year2007         -0.19553    0.21198   -0.92    0.3597
## Year2008         -0.33268    0.19057   -1.75    0.0855 .
## Year2009         -0.00479    0.49317   -0.01    0.9923
## Year2010         -0.31036    0.16881   -1.84    0.0705 .
## Year2011         -0.51327    0.20686   -2.48    0.0156 *
## Year2012         -0.05488    0.20624   -0.27    0.7910
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.291
## Multiple R-squared:  0.359, Adjusted R-squared:  0.165
## Convergence in 43 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 77 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.336  0.863   0.951   0.886   0.986   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.15e-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"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.888 1      1.699
## LastAuthorFemale  1.669 1      1.292
## Year              4.356 14      1.054

```

Residuals from first and last author



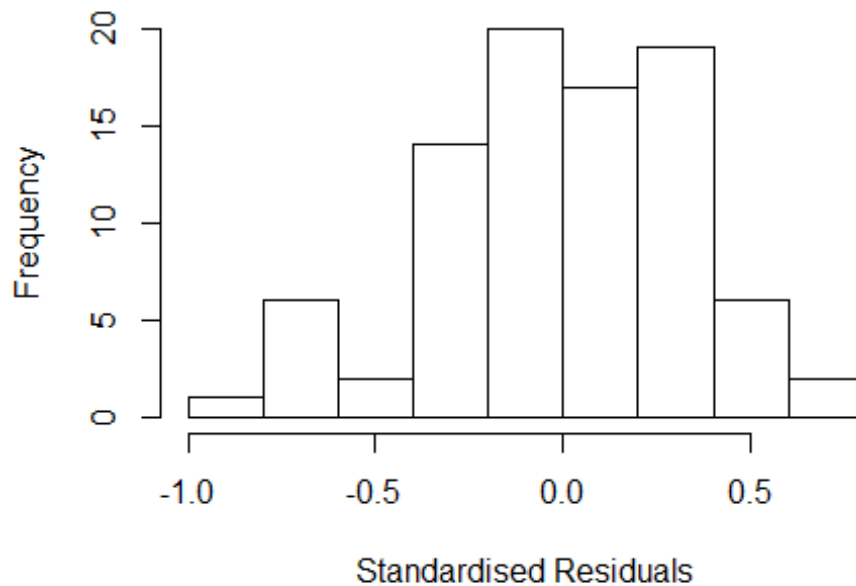
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8921 -0.2102 0.0116 0.2332 0.6064
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5534 0.1094 14.19 < 2e-16 ***
## FirstAuthorFemale1 -0.1615 0.0853 -1.89 0.0624 .
## LastAuthorFemale1 -0.1652 0.0830 -1.99 0.0506 .
## Year1997 -0.4656 0.1973 -2.36 0.0210 *
## Year1999 -0.1259 0.1079 -1.17 0.2475
## Year2001 -0.5461 0.1201 -4.55 2.2e-05 ***
## Year2002 -0.0643 0.3418 -0.19 0.8513
## Year2003 -0.7536 0.3909 -1.93 0.0580 .
## Year2004 -0.3274 0.1827 -1.79 0.0774 .
## Year2005 -0.3979 0.1324 -3.00 0.0037 **
## Year2006 -0.3346 0.1367 -2.45 0.0169 *
## Year2007 -0.2663 0.1615 -1.65 0.1036
```

```

## Year2008          -0.4033      0.1237    -3.26    0.0017 **
## Year2009          -0.0323      0.3451    -0.09    0.9258
## Year2010          -0.3419      0.1319    -2.59    0.0116 *
## Year2011          -0.5863      0.2653    -2.21    0.0304 *
## Year2012          -0.1754      0.1521    -1.15    0.2528
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.328
## Multiple R-squared:  0.256, Adjusted R-squared:  0.0865
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 80 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.440  0.872  0.952  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      1.15e-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"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.545 1          1.595
## Year              2.545 14          1.034

```

Residuals from first author



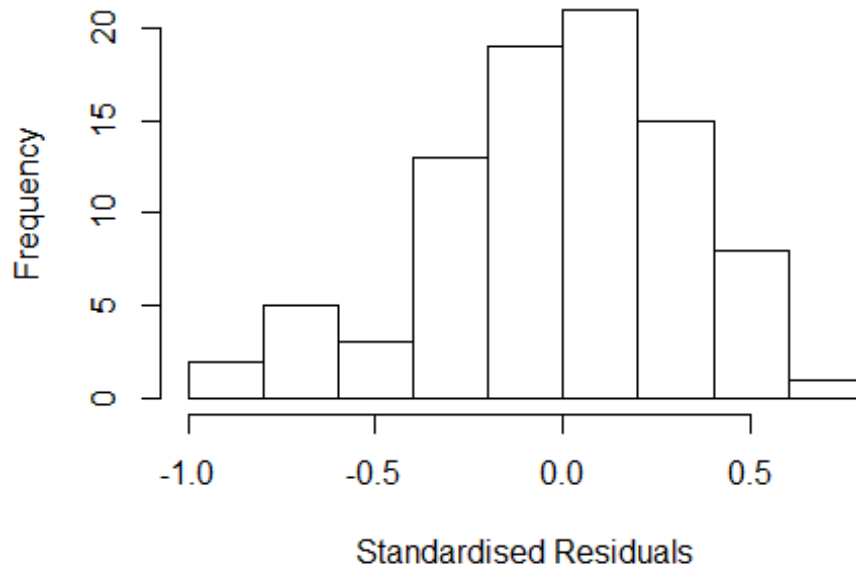
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.96378 -0.21121 0.00263 0.24216 0.62584
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5526 0.1083 14.33 < 2e-16 ***
## FirstAuthorFemale1 -0.1607 0.0827 -1.94 0.0560 .
## Year1997 -0.4635 0.1955 -2.37 0.0205 *
## Year1999 -0.1259 0.1075 -1.17 0.2457
## Year2001 -0.6283 0.0946 -6.64 5.3e-09 ***
## Year2002 -0.0694 0.3199 -0.22 0.8289
## Year2003 -0.7755 0.3810 -2.04 0.0455 *
## Year2004 -0.3263 0.1813 -1.80 0.0762 .
## Year2005 -0.4205 0.1316 -3.20 0.0021 **
## Year2006 -0.4281 0.1425 -3.00 0.0037 **
## Year2007 -0.3047 0.1604 -1.90 0.0616 .
## Year2008 -0.4437 0.1299 -3.41 0.0011 **
```

```

## Year2009          -0.0987      0.2872   -0.34   0.7321
## Year2010          -0.3789      0.1254   -3.02   0.0035 **
## Year2011          -0.6225      0.2336   -2.66   0.0095 **
## Year2012          -0.2257      0.1536   -1.47   0.1461
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.35
## Multiple R-squared:  0.225, Adjusted R-squared:  0.0609
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 81 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.428  0.905  0.954  0.915  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      1.15e-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 1.515 1          1.231
## Year            1.515 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
## -0.9742 -0.2148 0.0232 0.2359 0.5648
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4446 0.1302 11.10 <2e-16 ***
## LastAuthorFemale1 -0.1674 0.0913 -1.83 0.0709 .
## Year1997 -0.4138 0.2361 -1.75 0.0840 .
## Year1999 -0.1786 0.1376 -1.30 0.1985
## Year2001 -0.5169 0.1844 -2.80 0.0065 **
## Year2002 -0.0754 0.2539 -0.30 0.7673
## Year2003 -0.7699 0.3638 -2.12 0.0378 *
## Year2004 -0.2778 0.2152 -1.29 0.2009
## Year2005 -0.3551 0.1630 -2.18 0.0327 *
## Year2006 -0.3030 0.1669 -1.82 0.0737 .
## Year2007 -0.2579 0.2140 -1.20 0.2323
## Year2008 -0.3716 0.1537 -2.42 0.0182 *
```

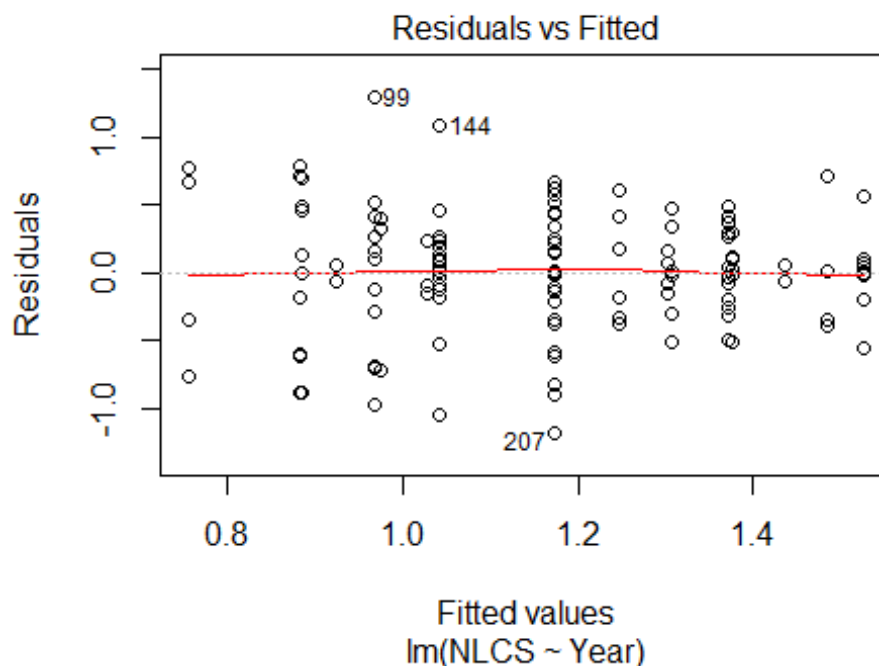
```

## Year2009          -0.0163      0.3796   -0.04    0.9660
## Year2010          -0.3248      0.1615   -2.01    0.0481 *
## Year2011          -0.5552      0.2406   -2.31    0.0240 *
## Year2012          -0.1489      0.1754   -0.85    0.3989
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.375
## Multiple R-squared:  0.216, Adjusted R-squared:  0.0498
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 81 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.480  0.891  0.959  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      1.15e-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 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
##    7    7    5    9    6   11    7    2    7   12    8    6   17   12   28
## 2011 2012
##   29   34
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    4    2    7    4    6    3    2    6    7    7    5   11    8   18
## 2011 2012
##   14   27
##

```



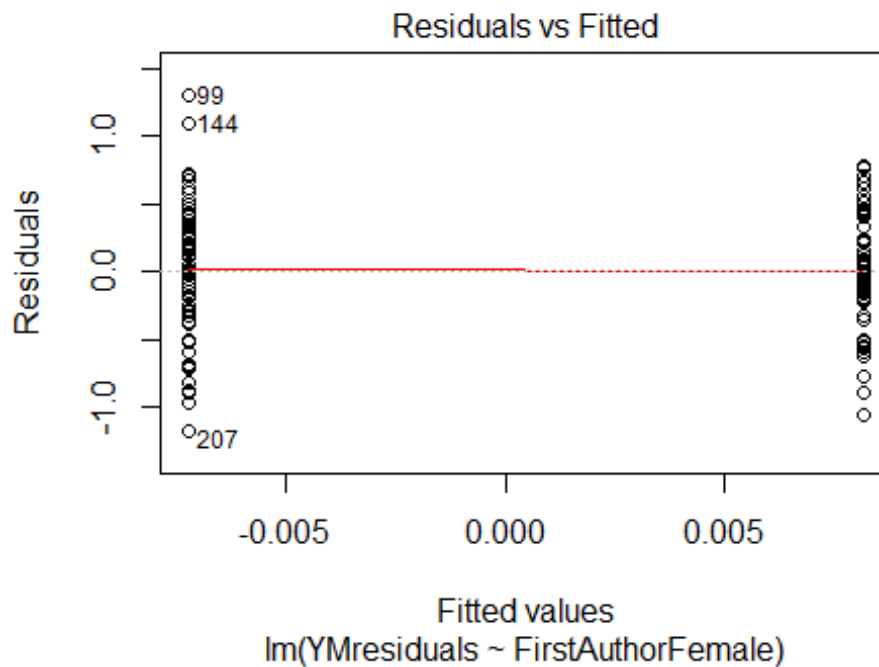
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    3    2    7    4    6    3    2    4    5    7    5   11    7   17
## 2011 2012
##   13   22
## [1] "Heteroscedasticity checks, confirming that there are problems with
##      these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 26, df = 16, p-value = 0.06
```



```
##
## 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: 4.84260121169902"
## [1] "Male first author team size 2018 geometric mean: 4.79603500329864"
##
## Warning in wilcox.test.default(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 = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.80362540373453"
## [1] "Male last author team size 2018 geometric mean: 4.83749854217512"

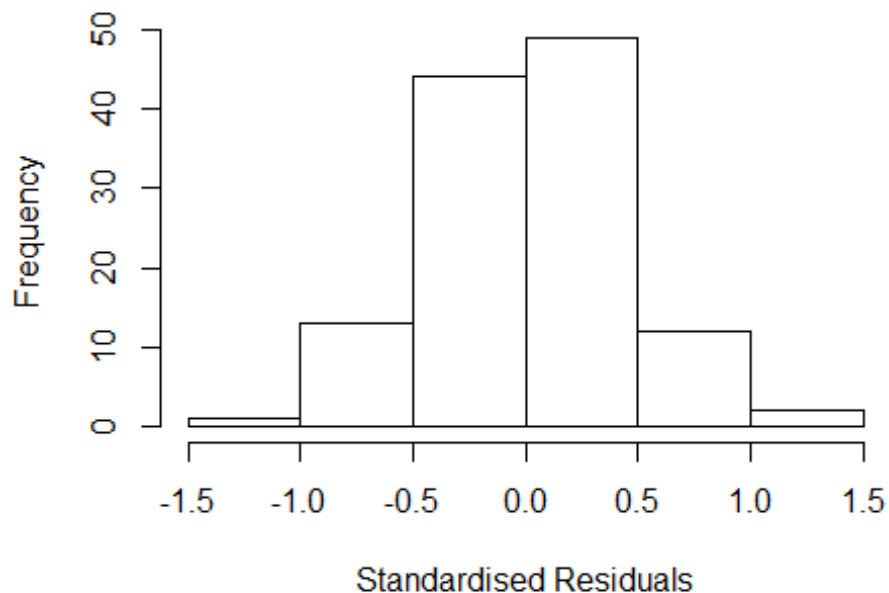
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 52, p-value = 0.8
## 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 | 4.393 | 1 | 2.096 |
| LastAuthorFemale | 2.138 | 1 | 1.462 |
| UniqueAuthors | 246.657 | 4 | 1.991 |
| Year | 552.554 | 16 | 1.218 |

Residuals from first and last author and team size



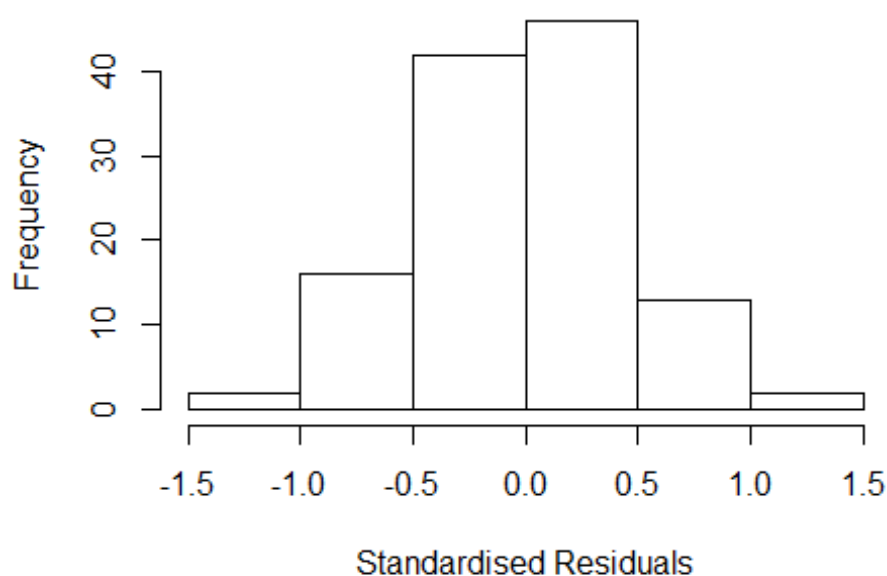
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0699 -0.2430 0.0133 0.2715 1.3760
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5362 0.2328 2.30 0.02339 *
## FirstAuthorFemale1 0.0107 0.1221 0.09 0.93015
## LastAuthorFemale1 0.1036 0.0917 1.13 0.26119
## UniqueAuthors2 0.2019 0.2240 0.90 0.36962
## UniqueAuthors3 0.4242 0.3323 1.28 0.20482
## UniqueAuthors4 0.3266 0.2906 1.12 0.26383
## UniqueAuthors5 0.4891 0.2093 2.34 0.02150 *
## Year1997 0.3497 0.4987 0.70 0.48479
## Year1998 0.0524 0.2102 0.25 0.80366
## Year1999 0.3228 0.2238 1.44 0.15229
```

```

## Year2000          0.4035      0.2217      1.82  0.07174 .
## Year2001          0.3903      0.2092      1.87  0.06507 .
## Year2002          0.2640      0.3838      0.69  0.49314
## Year2003          0.6498      0.2135      3.04  0.00300 **
## Year2004          0.7051      0.1956      3.60  0.00049 ***
## Year2005         -0.1904      0.7952     -0.24  0.81127
## Year2006          0.0713      0.3893      0.18  0.85508
## Year2007         -0.1188      0.3377     -0.35  0.72577
## Year2008          0.1459      0.2084      0.70  0.48560
## Year2009          0.5467      0.1728      3.16  0.00207 **
## Year2010          0.1963      0.1787      1.10  0.27468
## Year2011          0.3095      0.1615      1.92  0.05826 .
## Year2012          0.2799      0.1679      1.67  0.09875 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.392
## Multiple R-squared:  0.289, Adjusted R-squared:  0.13
## Convergence in 29 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 112 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.193  0.864  0.957   0.894   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      8.26e-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 4.018 1      2.004
## LastAuthorFemale  2.304 1      1.518
## Year              9.008 16      1.071

```

Residuals from first and last author



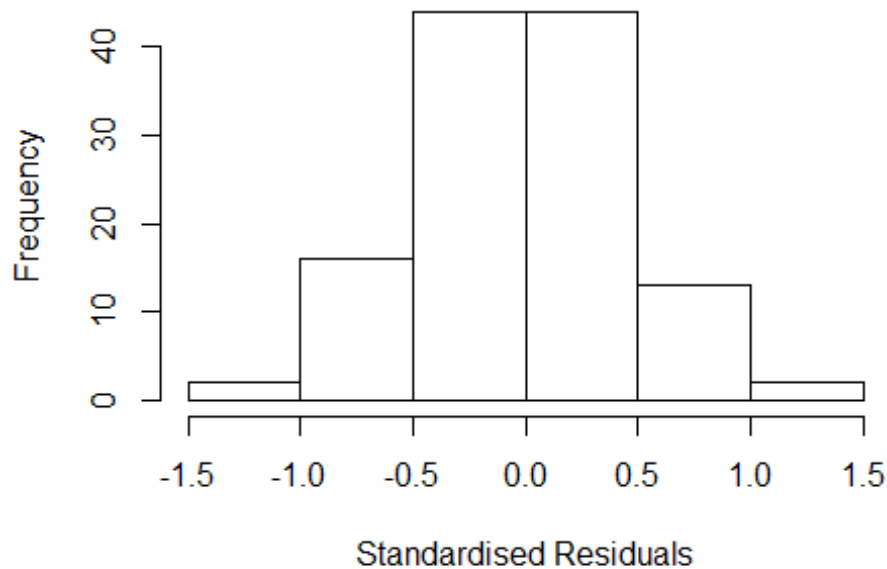
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.236161 -0.270405 0.000522 0.276692 1.366520
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0254 0.1012 10.14 < 2e-16 ***
## FirstAuthorFemale1 0.0252 0.1053 0.24 0.81104
## LastAuthorFemale1 0.0582 0.0974 0.60 0.55145
## Year1997 0.3588 0.4026 0.89 0.37499
## Year1998 -0.1246 0.1527 -0.82 0.41633
## Year1999 0.2034 0.1910 1.06 0.28946
## Year2000 0.2442 0.1445 1.69 0.09400 .
## Year2001 0.2611 0.1827 1.43 0.15613
## Year2002 0.0346 0.3911 0.09 0.92975
## Year2003 0.3980 0.1182 3.37 0.00108 **
## Year2004 0.4588 0.1236 3.71 0.00034 ***
## Year2005 -0.3635 0.4099 -0.89 0.37729
```

```

## Year2006          -0.2034      0.5421   -0.38  0.70828
## Year2007          -0.3933      0.4871   -0.81  0.42128
## Year2008          -0.1319      0.2480   -0.53  0.59588
## Year2009           0.4546      0.1839    2.47  0.01506 *
## Year2010           0.0139      0.1578    0.09  0.92989
## Year2011           0.2750      0.1584    1.74  0.08552 .
## Year2012           0.2108      0.1856    1.14  0.25878
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.394
## Multiple R-squared:  0.226, Adjusted R-squared:  0.0894
## Convergence in 28 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 112 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.204  0.838  0.951  0.885  0.986  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      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 4.723 1      2.173
## Year              4.723 16      1.050

```

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
## -1.26929 -0.25947 -0.00529  0.24481  1.35142
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0253     0.1013   10.12 < 2e-16 ***
## FirstAuthorFemale1  0.0202     0.1104    0.18  0.8552
## Year1997          0.3530     0.4035    0.87  0.3837
## Year1998         -0.1195     0.1565   -0.76  0.4471
## Year1999          0.2021     0.1922    1.05  0.2954
## Year2000          0.2624     0.1476    1.78  0.0783 .
## Year2001          0.2755     0.1795    1.53  0.1280
## Year2002          0.0427     0.3934    0.11  0.9138
## Year2003          0.4006     0.1201    3.34  0.0012 **
## Year2004          0.4731     0.1167    4.06 9.7e-05 ***
## Year2005         -0.3565     0.4342   -0.82  0.4134
## Year2006         -0.2049     0.6346   -0.32  0.7475
```

```

## Year2007          -0.3609      0.5066   -0.71    0.4778
## Year2008          -0.1167      0.2556   -0.46    0.6490
## Year2009           0.4872      0.1662     2.93    0.0041 **
## Year2010           0.0459      0.1423     0.32    0.7476
## Year2011           0.3008      0.1481     2.03    0.0448 *
## Year2012           0.2440      0.1803     1.35    0.1789
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.382
## Multiple R-squared:  0.23,   Adjusted R-squared:  0.103
## Convergence in 33 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 111 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.185  0.818  0.948  0.877  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+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 2.212 1          1.487
## Year              2.212 16          1.025
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.2607 -0.2564 -0.0102  0.2618  1.3566

```



```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0253    0.1013   10.12 < 2e-16 ***
## LastAuthorFemale1 0.0538    0.0970    0.55 0.58022
## Year1997          0.3632    0.3905    0.93 0.35452
## Year1998         -0.0993    0.1105   -0.90 0.37112
## Year1999          0.2021    0.1924    1.05 0.29610
## Year2000          0.2644    0.1167    2.27 0.02558 *
## Year2001          0.2724    0.1871    1.46 0.14848
## Year2002          0.0430    0.3952    0.11 0.91350
## Year2003          0.4107    0.1089    3.77 0.00027 ***
## Year2004          0.4600    0.1234    3.73 0.00032 ***
## Year2005         -0.3399    0.3986   -0.85 0.39572
## Year2006         -0.2039    0.6321   -0.32 0.74771
## Year2007         -0.3743    0.5091   -0.74 0.46385
## Year2008         -0.1219    0.2404   -0.51 0.61336
## Year2009          0.4737    0.1727    2.74 0.00718 **
## Year2010          0.0289    0.1422    0.20 0.83938
## Year2011          0.2903    0.1416    2.05 0.04293 *
## Year2012          0.2355    0.1657    1.42 0.15822
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.382
## Multiple R-squared:  0.23, Adjusted R-squared:  0.103
## Convergence in 31 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~ = 1. The remaining 112 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.180  0.829  0.947  0.878  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      8.26e-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] ""

```

```

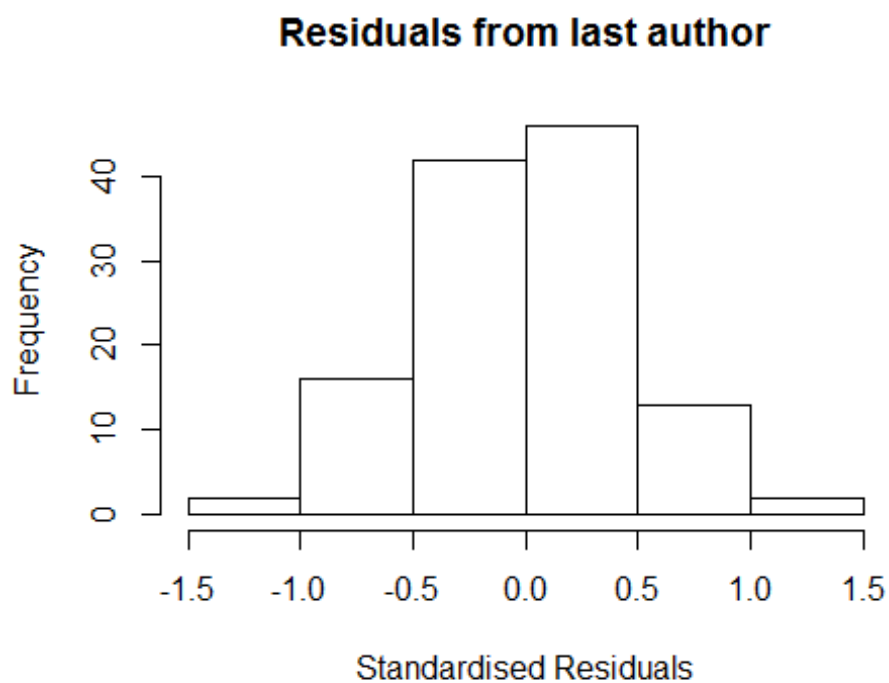
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2718"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    2    6    1    2    2    4    2    4    6    3    8    4    12    12
## 2012
##    10
##
## 1996 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    0    1    0    1    0    1    3    0    3    2    3    7    2    11    9
## 2012
##     7
##
## 1996 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    0    0    0    1    0    1    3    0    3    2    3    4    1    8    7
## 2012
##     6
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.73813722053759"
## [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 = 11, p-value = 0.5
## 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: 4.06860480487927"

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

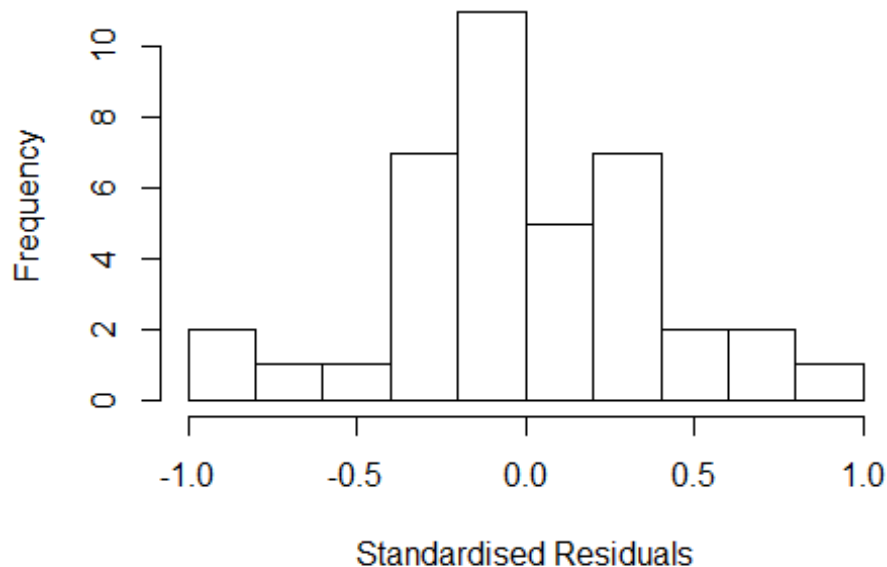
```



```
##
## 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 | -3.039e+14 | 1 | NaN |
| ## LastAuthorFemale | 6.002e+00 | 1 | 2.45 |
| ## UniqueAuthors | -1.119e+32 | 4 | NaN |
| ## Year | -4.074e+32 | 10 | NaN |

Residuals from first and last author and team size



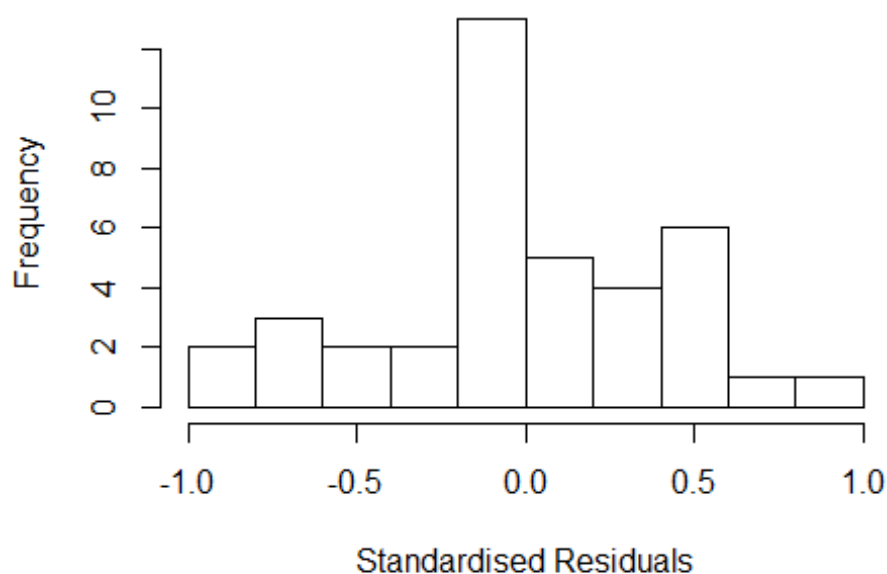
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -8.90e-01 -2.10e-01 -5.55e-16 2.18e-01 8.98e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7016 0.4079 1.72 0.0994 .
## FirstAuthorFemale1 0.1196 0.2446 0.49 0.6298
## LastAuthorFemale1 -0.0469 0.2999 -0.16 0.8771
## UniqueAuthors2 0.3574 0.2289 1.56 0.1328
## UniqueAuthors3 0.5129 0.2012 2.55 0.0183 *
## UniqueAuthors4 -0.2519 0.2138 -1.18 0.2513
## UniqueAuthors5 0.4838 0.2452 1.97 0.0612 .
## Year2002 1.8133 0.4444 4.08 0.0005 ***
## Year2003 0.3240 0.5425 0.60 0.5565
## Year2005 0.3470 0.4384 0.79 0.4371
```

```

## Year2006          0.4414      0.6698      0.66      0.5167
## Year2007         -0.7081      0.3268     -2.17      0.0414 *
## Year2008         -0.3768      0.3645     -1.03      0.3126
## Year2009          0.2295      0.4986      0.46      0.6499
## Year2010         -0.3695      0.4652     -0.79      0.4355
## Year2011          0.2779      0.4027      0.69      0.4974
## Year2012          0.0595      0.4428      0.13      0.8944
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.475
## Multiple R-squared:  0.581, Adjusted R-squared:  0.276
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 31 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.701  0.909  0.969  0.926  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      2.56e-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 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 5.629e+14 1      2.373e+07
## LastAuthorFemale  4.549e+00 1      2.133e+00
## Year              2.811e+15 10      5.922e+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
## -8.76e-01 -1.80e-01 -1.11e-16 3.52e-01 9.16e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.00509 0.20328 4.94 3.9e-05 ***
## FirstAuthorFemale1 0.29991 0.20328 1.48 0.15
## LastAuthorFemale1 -0.00397 0.24832 -0.02 0.99
## Year2002 1.25791 0.20328 6.19 1.5e-06 ***
## Year2003 -0.04476 0.40249 -0.11 0.91
## Year2005 0.00906 0.39879 0.02 0.98
## Year2006 0.62939 0.39752 1.58 0.13
## Year2007 -0.66243 0.14003 -4.73 6.8e-05 ***
## Year2008 -0.36836 0.25285 -1.46 0.16
## Year2009 0.43891 0.20328 2.16 0.04 *
## Year2010 -0.46764 0.32345 -1.45 0.16
## Year2011 0.32851 0.20664 1.59 0.12
```

```

## Year2012          -0.02566    0.27956   -0.09    0.93
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.542
## Multiple R-squared:  0.469, Adjusted R-squared:  0.225
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 34 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.757  0.909  0.961  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           2.56e-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 -1.141e+16  1           NaN
## Year              -1.141e+16 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
## -8.77e-01 -1.81e-01 -8.88e-16  3.51e-01  9.18e-01
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0070     0.1726    5.83 3.3e-06 ***
## FirstAuthorFemale1 0.2980     0.1726    1.73  0.096 .

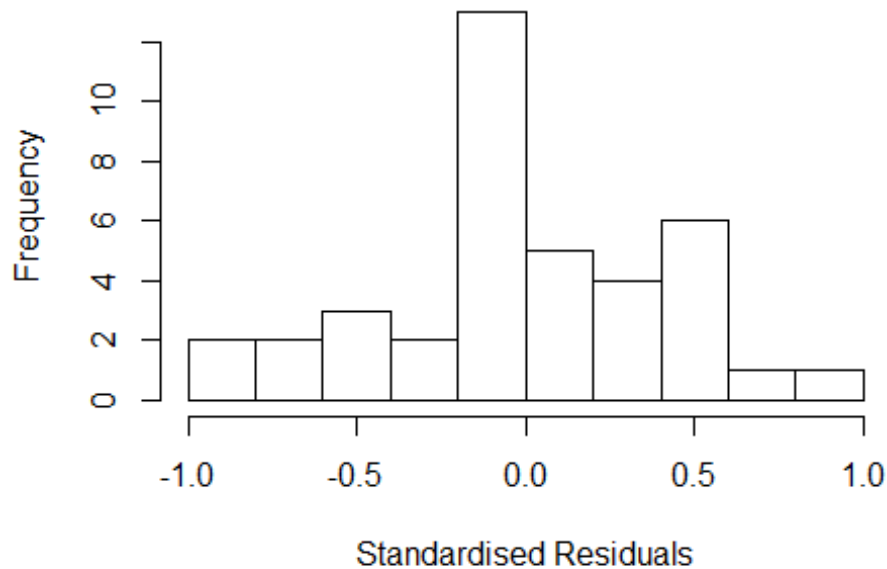
```

```

## Year2002          1.2560      0.1726      7.28  7.9e-08 ***
## Year2003         -0.0457      0.3232     -0.14   0.888
## Year2005          0.0052      0.3837      0.01   0.989
## Year2006          0.6255      0.3211      1.95   0.062 .
## Year2007         -0.6637      0.1202     -5.52  7.5e-06 ***
## Year2008         -0.3705      0.1895     -1.96   0.061 .
## Year2009          0.4370      0.1726      2.53   0.017 *
## Year2010         -0.4697      0.2395     -1.96   0.060 .
## Year2011          0.3258      0.1633      2.00   0.056 .
## Year2012         -0.0290      0.2912     -0.10   0.921
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.561
## Multiple R-squared:  0.467, Adjusted R-squared:  0.25
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  5 weights are ~= 1. The remaining 34 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.771  0.914  0.964  0.940  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      2.56e-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"
## 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

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q    Median        3Q      Max
## -8.64e-01 -2.34e-01  1.11e-16  2.97e-01  1.14e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.305      0.000    Inf < 2e-16 ***
## LastAuthorFemale1  0.160      0.210    0.76  0.45451
## Year2002          0.958      0.000    Inf < 2e-16 ***
## Year2003         -0.286      0.371   -0.77  0.44688
## Year2005         -0.292      0.343   -0.85  0.40193
## Year2006          0.248      0.229    1.08  0.28931
## Year2007         -0.864      0.114   -7.58  3.7e-08 ***
## Year2008         -0.632      0.169   -3.74  0.00087 ***
```

```

## Year2009          0.139      0.000      Inf < 2e-16 ***
## Year2010         -0.690      0.225     -3.06  0.00492 **
## Year2011          0.118      0.190      0.62  0.53820
## Year2012         -0.202      0.277     -0.73  0.47178
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.551
## Multiple R-squared:  0.444, Adjusted R-squared:  0.218
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 33 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.649  0.890  0.970  0.932  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-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      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 2719"
## [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
##    2    1    4    5    3    4    7    7    9   10   14   17   25   25   32
##
## 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    2    4    3    3    4    6    9    7   12   15   19   20   25
##
## 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    2    4    3    1    4    6    6    7   11   12   18   20   23
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"

```

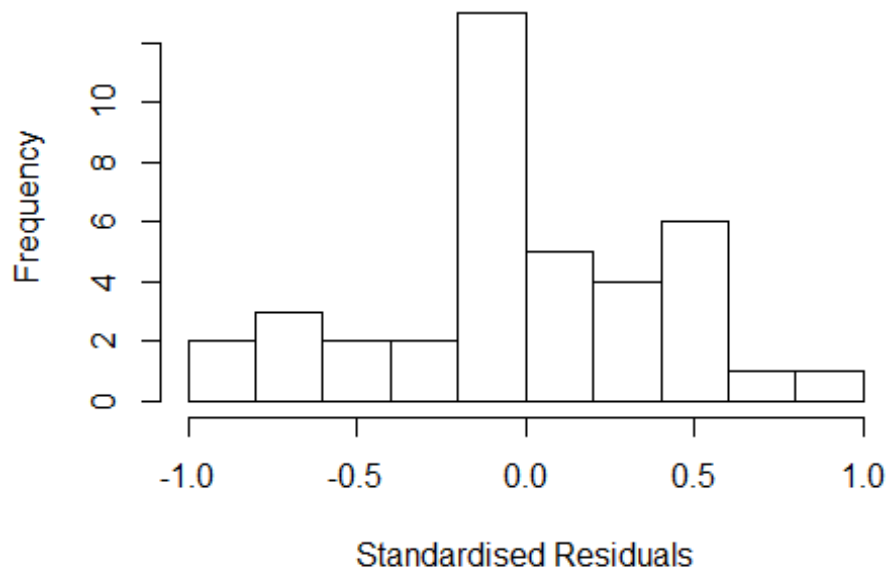
```
## [1] "Female first author team size 2018 geometric mean: 3.59593696387918"
## [1] "Male first author team size 2018 geometric mean: 2.85745885747044"

## Warning in wilcox.test.default(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 = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.00161193882347"
## [1] "Male last author team size 2018 geometric mean: 3.65209917575691"

## 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 = 0.4
## 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  9.327e+12  1      3.054e+06
## LastAuthorFemale   6.299e+00  1      2.510e+00
## UniqueAuthors     -1.309e+16  4      NaN
## Year               -3.078e+29 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 +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3400 -0.2781  0.0655  0.2692  1.1576
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.2405     0.1136   10.92 < 2e-16 ***
## FirstAuthorFemale1  0.2181     0.1214    1.80  0.07536 .
## LastAuthorFemale1 -0.1644     0.1255   -1.31  0.19340
## UniqueAuthors2     0.0115     0.1136    0.10  0.91924
## UniqueAuthors3     0.2086     0.1733    1.20  0.23146
## UniqueAuthors4     0.3024     0.1730    1.75  0.08365 .
## UniqueAuthors5    -0.0980     0.1803   -0.54  0.58799
## Year1999           0.0739     0.1214    0.61  0.54428
## Year2000          -0.0296     0.4874   -0.06  0.95176
## Year2001           0.1176     0.4756    0.25  0.80527
## Year2002          -0.5137     0.1278   -4.02  0.00011 ***
## Year2003          -0.1681     0.1933   -0.87  0.38678
## Year2004          -0.3828     0.2061   -1.86  0.06628 .
## Year2005          -0.1882     0.2224   -0.85  0.39939
## Year2006          -0.4846     0.1803   -2.69  0.00846 **
## Year2007           0.1146     0.2074    0.55  0.58189
## Year2008          -0.2088     0.2140   -0.98  0.33150
## Year2009          -0.2870     0.1546   -1.86  0.06648 .
## Year2010          -0.1244     0.1458   -0.85  0.39547
## Year2011          -0.1852     0.1786   -1.04  0.30235
## Year2012          -0.1091     0.2095   -0.52  0.60384
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.465
## Multiple R-squared:  0.138, Adjusted R-squared:  -0.0382

```

```

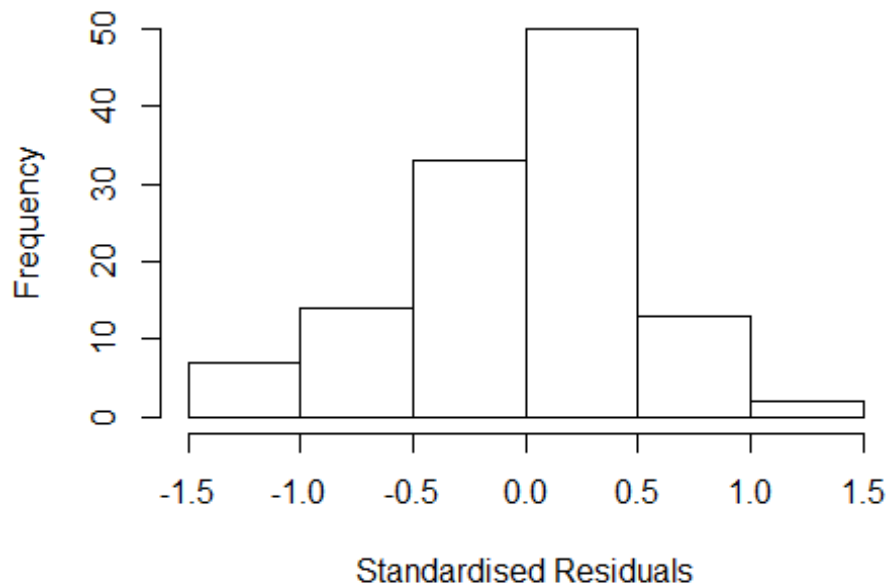
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 109 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.387  0.857  0.966   0.900   0.991   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      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      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"

## 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

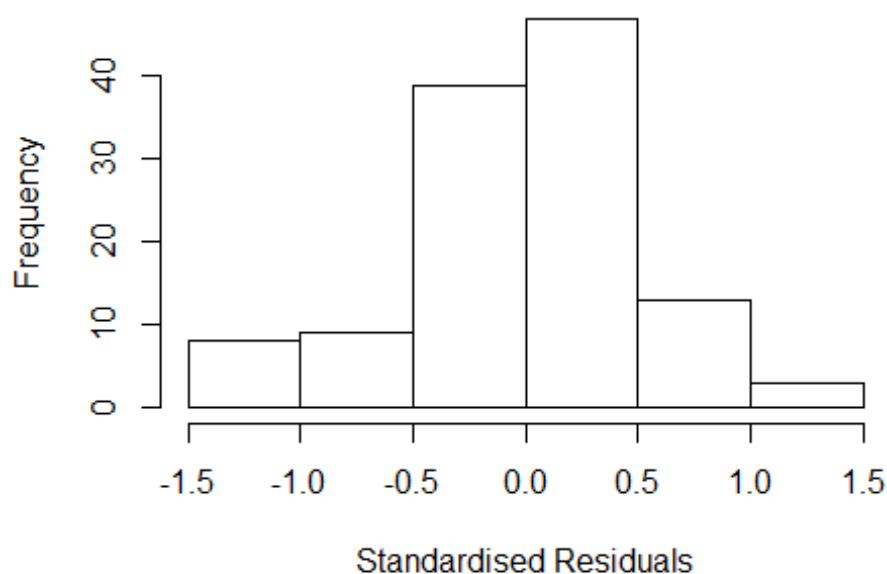
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2336 -0.3138  0.0292  0.2876  1.1628
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2520     0.0000      Inf <2e-16 ***
## FirstAuthorFemale1  0.1562     0.1403      1.11  0.268
## LastAuthorFemale1 -0.1106     0.1281     -0.86  0.390
## Year1999           0.1358     0.1403      0.97  0.335
## Year2000           0.0363     0.3654      0.10  0.921
## Year2001           0.1175     0.4874      0.24  0.810
## Year2002          -0.4310     0.0828     -5.20 1e-06 ***
```

```

## Year2003          0.0290      0.0000      Inf    <2e-16 ***
## Year2004         -0.3001      0.1537     -1.95     0.054 .
## Year2005         -0.0928      0.2135     -0.43     0.665
## Year2006         -0.3808      0.1562     -2.44     0.016 *
## Year2007          0.1623      0.1632      0.99     0.322
## Year2008         -0.1803      0.1870     -0.96     0.337
## Year2009         -0.2429      0.1518     -1.60     0.112
## Year2010         -0.0640      0.1263     -0.51     0.613
## Year2011         -0.1393      0.1876     -0.74     0.459
## Year2012         -0.0970      0.1826     -0.53     0.597
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.463
## Multiple R-squared:  0.0893, Adjusted R-squared:  -0.0535
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
##  15 weights are ~ = 1. The remaining 104 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.458  0.828  0.952  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      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"
## 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

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2792 -0.3033  0.0177  0.2641  1.2278
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2520     0.0000      Inf < 2e-16 ***
## FirstAuthorFemale1 0.1064     0.1202     0.89  0.3780
## Year1999          0.1856     0.1202     1.54  0.1255
## Year2000         -0.0190     0.4342    -0.04  0.9652
## Year2001          0.0953     0.4724     0.20  0.8404
## Year2002         -0.4147     0.0912    -4.55 1.5e-05 ***
## Year2003          0.0290     0.0000      Inf < 2e-16 ***
## Year2004         -0.2777     0.1536    -1.81  0.0736 .
```



```

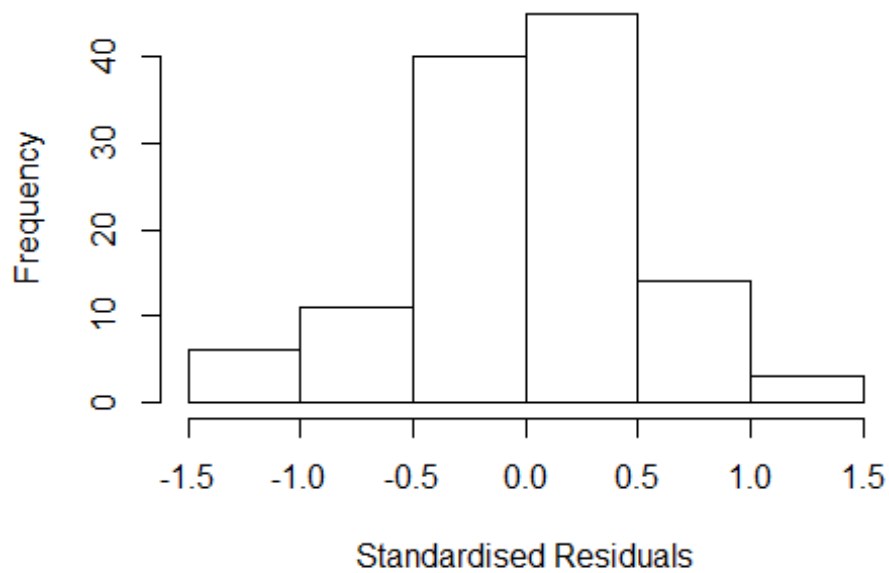
## Year2005          -0.1161      0.2068   -0.56   0.5758
## Year2006          -0.4101      0.1501   -2.73   0.0074 **
## Year2007           0.1172      0.1609    0.73   0.4681
## Year2008          -0.2091      0.1819   -1.15   0.2529
## Year2009          -0.2758      0.1454   -1.90   0.0606 .
## Year2010          -0.0792      0.1279   -0.62   0.5372
## Year2011          -0.1729      0.1768   -0.98   0.3304
## Year2012          -0.1330      0.1744   -0.76   0.4473
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.458
## Multiple R-squared:  0.0816, Adjusted R-squared:  -0.0522
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 107 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.416  0.843   0.950   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      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"

## 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

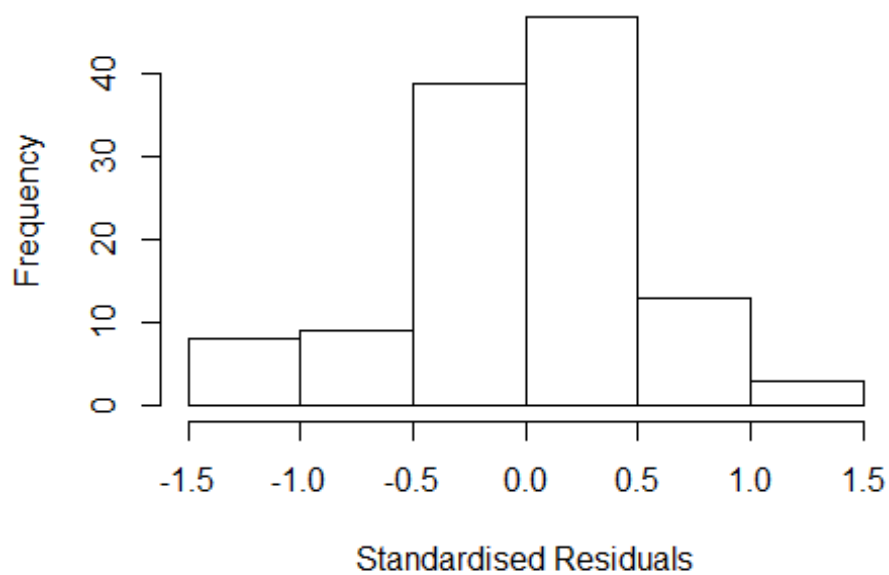
```

Residuals from first author



```
##          GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale  NaN  1          NaN
## Year              NaN 14          NaN
```

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
## -1.214 -0.289 0.026 0.262 1.272
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.25200 0.00000 Inf <2e-16 ***
## LastAuthorFemale1 -0.03976 0.10956 -0.36 0.7174
## Year1999 0.29200 0.00000 Inf <2e-16 ***
## Year2000 0.00088 0.40223 0.00 0.9983
## Year2001 0.14284 0.49118 0.29 0.7718
## Year2002 -0.38026 0.11413 -3.33 0.0012 **
## Year2003 0.02900 0.00000 Inf <2e-16 ***
## Year2004 -0.22866 0.15814 -1.45 0.1512
## Year2005 -0.04738 0.20479 -0.23 0.8175
## Year2006 -0.33900 0.16140 -2.10 0.0381 *
## Year2007 0.22607 0.13943 1.62 0.1080
## Year2008 -0.14222 0.17958 -0.79 0.4302
## Year2009 -0.20354 0.14187 -1.43 0.1544
## Year2010 -0.01699 0.11922 -0.14 0.8869
## Year2011 -0.07261 0.16316 -0.45 0.6572
## Year2012 -0.03805 0.16230 -0.23 0.8151
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.479
## Multiple R-squared: 0.0723, Adjusted R-squared: -0.0628
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 105 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.461 0.867 0.959 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 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
```

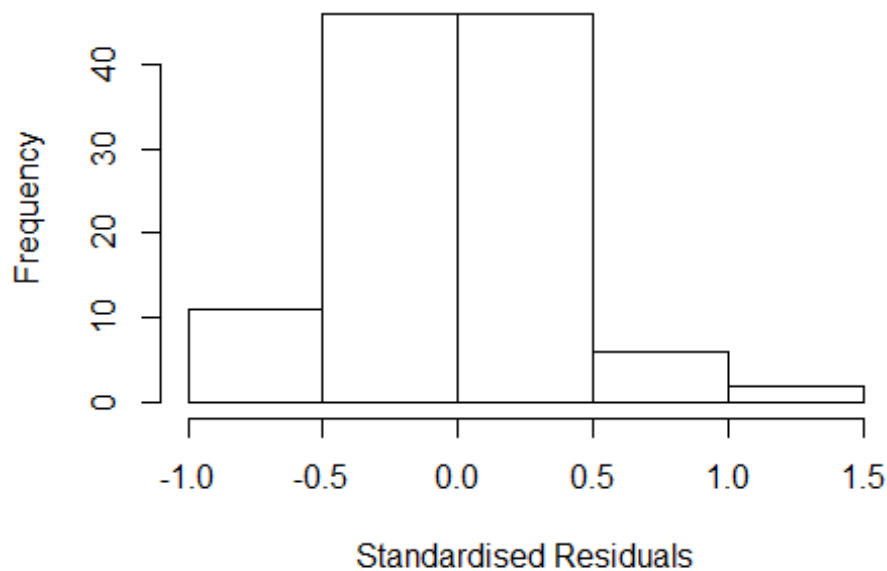
```

##           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 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
##   13    5   13    7   23   10   19   10   14   10   16    7   14   17   17
## 2011 2012
##   19   19
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    7    4   16    6   11    5    9    6    7    4   10   10   12
## 2011 2012
##   14    9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    5    4   15    4   11    3    8    3    6    4    6    9    9
## 2011 2012
##   14    8
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 7.09397094475071"
## [1] "Male first author team size 2018 geometric mean: 5"
##
## 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: 17"
## [1] "Male last author team size 2018 geometric mean: 4.71769398031653"
##
## 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] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 7.649e+12  1      2.766e+06
## LastAuthorFemale  2.755e+00  1      1.660e+00
## UniqueAuthors    1.275e+14  4      5.797e+01
## Year              7.066e+14 16      2.911e+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
## -9.24e-01 -2.65e-01 -7.77e-16  2.91e-01  1.43e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.0318      0.2582   0.12   0.9024
```

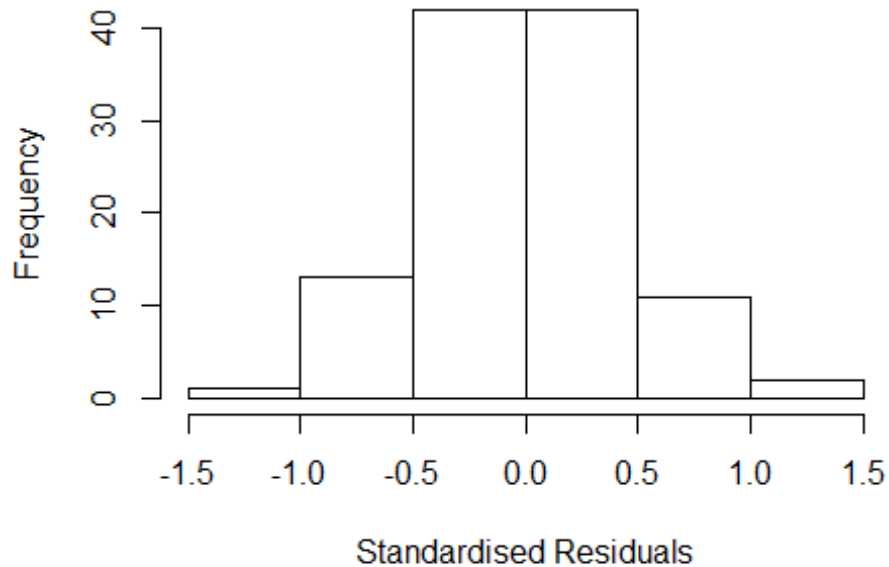
```

## FirstAuthorFemale1  0.0839    0.0948    0.89    0.3785
## LastAuthorFemale1  -0.0107    0.0822   -0.13    0.8965
## UniqueAuthors2     0.5734    0.2452    2.34    0.0216 *
## UniqueAuthors3     0.6764    0.2407    2.81    0.0061 **
## UniqueAuthors4     0.7392    0.2565    2.88    0.0050 **
## UniqueAuthors5     0.5591    0.2192    2.55    0.0125 *
## Year1997            0.8432    0.1759    4.79    6.6e-06 ***
## Year1998            0.4990    0.2874    1.74    0.0861 .
## Year1999            0.1761    0.1566    1.12    0.2638
## Year2000           -0.1024    0.2437   -0.42    0.6753
## Year2001            0.1646    0.2106    0.78    0.4365
## Year2002            0.4435    0.1691    2.62    0.0103 *
## Year2003            0.0124    0.2015    0.06    0.9509
## Year2004            0.4042    0.2094    1.93    0.0568 .
## Year2005            0.1115    0.1310    0.85    0.3971
## Year2006            0.3389    0.1775    1.91    0.0595 .
## Year2007            0.5046    0.3276    1.54    0.1271
## Year2008            0.2984    0.1566    1.91    0.0600 .
## Year2009            0.1493    0.1613    0.93    0.3573
## Year2010           -0.0604    0.1863   -0.32    0.7465
## Year2011            0.1234    0.1631    0.76    0.4514
## Year2012            0.0694    0.2090    0.33    0.7406
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.417
## Multiple R-squared:  0.295, Adjusted R-squared:  0.119
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 98 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.214  0.887   0.949   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           9.01e-04           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01                5.00e-01
## nResample      max.it    best.r.s    k.fast.s      k.max maxit.scale
##           500          50         2         1         1000         200
## trace.lev      mts    compute.rd
##           0          1000         0
##           psi                subsampling                cov
##           "bisquare"          "nonsingular"                ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as

```

```
factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.553e+14  1      1.246e+07
## LastAuthorFemale  2.524e+00  1      1.589e+00
## Year              3.799e+14 16      2.855e+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
## -1.17e+00 -2.92e-01  2.05e-15  3.00e-01  1.33e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.6659    0.1029   6.47 4.6e-09 ***
## FirstAuthorFemale1  0.1261    0.1029   1.23 0.22334
## LastAuthorFemale1 -0.0149    0.0870  -0.17 0.86467
## Year1997         0.7681    0.1029   7.47 4.6e-11 ***
## Year1998         0.5002    0.4325   1.16 0.25043
## Year1999         0.1598    0.1507   1.06 0.29149
```

```

## Year2000          -0.2090      0.2305    -0.91  0.36709
## Year2001          0.1255      0.1515     0.83  0.40967
## Year2002          0.3831      0.1101     3.48  0.00077 ***
## Year2003          0.0322      0.1893     0.17  0.86540
## Year2004          0.3267      0.1618     2.02  0.04643 *
## Year2005          0.1222      0.1104     1.11  0.27150
## Year2006          0.2914      0.1270     2.29  0.02407 *
## Year2007          0.4573      0.3309     1.38  0.17030
## Year2008          0.1388      0.1151     1.21  0.23116
## Year2009          0.0702      0.1349     0.52  0.60416
## Year2010         -0.1433      0.1639    -0.87  0.38433
## Year2011          0.0674      0.1419     0.47  0.63610
## Year2012          0.0577      0.1910     0.30  0.76328
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.441
## Multiple R-squared:  0.199, Adjusted R-squared:  0.0427
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 100 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.344  0.879  0.953  0.912  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          9.01e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.953e+13 1      5.434e+06
## Year              2.953e+13 16      2.636e+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:

```



```

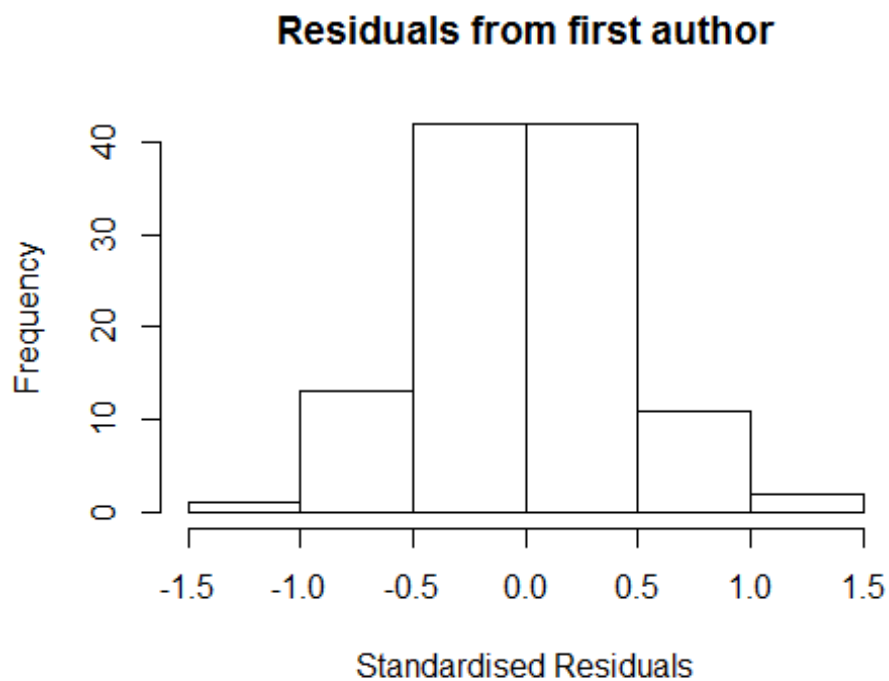
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q      Median        3Q        Max
## -1.17e+00 -3.00e-01  7.77e-16  3.04e-01  1.34e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.6647      0.1023   6.50 4.0e-09 ***
## FirstAuthorFemale1 0.1273      0.1023   1.24 0.21651
## Year1997          0.7693      0.1023   7.52 3.4e-11 ***
## Year1998          0.5015      0.4285   1.17 0.24492
## Year1999          0.1571      0.1492   1.05 0.29493
## Year2000         -0.2152      0.2320  -0.93 0.35599
## Year2001          0.1187      0.1513   0.78 0.43452
## Year2002          0.3809      0.1096   3.48 0.00078 ***
## Year2003          0.0272      0.1908   0.14 0.88716
## Year2004          0.3241      0.1616   2.00 0.04790 *
## Year2005          0.1230      0.1102   1.12 0.26728
## Year2006          0.2838      0.1174   2.42 0.01759 *
## Year2007          0.4489      0.3208   1.40 0.16501
## Year2008          0.1370      0.1144   1.20 0.23392
## Year2009          0.0652      0.1253   0.52 0.60425
## Year2010         -0.1475      0.1616  -0.91 0.36364
## Year2011          0.0631      0.1407   0.45 0.65508
## Year2012          0.0522      0.1824   0.29 0.77550
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.435
## Multiple R-squared:  0.201, Adjusted R-squared:  0.0553
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 101 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.325  0.876  0.952  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          9.01e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500            50          2            1            1000      200
##      trace.lev      mts      compute.rd
##      0              1000          0

```

```
##               psi               subsampling               cov
##           "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"

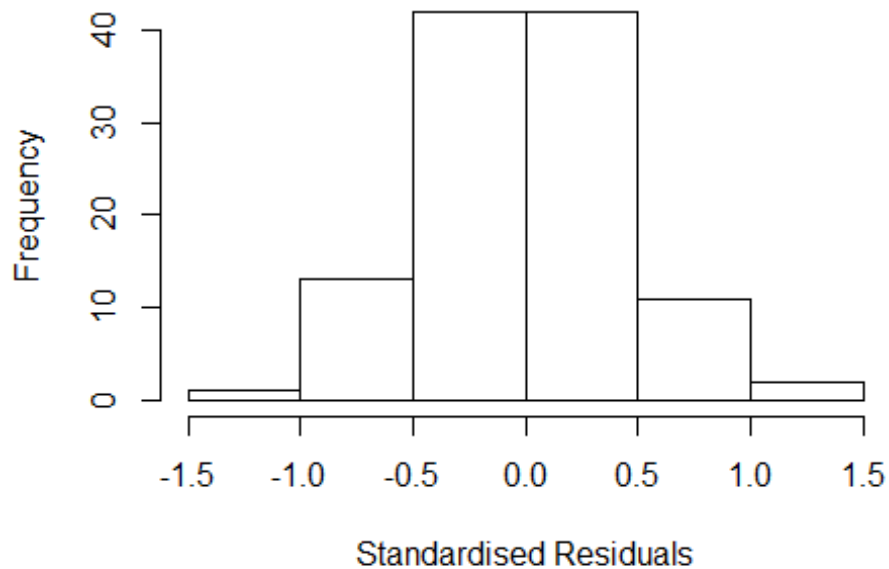
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

## 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



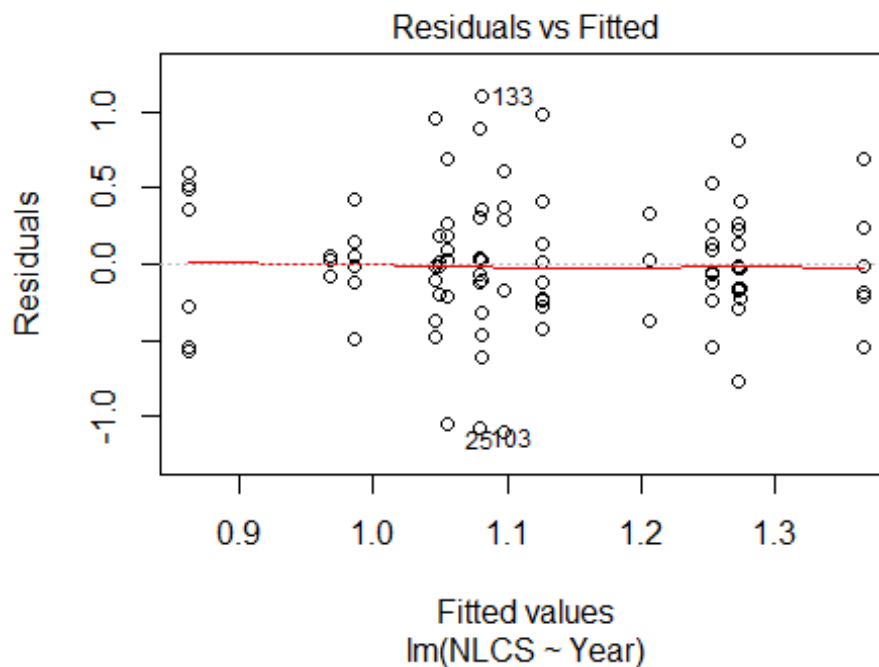
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2578 -0.2603 -0.0113 0.3361 1.2385
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.79200 0.00000 Inf <2e-16 ***
## LastAuthorFemale1 -0.02187 0.08964 -0.24 0.8078
## Year1997 0.64200 0.00000 Inf <2e-16 ***
## Year1998 0.46583 0.50860 0.92 0.3621
## Year1999 0.06849 0.15189 0.45 0.6531
## Year2000 -0.24448 0.22590 -1.08 0.2819
## Year2001 0.06652 0.17343 0.38 0.7022
## Year2002 0.32100 0.10142 3.16 0.0021 **
## Year2003 -0.00285 0.14991 -0.02 0.9849
## Year2004 0.22611 0.12520 1.81 0.0741 .
## Year2005 0.03941 0.10139 0.39 0.6984
## Year2006 0.25103 0.12711 1.97 0.0513 .
```

```

## Year2007      0.43166      0.37309      1.16      0.2502
## Year2008      0.05476      0.11558      0.47      0.6368
## Year2009      0.02120      0.14879      0.14      0.8870
## Year2010     -0.17444      0.16524     -1.06      0.2939
## Year2011     -0.00339      0.11449     -0.03      0.9765
## Year2012      0.04628      0.18531      0.25      0.8034
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.437
## Multiple R-squared:  0.186, Adjusted R-squared:  0.0368
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 98 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.387  0.871  0.942  0.907  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      9.01e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 111"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2721"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    4    9    8   14   12    6    6   11    5    5    9    2    9   10
## 2011 2012
##   13   12
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    3    6    5    9   10    4    6    7    4    3    7    0    5    7
## 2011 2012

```

```
## 10 8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 0 2 5 5 6 9 3 6 3 2 3 5 0 2 5
## 2011 2012
## 10 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 = 17, df = 14, p-value = 0.2
```



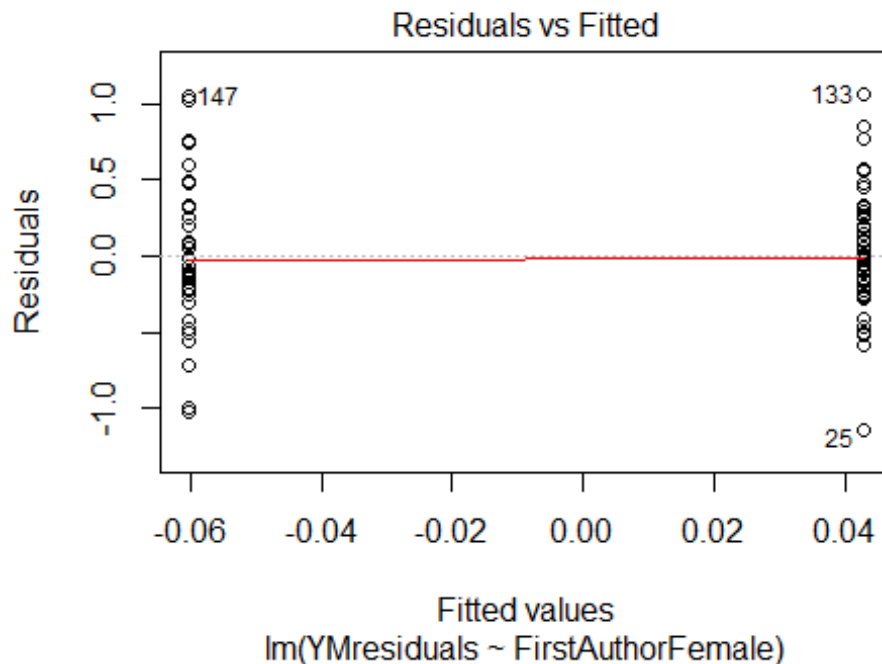
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3, df = 1, p-value = 0.08

## [1] "Female first author team size 2018 geometric mean: 6.92820323027551"
## [1] "Male first author team size 2018 geometric mean: 10.4400868170485"

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

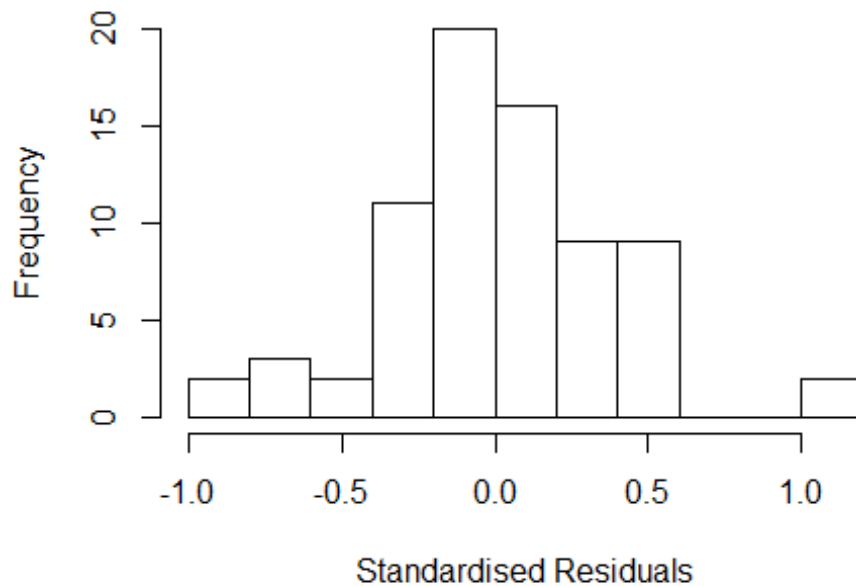
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2.5, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6"
## [1] "Male last author team size 2018 geometric mean: 9.89877118707762"

## Warning in wilcox.test.default(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] "Regression 1: First author gender, last author gender, team size,
##      Year as factors"
##
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale    6.201  1      2.490
## LastAuthorFemale     5.835  1      2.416
## UniqueAuthors      390.059  4      2.108
## Year                1280.589 14      1.291
```

Residuals from first and last author and team size



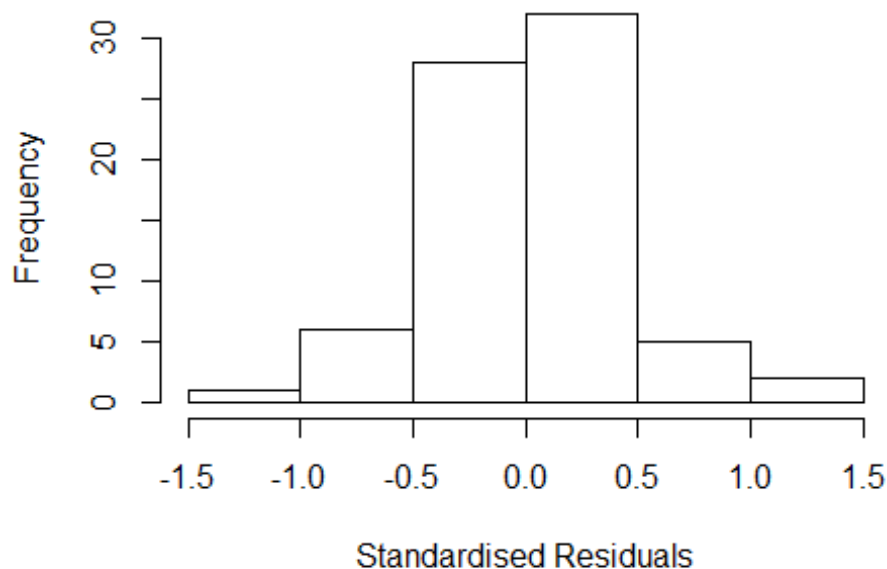
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.89883 -0.19926 -0.00286 0.21401 1.16294
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.05025 0.50404 2.08 0.0420 *
## FirstAuthorFemale1 -0.08325 0.14921 -0.56 0.5792
## LastAuthorFemale1 -0.25466 0.15505 -1.64 0.1064
## UniqueAuthors2 0.06952 0.46209 0.15 0.8810
## UniqueAuthors3 -0.44995 0.58026 -0.78 0.4415
## UniqueAuthors4 -0.24589 0.44867 -0.55 0.5860
## UniqueAuthors5 -0.07277 0.46686 -0.16 0.8767
## Year1998 0.37403 0.12815 2.92 0.0051 **
## Year1999 0.29853 0.27113 1.10 0.2758
## Year2000 0.37156 0.16101 2.31 0.0249 *
```

```

## Year2001          0.43729      0.13961      3.13      0.0028 **
## Year2002          0.36307      0.21967      1.65      0.1043
## Year2003          0.20908      0.16677      1.25      0.2154
## Year2004          0.56725      0.24446      2.32      0.0242 *
## Year2005          0.20973      0.19232      1.09      0.2804
## Year2006          0.41002      0.18551      2.21      0.0314 *
## Year2007         -0.00879      0.31146     -0.03      0.9776
## Year2009          0.41486      0.37643      1.10      0.2754
## Year2010          0.04358      0.39475      0.11      0.9125
## Year2011          0.15503      0.15997      0.97      0.3369
## Year2012          0.00229      0.27679      0.01      0.9934
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.338
## Multiple R-squared:  0.303, Adjusted R-squared:  0.0399
## Convergence in 29 IRWLS iterations
##
## Robustness weights:
## 2 weights are ~= 1. The remaining 72 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.213  0.874  0.962  0.899  0.987  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.35e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.449 1      1.857
## LastAuthorFemale  2.751 1      1.659
## Year              9.025 14      1.082

```


Residuals from first and last author



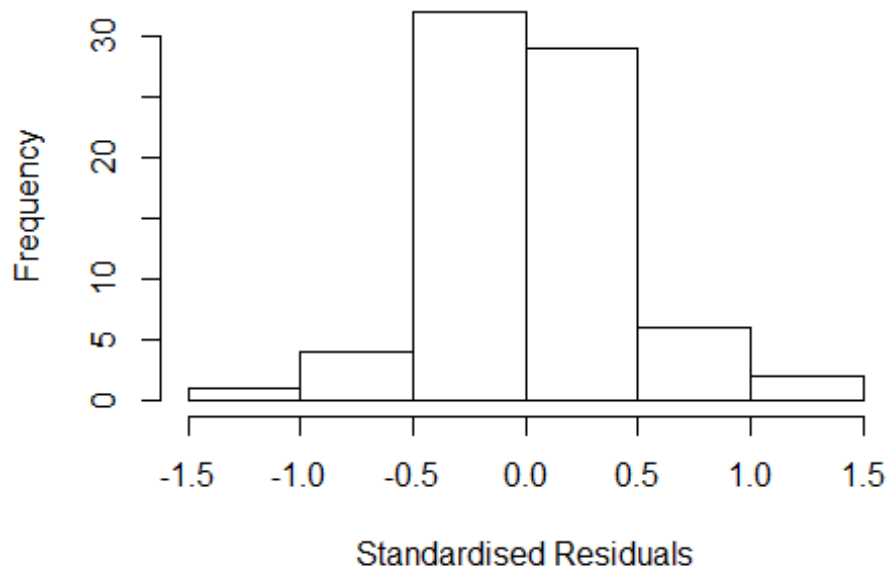
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.38801 -0.20486  0.00524  0.21093  1.28093
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.07005    0.06995   15.30  <2e-16 ***
## FirstAuthorFemale1 -0.12610    0.12293   -1.03   0.3093
## LastAuthorFemale1 -0.17195    0.11048   -1.56   0.1252
## Year1998         0.24385    0.11310    2.16   0.0353 *
## Year1999         0.31796    0.18238    1.74   0.0867 .
## Year2000         0.31006    0.17254    1.80   0.0776 .
## Year2001         0.32573    0.10291    3.17   0.0025 **
## Year2002         0.28022    0.22662    1.24   0.2213
## Year2003         0.00648    0.12746    0.05   0.9596
## Year2004         0.45443    0.26561    1.71   0.0925 .
## Year2005        -0.05000    0.04617   -1.08   0.2834
## Year2006         0.26204    0.24436    1.07   0.2881
```

```

## Year2007          -0.12992    0.19219   -0.68    0.5018
## Year2009          -0.13760    0.19842   -0.69    0.4908
## Year2010          -0.16698    0.32238   -0.52    0.6065
## Year2011           0.03474    0.13165    0.26    0.7928
## Year2012          -0.14941    0.25273   -0.59    0.5567
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.335
## Multiple R-squared:  0.275, Adjusted R-squared:  0.0712
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0471 0.8350 0.9440 0.8790 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      1.35e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.667 1          1.633
## Year              2.667 14          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
## -1.372534 -0.207167  0.000809  0.242135  1.248087
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0735     0.0684   15.70  <2e-16 ***
## FirstAuthorFemale1 -0.1330     0.1164   -1.14   0.2581
## Year1998          0.2050     0.1052    1.95   0.0561 .
## Year1999          0.2990     0.1937    1.54   0.1281
## Year2000          0.3069     0.1714    1.79   0.0787 .
## Year2001          0.3043     0.1038    2.93   0.0048 **
## Year2002          0.2804     0.2275    1.23   0.2228
## Year2003         -0.0276     0.1190   -0.23   0.8172
## Year2004          0.3364     0.3222    1.04   0.3008
## Year2005         -0.0500     0.0464   -1.08   0.2854
## Year2006          0.1897     0.1987    0.95   0.3436
## Year2007         -0.1412     0.1879   -0.75   0.4555
```

```

## Year2009          -0.3130      0.1543   -2.03   0.0471 *
## Year2010          -0.1376      0.3421   -0.40   0.6890
## Year2011           0.0434      0.1391    0.31   0.7559
## Year2012          -0.1516      0.2382   -0.64   0.5268
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.351
## Multiple R-squared:  0.25,   Adjusted R-squared:  0.0556
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0909 0.8490 0.9480 0.8860 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      1.35e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 3.221 1          1.795
## Year              3.221 14          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
## -1.3646 -0.1894  0.0136  0.1787  1.3935
##
## Coefficients:

```

```

##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0070      0.0113  88.97 < 2e-16 ***
## LastAuthorFemale1 -0.1811      0.1097  -1.65  0.10400
## Year1998          0.2624      0.1327   1.98  0.05266 .
## Year1999          0.3576      0.2130   1.68  0.09860 .
## Year2000          0.3349      0.1375   2.44  0.01793 *
## Year2001          0.3439      0.0965   3.56  0.00074 ***
## Year2002          0.3113      0.1724   1.81  0.07619 .
## Year2003          0.0209      0.1272   0.16  0.87017
## Year2004          0.4272      0.2140   2.00  0.05059 .
## Year2005         -0.0500      0.0787  -0.64  0.52783
## Year2006          0.2784      0.2284   1.22  0.22781
## Year2007         -0.0477      0.1460  -0.33  0.74527
## Year2009         -0.0654      0.1777  -0.37  0.71438
## Year2010         -0.2165      0.2867  -0.76  0.45313
## Year2011          0.0337      0.1148   0.29  0.77036
## Year2012         -0.1364      0.3253  -0.42  0.67652
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.329
## Multiple R-squared:  0.264, Adjusted R-squared:  0.074
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0335 0.8410 0.9590 0.8720 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           1.35e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01           5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500           50           2           1           1000           200
## trace.lev      mts      compute.rd
##           0           1000           0
##           psi           subsampling           cov
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 74"
## [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
##    4    7    1    4    1    3    6   10    8    6    6    8    9    9   10
## 2011 2012
##    5    4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    0    1    1    0    2    4    3    3    3    5    4    8    8
## 2011 2012
##    3    2
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    0    1    0    0    0    3    2    2    3    4    3    7    7
## 2011 2012
##    2    2
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.88449914061482"
## [1] "Male first 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 = 1, p-value = 1
## 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: 2"

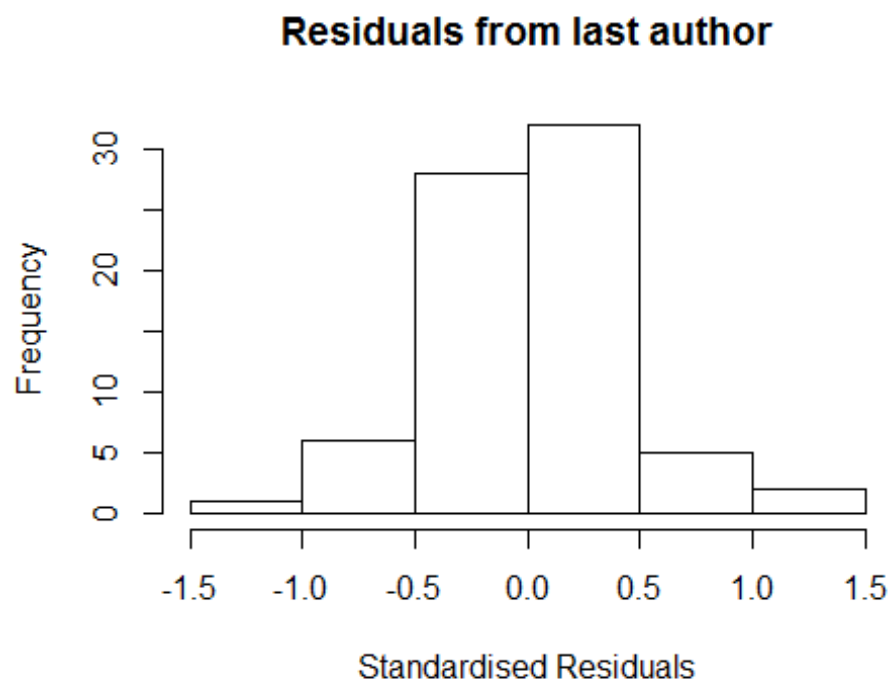
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## 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] "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

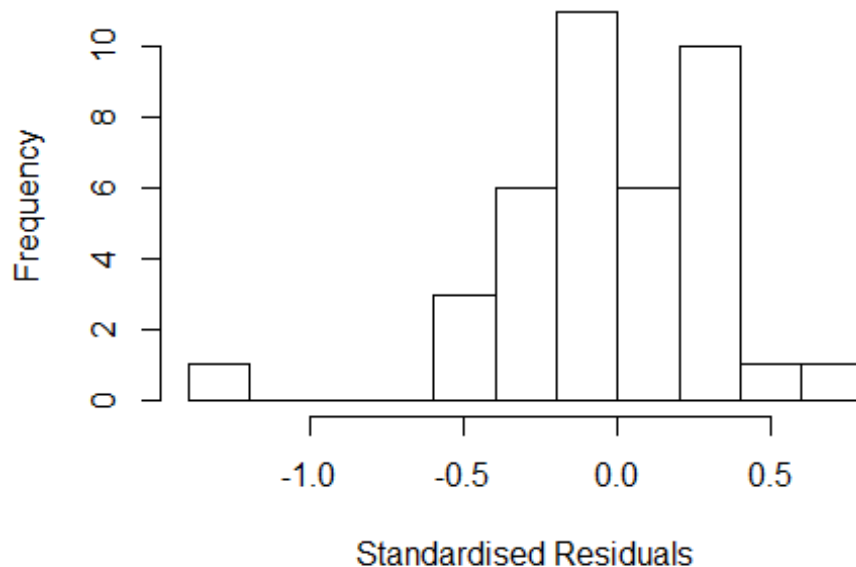
```

```
## 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 | 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
## -1.29e+00 -2.20e-01 8.33e-17 2.27e-01 6.72e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 9.12e-01 1.38e-08 6.61e+07 <2e-16 ***
## FirstAuthorFemale1 -1.25e-01 3.10e-01 -4.00e-01 0.691
## LastAuthorFemale1 3.05e-01 2.38e-01 1.28e+00 0.214
## UniqueAuthors2 4.84e-01 5.27e-01 9.20e-01 0.368
## UniqueAuthors3 2.44e-01 0.00e+00 Inf <2e-16 ***
## UniqueAuthors4 1.90e-01 2.52e-01 7.60e-01 0.459
## UniqueAuthors5 -2.48e-02 1.67e-01 -1.50e-01 0.883
## Year1997 3.24e-01 2.49e-08 1.30e+07 <2e-16 ***
## Year1999 -4.73e-01 5.27e-01 -9.00e-01 0.379
## Year2003 -1.29e-01 2.35e-01 -5.50e-01 0.588
```

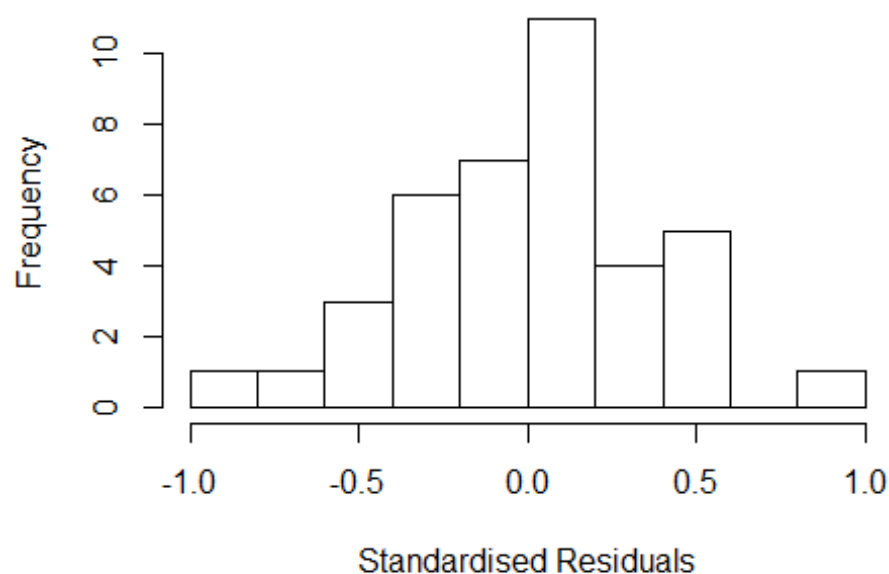


```

## Year2004          -2.41e-01    2.58e-01 -9.30e-01    0.362
## Year2005           2.41e-01    2.26e-01  1.07e+00    0.299
## Year2006          -1.02e-01    4.24e-01 -2.40e-01    0.813
## Year2007          -7.61e-02    3.09e-01 -2.50e-01    0.808
## Year2008          -1.58e-02    3.88e-01 -4.00e-02    0.968
## Year2009          -2.90e-01    2.15e-01 -1.35e+00    0.192
## Year2010           1.31e-01    2.49e-01  5.30e-01    0.605
## Year2011           1.89e-01    3.68e-01  5.10e-01    0.612
## Year2012           4.45e-01    1.82e-01  2.45e+00    0.024 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.428
## Multiple R-squared:  0.271, Adjusted R-squared:  -0.386
## Convergence in 32 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 29 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.347  0.923  0.964   0.928   0.978   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          2.56e-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 2: First author gender, Last author gender, Year as
factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 14.369 1          3.791
## LastAuthorFemale  5.777 1          2.404
## Year              67.090 12         1.192

```

Residuals from first and last author



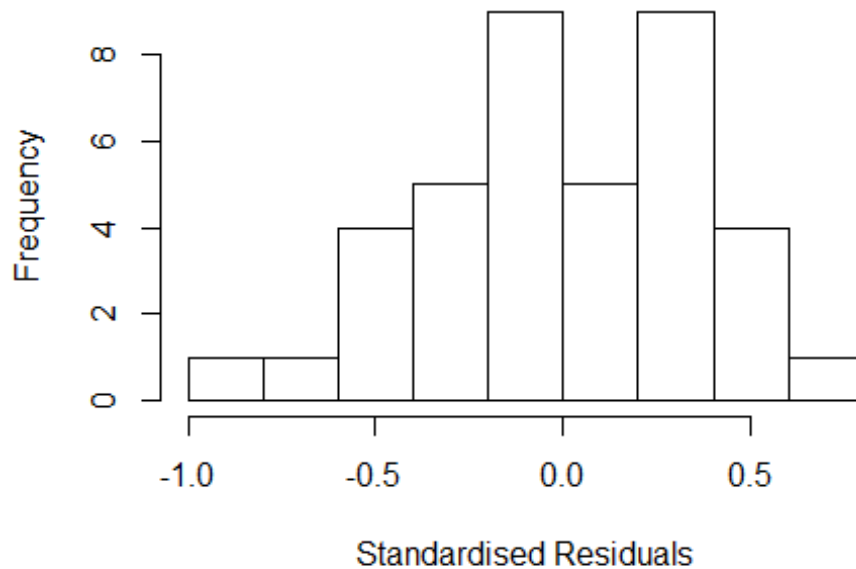
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.93574 -0.22349 0.00782 0.19816 0.81304
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0340 0.0876 11.81 1.7e-11 ***
## FirstAuthorFemale1 -0.1966 0.1668 -1.18 0.250
## LastAuthorFemale1 0.1664 0.1659 1.00 0.326
## Year1997 0.2020 0.0876 2.31 0.030 *
## Year1999 -0.1110 0.0876 -1.27 0.217
## Year2003 -0.0112 0.2482 -0.05 0.964
## Year2004 -0.1754 0.1851 -0.95 0.353
## Year2005 0.2978 0.1830 1.63 0.117
## Year2006 0.1767 0.3150 0.56 0.580
## Year2007 0.1267 0.2536 0.50 0.622
## Year2008 0.1148 0.2929 0.39 0.699
## Year2009 -0.0680 0.2791 -0.24 0.809
```

```

## Year2010          0.2327      0.1883      1.24      0.229
## Year2011          0.1858      0.5703      0.33      0.747
## Year2012          0.3338      0.1238      2.70      0.013 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.429
## Multiple R-squared:  0.167, Adjusted R-squared:  -0.318
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 31 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.614  0.903   0.967   0.926   0.987   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      2.56e-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 1.163 1          1.078
## Year              1.163 12          1.006

```

Residuals from first author



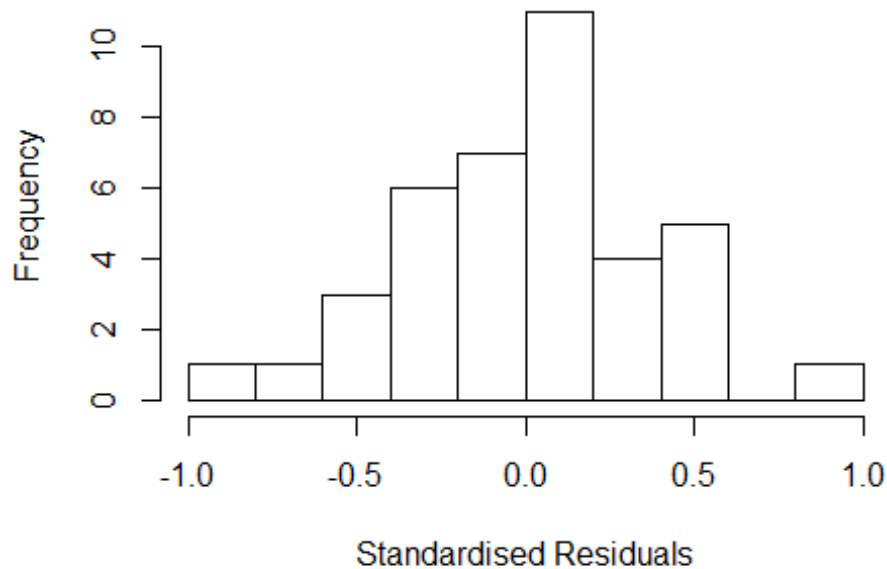
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8.75e-01 -2.41e-01  2.50e-16  2.55e-01  7.21e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0340     0.0873   11.85 9.4e-12 ***
## FirstAuthorFemale1 -0.1828     0.1698   -1.08  0.292
## Year1997          0.2020     0.0873    2.31  0.029 *
## Year1999         -0.1110     0.0873   -1.27  0.215
## Year2003          0.0959     0.2467    0.39  0.701
## Year2004         -0.0991     0.1967   -0.50  0.619
## Year2005          0.3810     0.2211    1.72  0.097 .
## Year2006          0.3290     0.2297    1.43  0.165
## Year2007          0.2068     0.2544    0.81  0.424
## Year2008          0.1049     0.2873    0.37  0.718
## Year2009          0.0242     0.2172    0.11  0.912
## Year2010          0.3097     0.1781    1.74  0.094 .
```

```

## Year2011          0.1789      0.5021      0.36      0.725
## Year2012          0.3269      0.1237      2.64      0.014 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.489
## Multiple R-squared:  0.139, Adjusted R-squared:  -0.309
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 32 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.730  0.919  0.970  0.943  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      2.56e-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 5.369 1          2.317
## Year            5.369 12          1.073

```

Residuals from last author



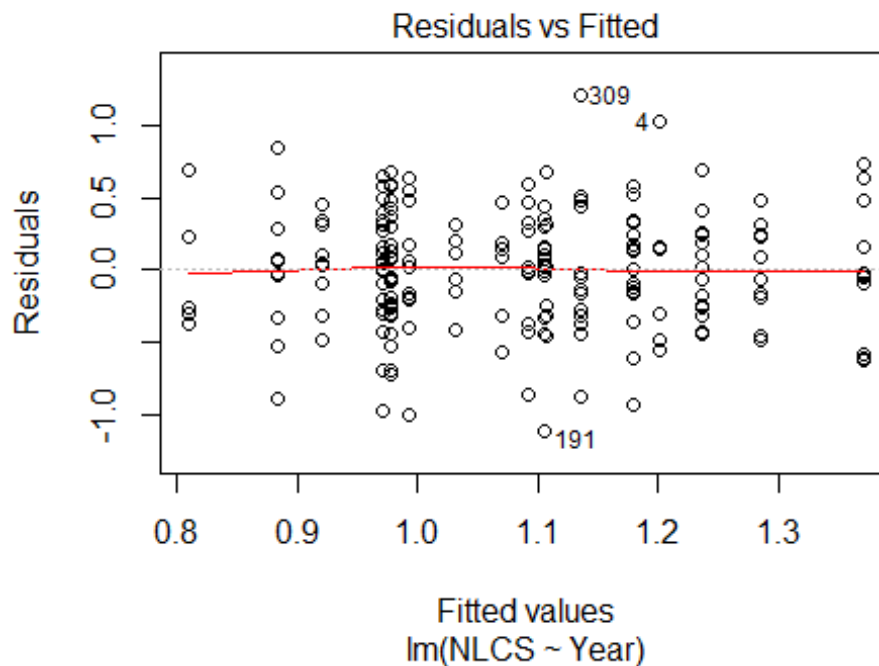
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.10e+00 -2.10e-01 -5.13e-16 2.33e-01 8.34e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.03400 0.08774 11.78 1.1e-11 ***
## LastAuthorFemale1 0.15222 0.18584 0.82 0.420
## Year1997 0.20200 0.08774 2.30 0.030 *
## Year1999 -0.11100 0.08774 -1.27 0.217
## Year2003 -0.05843 0.20625 -0.28 0.779
## Year2004 -0.26661 0.21268 -1.25 0.222
## Year2005 0.30489 0.19254 1.58 0.126
## Year2006 -0.00551 0.24011 -0.02 0.982
## Year2007 0.08039 0.23628 0.34 0.737
## Year2008 -0.03014 0.32900 -0.09 0.928
## Year2009 -0.08887 0.29097 -0.31 0.763
## Year2010 0.11178 0.16313 0.69 0.500
```

```

## Year2011          0.08750    0.41559    0.21    0.835
## Year2012          0.23550    0.09769    2.41    0.024 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.403
## Multiple R-squared:  0.138, Adjusted R-squared:  -0.31
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 3 weights are ~= 1. The remaining 36 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.438  0.900  0.970  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      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 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
##   20   11   10   15   21   13   15   15   18   17   17   14   19   22   37
## 2011 2012
##   26   30
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    6    5   10   14    6   10   11   15   12   11    6   12   13   27
## 2011 2012
##   17   22
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    4    5    7   13    4    8   11    8    9    8    5    9    9   23

```

```
## 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 = 12, df = 16, p-value = 0.8
```

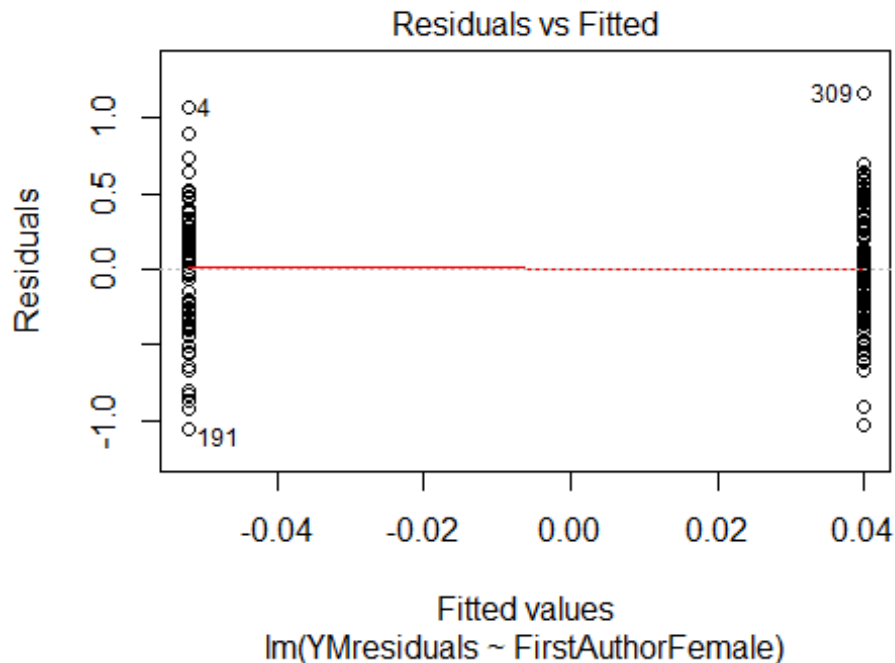


```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.1, df = 1, p-value = 0.3
## [1] "Female first author team size 2018 geometric mean: 6.01198451522228"
## [1] "Male first author team size 2018 geometric mean: 4.7622031559046"
## Warning in wilcox.test.default(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] "Female last author team size 2018 geometric mean: 7.96856520746051"
## [1] "Male last author team size 2018 geometric mean: 5.09236788003047"

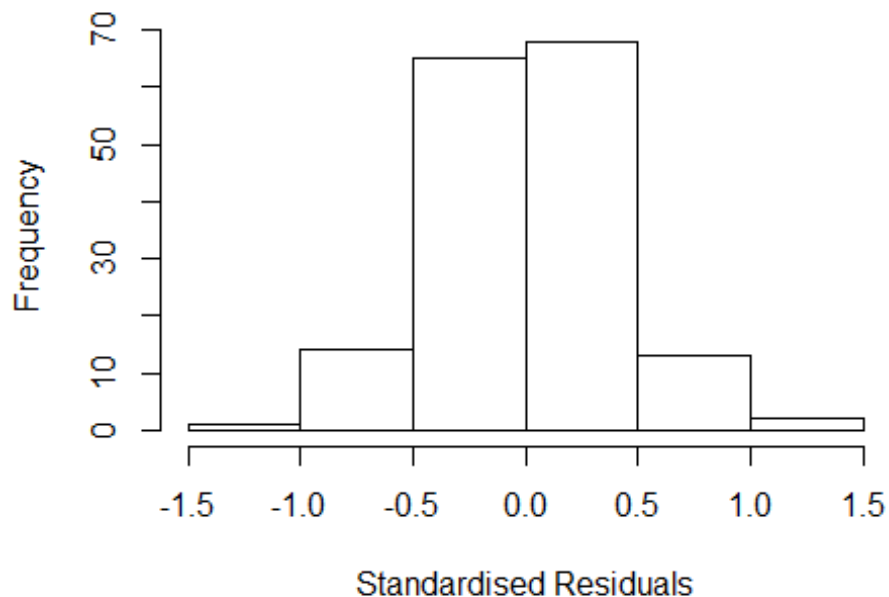
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 34, 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 | 1.788 | 1 | 1.337 |
| LastAuthorFemale | 1.182 | 1 | 1.087 |
| UniqueAuthors | 4.311 | 4 | 1.200 |
| Year | 5.492 | 16 | 1.055 |

Residuals from first and last author and team size



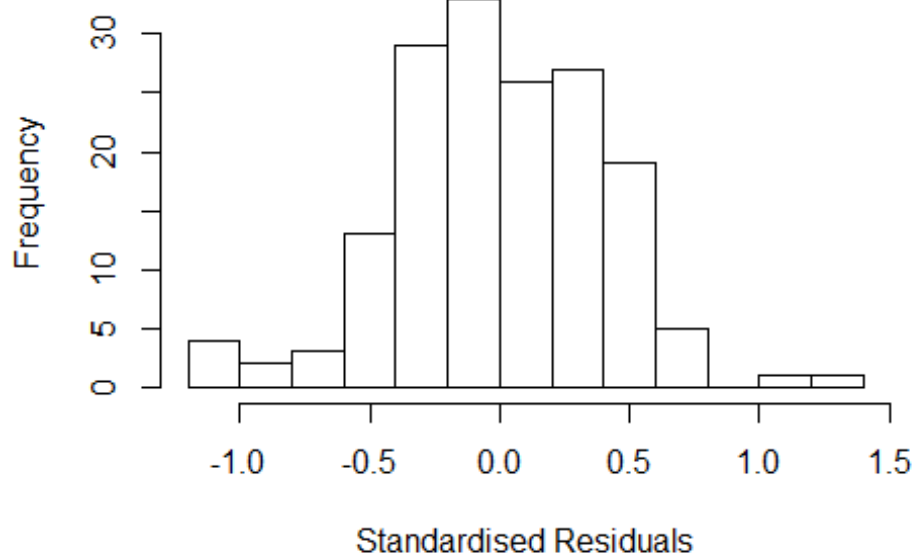
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3080 -0.2525 0.0156 0.2786 1.2942
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.05573 0.31501 3.35 0.001 **
## FirstAuthorFemale1 0.08083 0.07695 1.05 0.295
## LastAuthorFemale1 -0.08610 0.08211 -1.05 0.296
## UniqueAuthors2 -0.24953 0.19879 -1.26 0.211
## UniqueAuthors3 -0.21375 0.20331 -1.05 0.295
## UniqueAuthors4 -0.12995 0.20529 -0.63 0.528
## UniqueAuthors5 0.00400 0.19188 0.02 0.983
## Year1997 0.21527 0.28409 0.76 0.450
## Year1998 -0.20547 0.32172 -0.64 0.524
## Year1999 0.17026 0.29445 0.58 0.564
```

```

## Year2000      0.26553      0.27843      0.95      0.342
## Year2001      0.20219      0.29374      0.69      0.492
## Year2002     -0.02776      0.30033     -0.09      0.926
## Year2003     -0.00832      0.28401     -0.03      0.977
## Year2004      0.25223      0.27695      0.91      0.364
## Year2005      0.03696      0.26857      0.14      0.891
## Year2006      0.57026      0.28410      2.01      0.047 *
## Year2007      0.24146      0.32528      0.74      0.459
## Year2008      0.21968      0.29610      0.74      0.459
## Year2009      0.13757      0.28926      0.48      0.635
## Year2010      0.03977      0.26520      0.15      0.881
## Year2011      0.15481      0.26782      0.58      0.564
## Year2012      0.04822      0.28069      0.17      0.864
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.405
## Multiple R-squared:  0.203, Adjusted R-squared:  0.0777
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 154 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.275  0.900  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      6.13e-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 1.569 1      1.253
## LastAuthorFemale  1.140 1      1.068
## Year              1.754 16      1.018

```

Residuals from first and last author



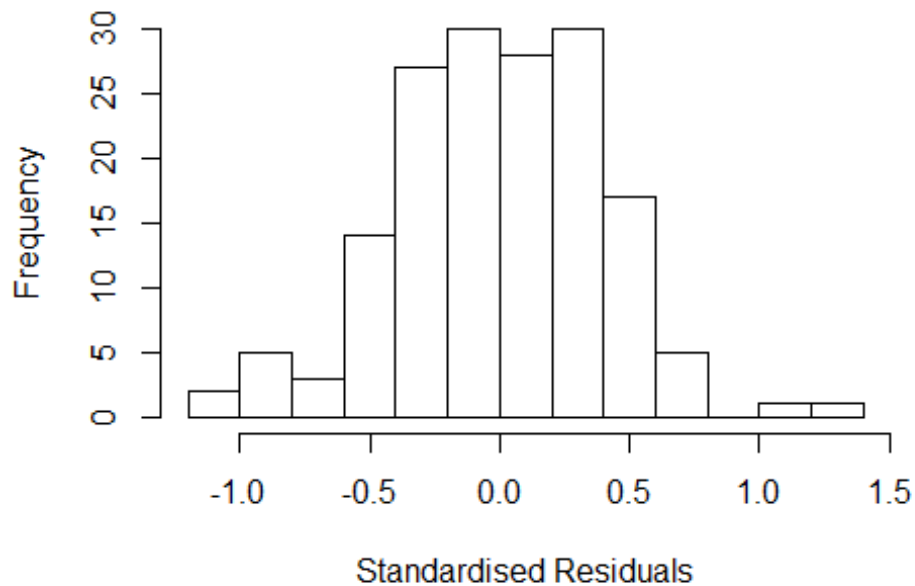
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1595 -0.2653 -0.0148 0.2849 1.2660
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.95402 0.28011 3.41 0.00086 ***
## FirstAuthorFemale1 0.09736 0.07367 1.32 0.18838
## LastAuthorFemale1 -0.11923 0.08194 -1.46 0.14782
## Year1997 0.20289 0.30972 0.66 0.51347
## Year1998 -0.24993 0.35094 -0.71 0.47752
## Year1999 0.20922 0.30834 0.68 0.49852
## Year2000 0.24375 0.28916 0.84 0.40066
## Year2001 0.09876 0.31954 0.31 0.75771
## Year2002 0.00232 0.32696 0.01 0.99434
## Year2003 0.01841 0.30214 0.06 0.95151
## Year2004 0.20549 0.29560 0.70 0.48807
## Year2005 0.03906 0.28962 0.13 0.89290
```

```

## Year2006          0.59265    0.29632    2.00  0.04737 *
## Year2007          0.15601    0.32794    0.48  0.63499
## Year2008          0.19362    0.31707    0.61  0.54240
## Year2009          0.13007    0.30485    0.43  0.67026
## Year2010         -0.02149    0.28382   -0.08  0.93974
## Year2011          0.17957    0.28843    0.62  0.53456
## Year2012          0.06505    0.30219    0.22  0.82987
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.398
## Multiple R-squared:  0.153, Adjusted R-squared:  0.0473
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 147 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.291  0.888  0.950  0.905  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      6.13e-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 1.557 1      1.248
## Year              1.557 16      1.014

```

Residuals from first author



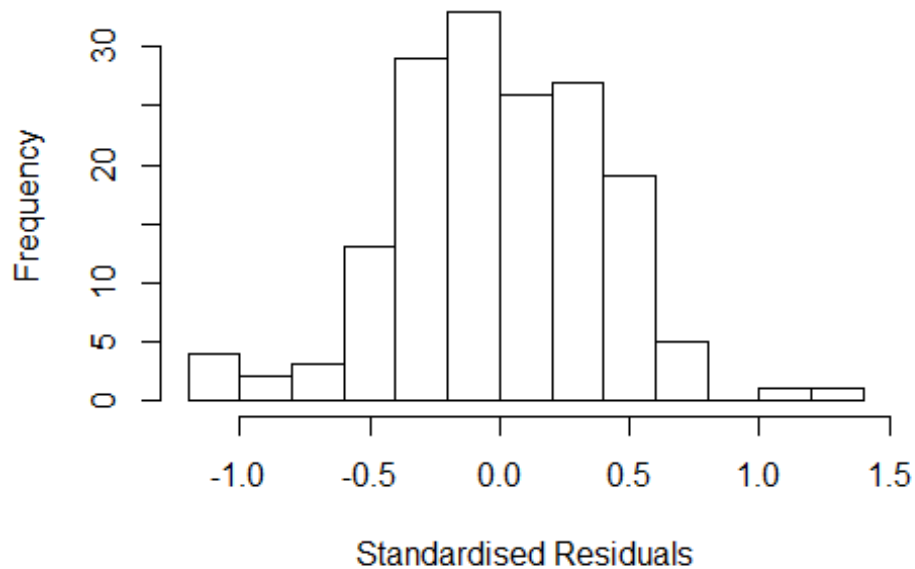
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.145625 -0.259905 0.000921 0.286049 1.257326
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.96267 0.27903 3.45 0.00073 ***
## FirstAuthorFemale1 0.08599 0.07418 1.16 0.24828
## Year1997 0.17541 0.32173 0.55 0.58645
## Year1998 -0.24994 0.34757 -0.72 0.47324
## Year1999 0.20508 0.30582 0.67 0.50355
## Year2000 0.21314 0.28896 0.74 0.46195
## Year2001 0.06315 0.30950 0.20 0.83860
## Year2002 -0.02264 0.32619 -0.07 0.94476
## Year2003 0.00257 0.29985 0.01 0.99316
## Year2004 0.18295 0.29581 0.62 0.53724
## Year2005 0.00900 0.28460 0.03 0.97482
## Year2006 0.57307 0.29303 1.96 0.05242 .
```

```

## Year2007          0.15147      0.32876      0.46  0.64569
## Year2008          0.15842      0.31439      0.50  0.61511
## Year2009          0.10271      0.30212      0.34  0.73438
## Year2010         -0.05260      0.28203     -0.19  0.85232
## Year2011          0.18005      0.28739      0.63  0.53197
## Year2012          0.02752      0.29865      0.09  0.92670
## ---
## 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:  0.0423
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 146 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.300  0.885   0.951   0.904   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      6.13e-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 1.14 1      1.068
## Year              1.14 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
## -1.2043 -0.2436 -0.0382 0.2654 1.2040
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0160 0.2598 3.91 0.00014 ***
## LastAuthorFemale1 -0.1042 0.0829 -1.26 0.21050
## Year1997 0.1862 0.2883 0.65 0.51940
## Year1998 -0.2422 0.3222 -0.75 0.45342
## Year1999 0.1866 0.2861 0.65 0.51531
## Year2000 0.2332 0.2740 0.85 0.39604
## Year2001 0.0563 0.3126 0.18 0.85743
## Year2002 -0.0112 0.3197 -0.04 0.97200
## Year2003 0.0191 0.2950 0.06 0.94843
## Year2004 0.1883 0.2841 0.66 0.50845
## Year2005 0.0130 0.2740 0.05 0.96211
## Year2006 0.5862 0.2834 2.07 0.04042 *
```

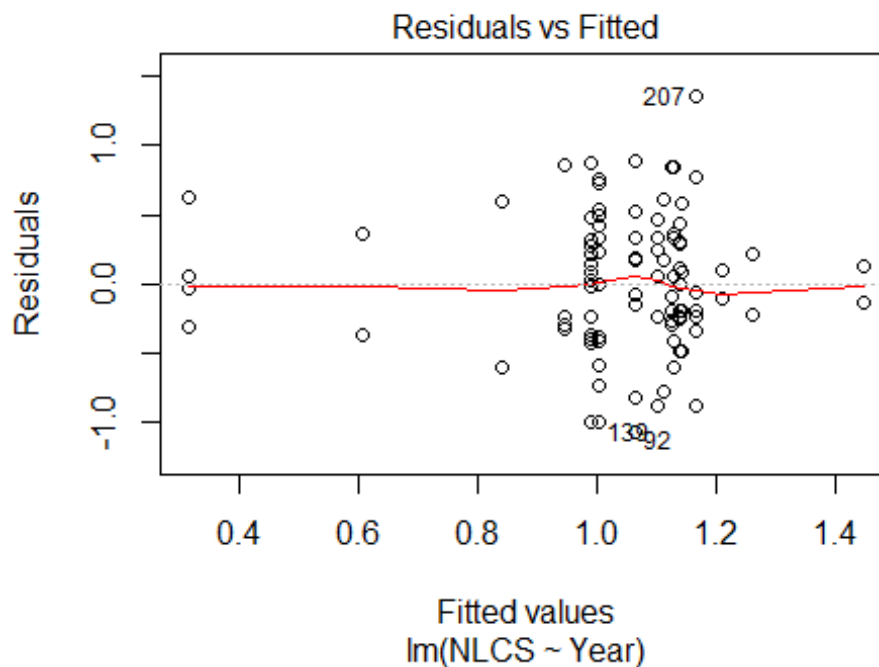


```

## Year2007          0.1278      0.3324      0.38  0.70110
## Year2008          0.1887      0.3018      0.63  0.53275
## Year2009          0.1305      0.2953      0.44  0.65925
## Year2010         -0.0263      0.2718     -0.10  0.92302
## Year2011          0.1979      0.2792      0.71  0.47972
## Year2012          0.0742      0.2949      0.25  0.80177
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.395
## Multiple R-squared:  0.143, Adjusted R-squared:  0.042
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 149 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.333  0.887  0.954  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      6.13e-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: 163"
## [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
##   11   11   9   4   12   8   3   9   2   12   6   22   6   18   16
## 2011 2012
##   14   12
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    2    2    4    6    3    2    5    2    9    2   14    4    9   14
## 2011 2012

```

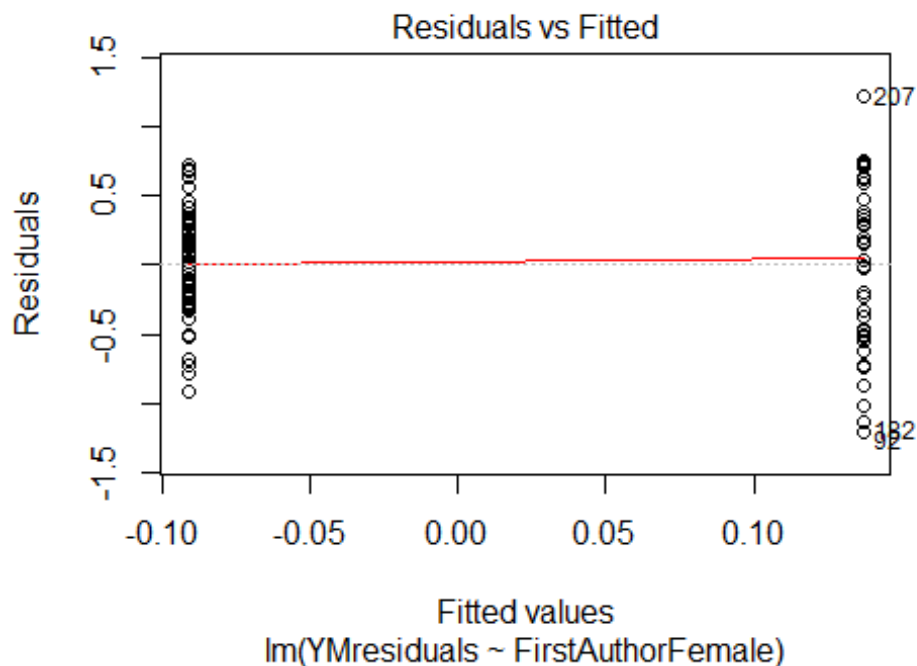
```
##      9      7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      6      2      2      3      5      2      2      5      2      9      2     13      3      8      8
## 2011 2012
##      7      6
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 9.7, df = 16, p-value = 0.9
```



```
##
## 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: 8.14325284978472"
## [1] "Male first 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
```

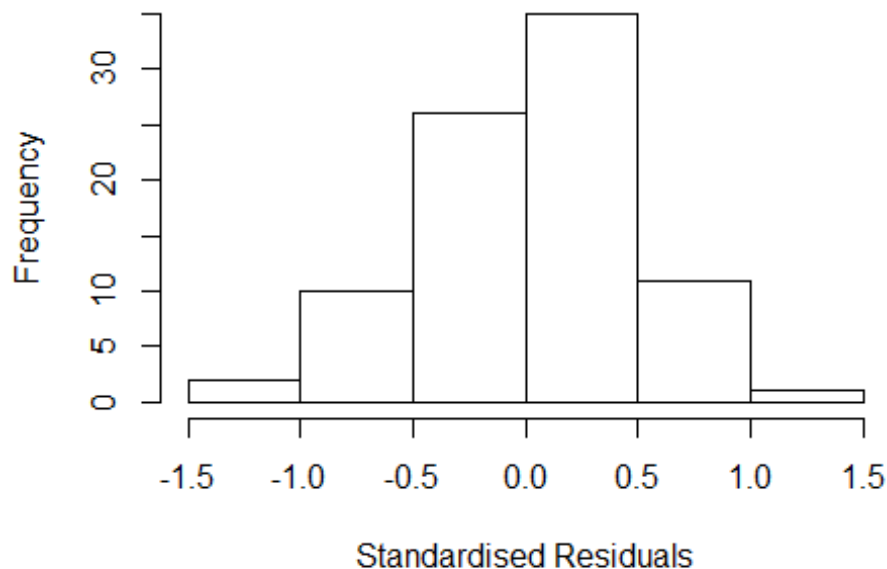
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 12, p-value = 0.07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6"
## [1] "Male last author team size 2018 geometric mean: 5.88566191276542"

## Warning in wilcox.test.default(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 2.612 1      1.616
## LastAuthorFemale  2.533 1      1.591
## UniqueAuthors    43.504 4      1.603
## Year              84.138 16     1.149
```

Residuals from first and last author and team size



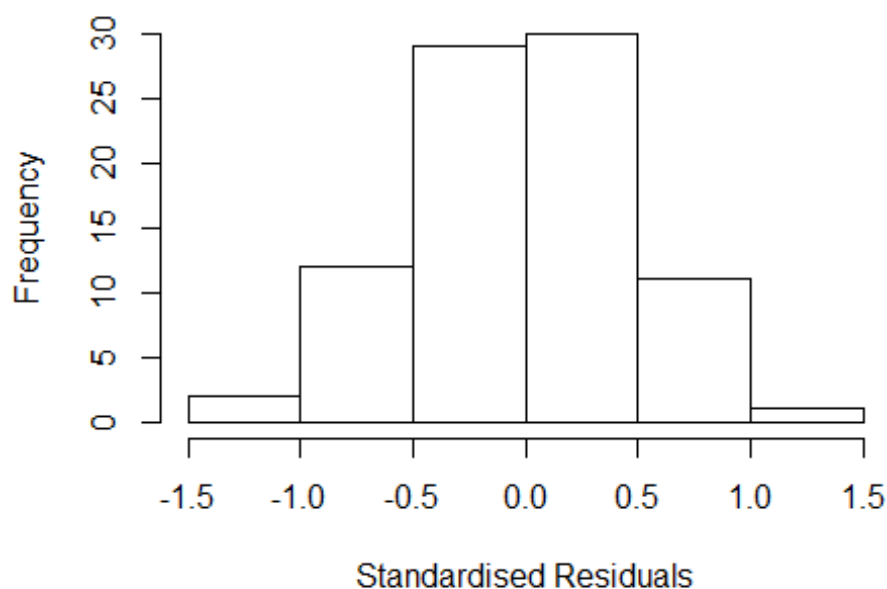
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2679 -0.2669 0.0233 0.2476 1.2374
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.2669 0.1576 1.69 0.09538 .
## FirstAuthorFemale1 0.2711 0.1283 2.11 0.03854 *
## LastAuthorFemale1 0.2727 0.1766 1.54 0.12760
## UniqueAuthors2 0.1551 0.2260 0.69 0.49525
## UniqueAuthors3 0.2357 0.3025 0.78 0.43881
## UniqueAuthors4 0.0500 0.1728 0.29 0.77332
## UniqueAuthors5 0.2578 0.1609 1.60 0.11407
## Year1997 0.7377 0.2233 3.30 0.00159 **
## Year1998 0.7882 0.2263 3.48 0.00091 ***
## Year1999 0.6721 0.3212 2.09 0.04053 *
```

```

## Year2000          0.5689      0.2536      2.24  0.02846 *
## Year2001          0.8886      0.2276      3.90  0.00024 ***
## Year2002          0.8371      0.2090      4.00  0.00017 ***
## Year2003          0.5998      0.2335      2.57  0.01262 *
## Year2004          0.0470      0.5038      0.09  0.92597
## Year2005          0.5748      0.2399      2.40  0.01959 *
## Year2006          0.0555      0.6608      0.08  0.93339
## Year2007          0.2822      0.2330      1.21  0.23042
## Year2008          0.4727      0.3485      1.36  0.17990
## Year2009          0.4217      0.2089      2.02  0.04787 *
## Year2010          0.4412      0.2132      2.07  0.04266 *
## Year2011          0.4858      0.2822      1.72  0.09018 .
## Year2012          0.5727      0.3129      1.83  0.07198 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.498
## Multiple R-squared:  0.338, Adjusted R-squared:  0.103
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 74 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.496  0.889  0.955  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      1.18e-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 3.594 1      1.896
## LastAuthorFemale  2.223 1      1.491
## Year              5.009 16      1.052

```

Residuals from first and last author



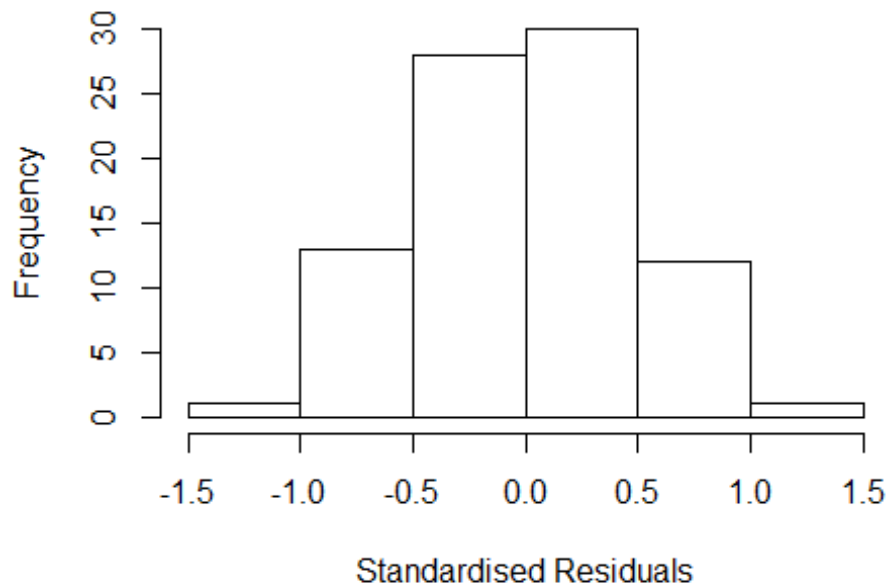
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2436 -0.2992 -0.0112 0.3089 1.3159
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.299 0.129 2.32 0.02337 *
## FirstAuthorFemale1 0.312 0.141 2.20 0.03105 *
## LastAuthorFemale1 0.272 0.183 1.49 0.14222
## Year1997 0.912 0.149 6.11 5.9e-08 ***
## Year1998 0.962 0.204 4.71 1.3e-05 ***
## Year1999 0.788 0.326 2.42 0.01847 *
## Year2000 0.676 0.230 2.94 0.00458 **
## Year2001 1.042 0.155 6.74 4.7e-09 ***
## Year2002 0.991 0.149 6.67 6.3e-09 ***
## Year2003 0.679 0.190 3.57 0.00066 ***
## Year2004 0.149 0.504 0.30 0.76776
## Year2005 0.633 0.224 2.82 0.00630 **
```

```

## Year2006          0.250      0.697      0.36  0.72091
## Year2007          0.392      0.246      1.59  0.11567
## Year2008          0.487      0.355      1.37  0.17488
## Year2009          0.574      0.180      3.18  0.00221 **
## Year2010          0.627      0.201      3.11  0.00272 **
## Year2011          0.592      0.260      2.28  0.02616 *
## Year2012          0.743      0.251      2.96  0.00430 **
## ---
## 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:  0.138
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 74 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.399  0.864  0.944  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      1.18e-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 2.705 1      1.645
## Year              2.705 16      1.032

```

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
## -1.32686 -0.30232 0.00622 0.28833 1.23778
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.299 0.129 2.32 0.02344 *
## FirstAuthorFemale1 0.385 0.128 3.01 0.00369 **
## Year1997 0.912 0.149 6.12 5.5e-08 ***
## Year1998 0.963 0.205 4.71 1.3e-05 ***
## Year1999 0.788 0.327 2.41 0.01882 *
## Year2000 0.659 0.227 2.91 0.00497 **
## Year2001 1.006 0.145 6.93 2.0e-09 ***
## Year2002 0.955 0.150 6.38 1.9e-08 ***
## Year2003 0.656 0.187 3.51 0.00079 ***
## Year2004 0.113 0.568 0.20 0.84269
## Year2005 0.643 0.226 2.85 0.00583 **
## Year2006 0.350 1.537 0.23 0.82067
```

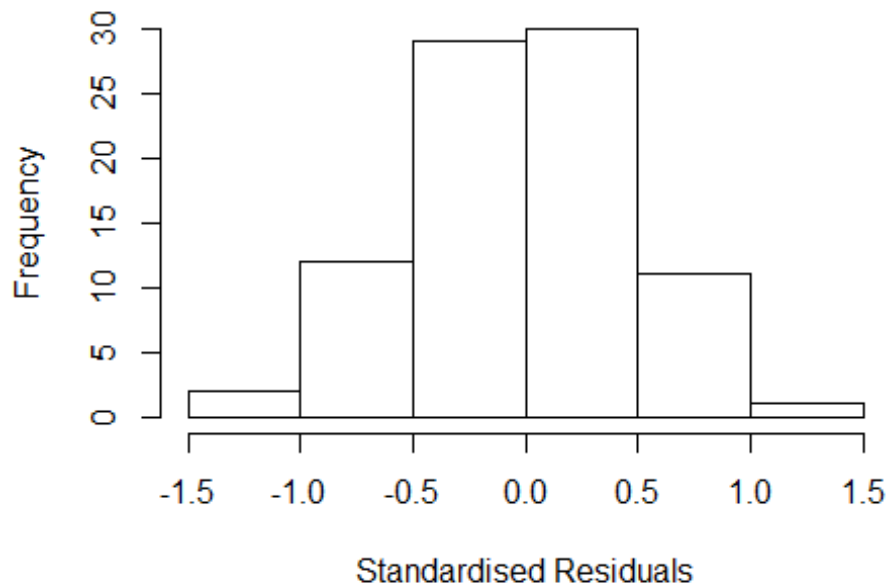


```

## Year2007          0.416      0.241      1.73  0.08832 .
## Year2008          0.447      0.351      1.27  0.20778
## Year2009          0.593      0.189      3.14  0.00252 **
## Year2010          0.718      0.190      3.79  0.00033 ***
## Year2011          0.598      0.305      1.96  0.05425 .
## Year2012          0.776      0.242      3.21  0.00207 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.457
## Multiple R-squared:  0.304, Adjusted R-squared:  0.127
## Convergence in 29 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 73 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.380  0.838  0.934  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      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"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.549 1      1.244
## Year            1.549 16      1.014

```

Residuals from last author



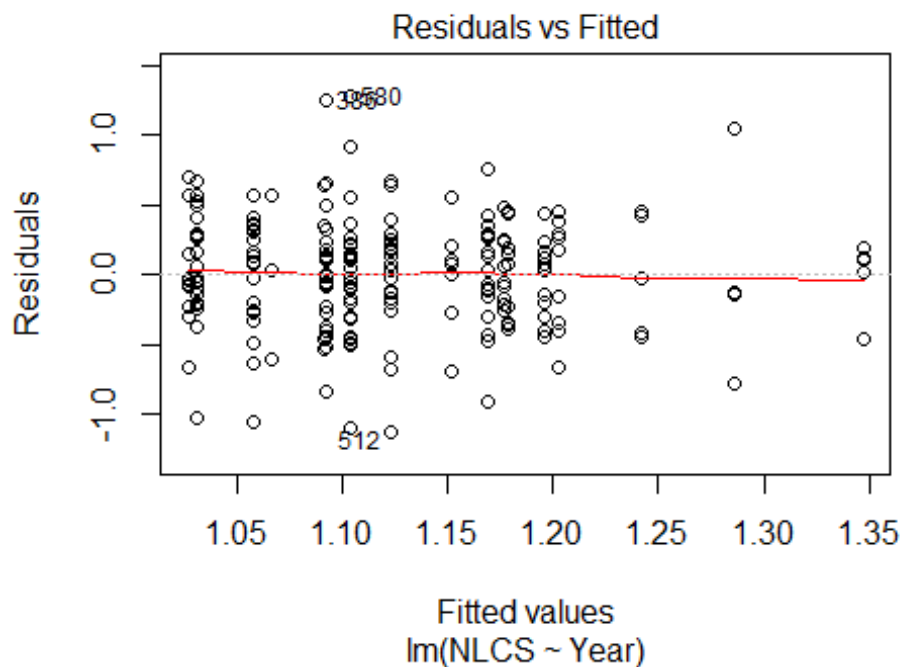
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0512 -0.2436 -0.0349 0.3498 1.4639
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.301 0.129 2.34 0.02220 *
## LastAuthorFemale1 0.397 0.169 2.35 0.02159 *
## Year1997 0.910 0.149 6.11 5.7e-08 ***
## Year1998 0.960 0.204 4.72 1.3e-05 ***
## Year1999 0.791 0.321 2.47 0.01621 *
## Year2000 0.738 0.245 3.02 0.00360 **
## Year2001 1.196 0.206 5.80 1.9e-07 ***
## Year2002 1.145 0.161 7.11 9.7e-10 ***
## Year2003 0.780 0.223 3.50 0.00082 ***
## Year2004 0.303 0.313 0.97 0.33606
## Year2005 0.750 0.238 3.15 0.00247 **
## Year2006 0.341 0.357 0.96 0.34255
```

```

## Year2007          0.553      0.232      2.39  0.01973 *
## Year2008          0.667      0.397      1.68  0.09762 .
## Year2009          0.769      0.160      4.82  8.7e-06 ***
## Year2010          0.618      0.211      2.93  0.00463 **
## Year2011          0.754      0.250      3.01  0.00366 **
## Year2012          0.812      0.279      2.91  0.00484 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.499
## Multiple R-squared:  0.267, Adjusted R-squared:  0.0806
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 79 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.370  0.890  0.964  0.920  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      1.18e-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 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
##   11   15   16    8   25   20   19   15   10   19   24   30   39   29   51
## 2011 2012
##   48   55
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    4    4    5   10    7   10    9    5    9   14   20   23   22   24
## 2011 2012

```

```
## 23 28
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2 3 4 5 8 6 5 7 3 7 7 18 19 18 18
## 2011 2012
## 22 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 = 15, df = 16, p-value = 0.5
```



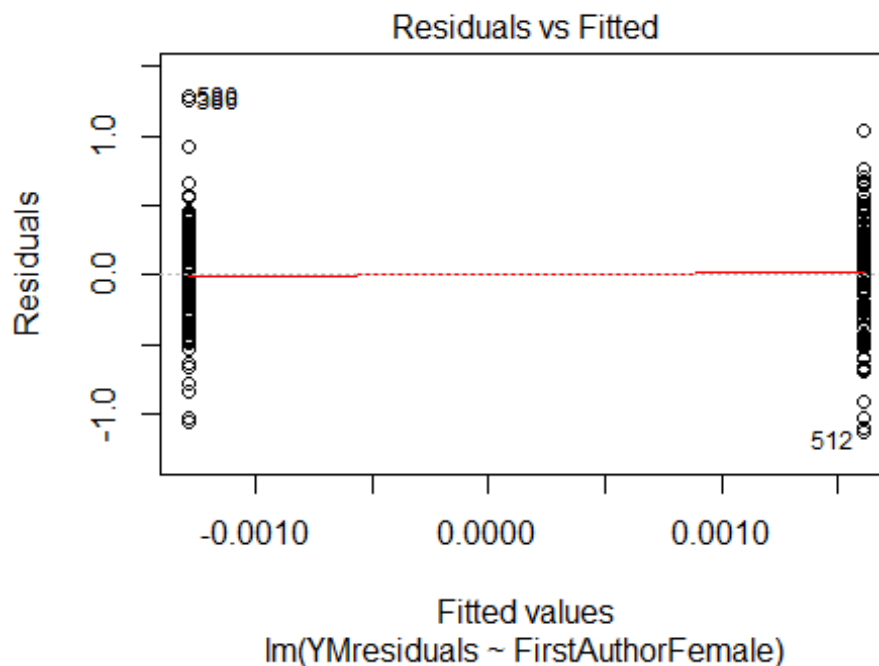
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1, df = 1, p-value = 0.3

## [1] "Female first author team size 2018 geometric mean: 4.74152530679434"
## [1] "Male first author team size 2018 geometric mean: 4.2831900971392"

## Warning in wilcox.test.default(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 = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.17416387779665"
## [1] "Male last author team size 2018 geometric mean: 3.97839213700271"

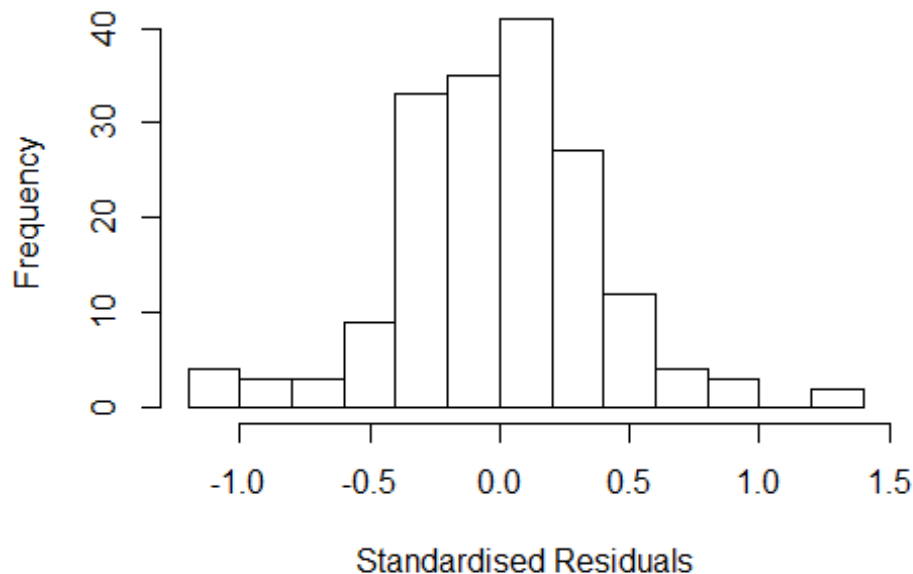
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## 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 | 1.621 | 1 | 1.273 |
| LastAuthorFemale | 1.759 | 1 | 1.326 |
| UniqueAuthors | 5.315 | 4 | 1.232 |
| Year | 7.583 | 16 | 1.065 |

Residuals from first and last author and team size



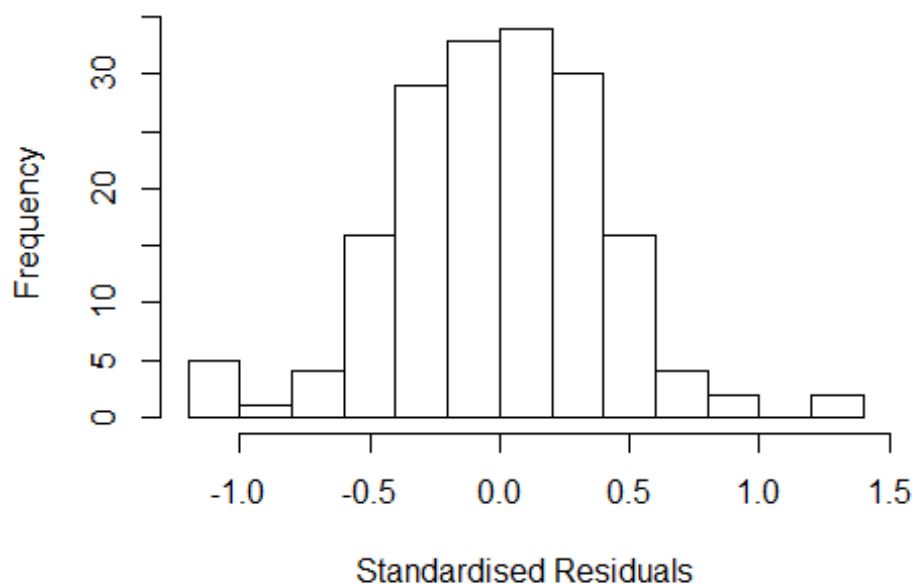
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1627 -0.2501 0.0146 0.2187 1.2635
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.13722 0.32814 3.47 0.00069 ***
## FirstAuthorFemale1 -0.04873 0.06197 -0.79 0.43281
## LastAuthorFemale1 -0.02169 0.06897 -0.31 0.75360
## UniqueAuthors2 -0.32201 0.15264 -2.11 0.03652 *
## UniqueAuthors3 -0.14311 0.15882 -0.90 0.36895
## UniqueAuthors4 0.00346 0.13637 0.03 0.97977
## UniqueAuthors5 -0.05071 0.12904 -0.39 0.69487
## Year1997 0.31014 0.50963 0.61 0.54372
## Year1998 0.00179 0.43461 0.00 0.99672
## Year1999 0.18881 0.38788 0.49 0.62711
```

```

## Year2000          0.12463      0.31622      0.39  0.69405
## Year2001          0.23798      0.33114      0.72  0.47345
## Year2002          0.16492      0.37564      0.44  0.66125
## Year2003          0.09481      0.33051      0.29  0.77461
## Year2004          0.38194      0.30844      1.24  0.21749
## Year2005          0.14143      0.33849      0.42  0.67666
## Year2006          0.17740      0.31835      0.56  0.57817
## Year2007          0.08085      0.31344      0.26  0.79680
## Year2008          0.14765      0.31209      0.47  0.63683
## Year2009          0.05625      0.31330      0.18  0.85776
## Year2010          0.07617      0.30234      0.25  0.80142
## Year2011          0.14747      0.30649      0.48  0.63109
## Year2012          0.08270      0.30828      0.27  0.78886
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.345
## Multiple R-squared:  0.11,   Adjusted R-squared:  -0.0174
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 166 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.152  0.883   0.954   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           5.68e-04           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01           5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500           50           2           1           1000           200
## trace.lev      mts      compute.rd
##           0           1000           0
##           psi           subsampling           cov
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.281 1           1.132
## LastAuthorFemale  1.415 1           1.190
## Year              1.798 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
## -1.158602 -0.259106 -0.000447 0.251904 1.290440
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7783 0.2814 2.77 0.0064 **
## FirstAuthorFemale1 -0.0157 0.0614 -0.26 0.7982
## LastAuthorFemale1 0.0192 0.0645 0.30 0.7661
## Year1997 0.6389 0.4998 1.28 0.2030
## Year1998 0.2994 0.4727 0.63 0.5273
## Year1999 0.4728 0.3490 1.35 0.1774
## Year2000 0.3911 0.2905 1.35 0.1802
## Year2001 0.4560 0.3159 1.44 0.1509
## Year2002 0.4676 0.3356 1.39 0.1655
## Year2003 0.3270 0.3184 1.03 0.3061
## Year2004 0.6602 0.2787 2.37 0.0191 *
## Year2005 0.4325 0.3164 1.37 0.1737
```

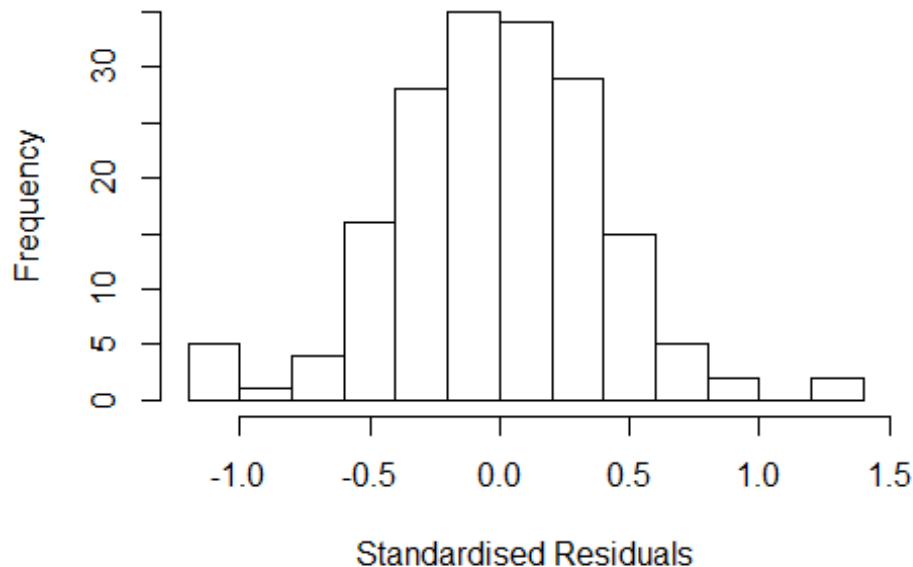


```

## Year2006          0.4824      0.2902      1.66      0.0984 .
## Year2007          0.3177      0.2918      1.09      0.2780
## Year2008          0.3945      0.2930      1.35      0.1801
## Year2009          0.3002      0.2918      1.03      0.3051
## Year2010          0.2737      0.2923      0.94      0.3506
## Year2011          0.3803      0.2877      1.32      0.1880
## Year2012          0.3310      0.2906      1.14      0.2564
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.374
## Multiple R-squared:  0.0525, Adjusted R-squared:  -0.0561
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 162 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.209  0.879   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      5.68e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##      trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.276 1      1.130
## Year              1.276 16      1.008

```

Residuals from first author



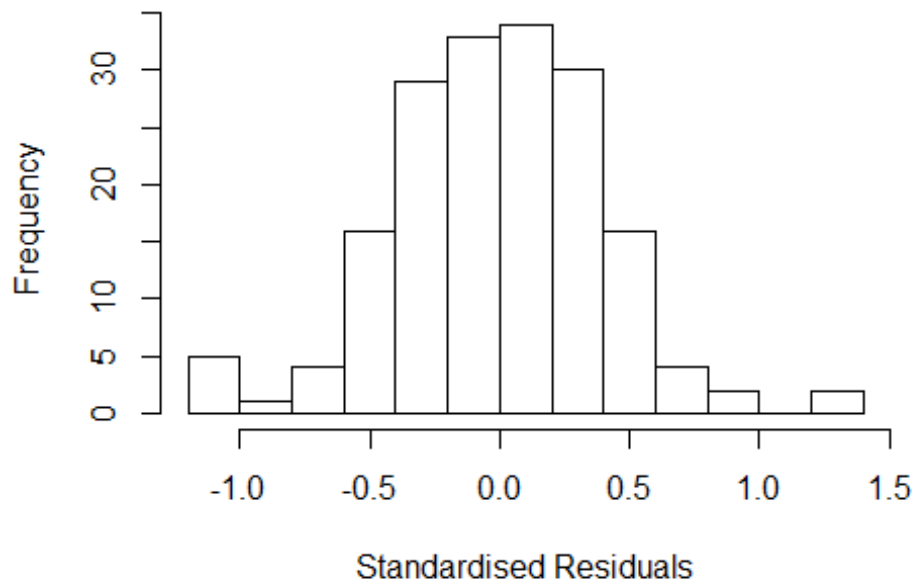
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.165460 -0.261262 -0.000374 0.257219 1.281419
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7866 0.2669 2.95 0.0037 **
## FirstAuthorFemale1 -0.0132 0.0615 -0.21 0.8304
## Year1997 0.6274 0.4833 1.30 0.1962
## Year1998 0.2952 0.4618 0.64 0.5236
## Year1999 0.4626 0.3389 1.36 0.1742
## Year2000 0.3914 0.2782 1.41 0.1614
## Year2001 0.4471 0.3056 1.46 0.1455
## Year2002 0.4582 0.3244 1.41 0.1598
## Year2003 0.3197 0.3084 1.04 0.3015
## Year2004 0.6566 0.2673 2.46 0.0151 *
## Year2005 0.4254 0.3078 1.38 0.1688
## Year2006 0.4727 0.2774 1.70 0.0904 .
```

```

## Year2007          0.3159      0.2844      1.11      0.2685
## Year2008          0.3860      0.2810      1.37      0.1715
## Year2009          0.2944      0.2775      1.06      0.2902
## Year2010          0.2715      0.2840      0.96      0.3406
## Year2011          0.3789      0.2790      1.36      0.1764
## Year2012          0.3292      0.2783      1.18      0.2387
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.373
## Multiple R-squared:  0.0521, Adjusted R-squared:  -0.0499
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 162 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.213  0.877   0.953   0.897   0.986   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      5.68e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.413 1          1.189
## Year            1.413 16          1.011

```

Residuals from last author



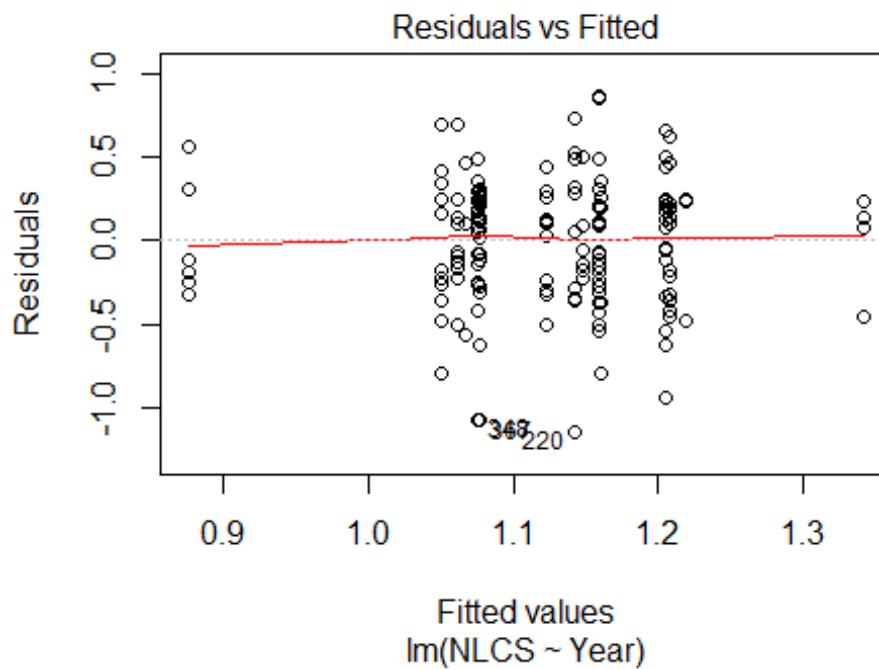
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.14925 -0.25555 0.00693 0.25372 1.28223
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7720 0.2688 2.87 0.0046 **
## LastAuthorFemale1 0.0160 0.0649 0.25 0.8055
## Year1997 0.6365 0.4948 1.29 0.2002
## Year1998 0.2983 0.4756 0.63 0.5315
## Year1999 0.4677 0.3419 1.37 0.1732
## Year2000 0.3907 0.2805 1.39 0.1656
## Year2001 0.4569 0.3052 1.50 0.1364
## Year2002 0.4678 0.3293 1.42 0.1574
## Year2003 0.3235 0.3092 1.05 0.2970
## Year2004 0.6570 0.2682 2.45 0.0154 *
## Year2005 0.4264 0.3078 1.39 0.1679
## Year2006 0.4798 0.2799 1.71 0.0885 .
```

```

## Year2007          0.3156      0.2820      1.12      0.2647
## Year2008          0.3922      0.2826      1.39      0.1672
## Year2009          0.2994      0.2814      1.06      0.2891
## Year2010          0.2740      0.2823      0.97      0.3331
## Year2011          0.3773      0.2765      1.36      0.1744
## Year2012          0.3298      0.2804      1.18      0.2413
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.375
## Multiple R-squared:  0.0518, Adjusted R-squared:  -0.0502
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 160 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.217  0.877   0.954   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      5.68e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 176"
## [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
##    9    8   13   10   18   14   16   12   13   16   20   16   27   22   35
## 2011 2012
##   38   38
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    3    6    8    6    3    6    3    4   10   13   11   17   13   19
## 2011 2012

```

```
## 19 24
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 0 2 5 7 6 2 5 3 1 10 9 9 13 11 16
## 2011 2012
## 17 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 = 8.1, df = 15, p-value = 0.9
```



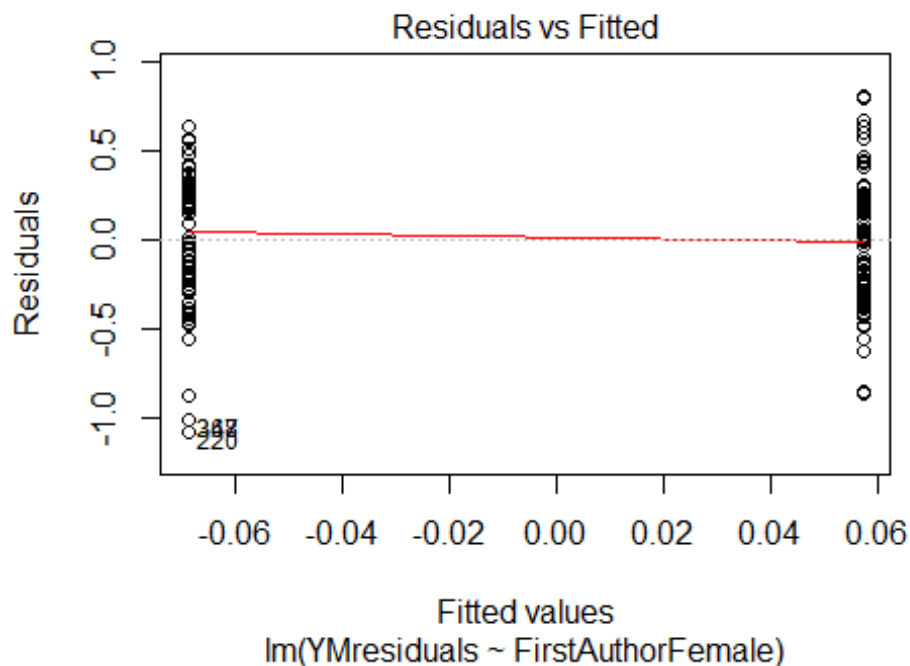
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.42, df = 1, p-value = 0.5

## [1] "Female first author team size 2018 geometric mean: 4.79035262372259"
## [1] "Male first author team size 2018 geometric mean: 5.04643110849529"

## Warning in wilcox.test.default(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 = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 7.16360056283859"
## [1] "Male last author team size 2018 geometric mean: 4.40150079971845"

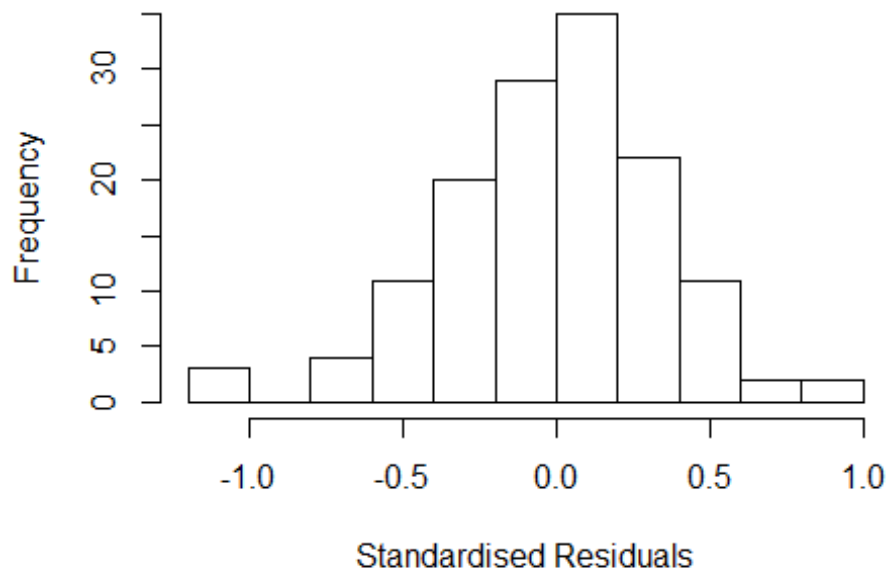
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 72, p-value = 0.02
## 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 | 6.265 | 1 | 2.503 |
| LastAuthorFemale | 10.145 | 1 | 3.185 |
| UniqueAuthors | 81.858 | 4 | 1.734 |
| Year | 643.046 | 15 | 1.241 |

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
## -1.1449 -0.2177 0.0102 0.2143 0.8522
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8587 0.2905 2.96 0.0038 **
## FirstAuthorFemale1 0.0801 0.0687 1.17 0.2462
## LastAuthorFemale1 -0.0758 0.0740 -1.02 0.3077
## UniqueAuthors2 0.3537 0.2206 1.60 0.1116
## UniqueAuthors3 0.4843 0.2149 2.25 0.0261 *
## UniqueAuthors4 0.4531 0.2315 1.96 0.0527 .
## UniqueAuthors5 0.4501 0.2104 2.14 0.0345 *
## Year1998 -0.4520 0.2692 -1.68 0.0958 .
## Year1999 -0.2118 0.2601 -0.81 0.4172
## Year2000 -0.0890 0.2052 -0.43 0.6654
```

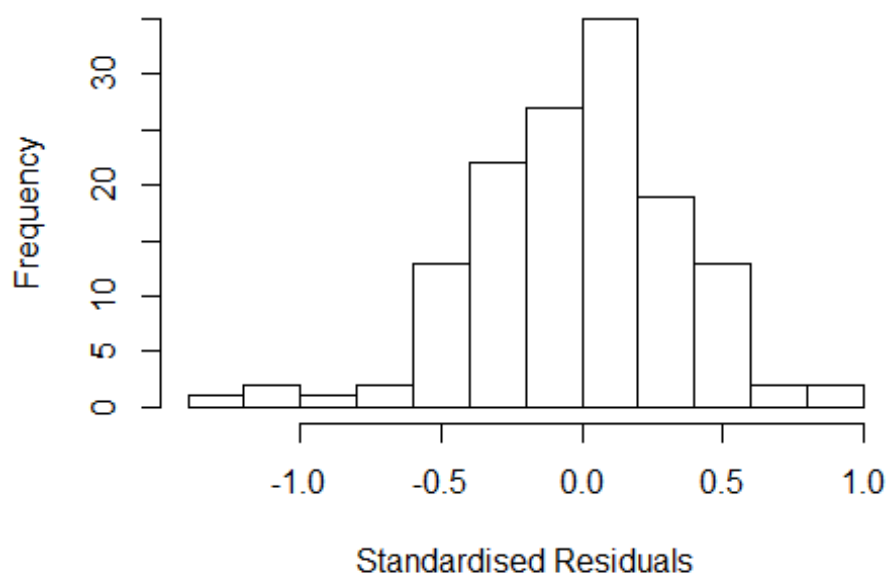


```

## Year2001          -0.0342      0.2017   -0.17   0.8657
## Year2002           0.0344      0.2203    0.16   0.8761
## Year2003           0.0209      0.2055    0.10   0.9192
## Year2004           0.1973      0.2024    0.97   0.3316
## Year2005          -0.2867      0.2121   -1.35   0.1790
## Year2006           0.0232      0.2264    0.10   0.9185
## Year2007          -0.1301      0.2613   -0.50   0.6196
## Year2008          -0.0170      0.2112   -0.08   0.9360
## Year2009          -0.1699      0.2384   -0.71   0.4776
## Year2010          -0.1732      0.1961   -0.88   0.3787
## Year2011          -0.1981      0.1898   -1.04   0.2988
## Year2012          -0.2181      0.1962   -1.11   0.2687
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.322
## Multiple R-squared:  0.175, Adjusted R-squared:  0.0266
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 126 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.180  0.855  0.951  0.893  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      7.19e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 112.89 1          10.625
## LastAuthorFemale  83.38 1           9.131
## Year              304.65 15          1.210

```

Residuals from first and last author



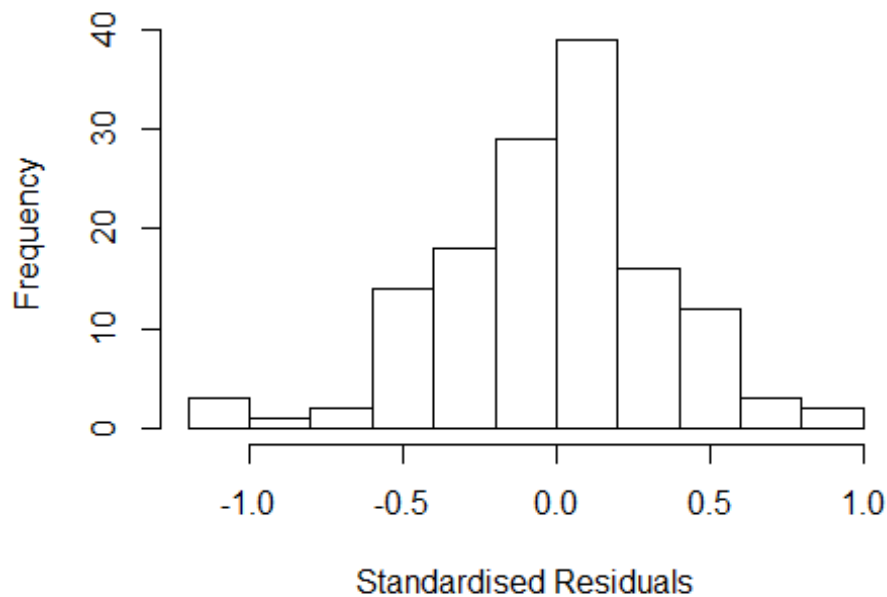
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2118 -0.2566 0.0105 0.2059 0.8430
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.31777 0.20271 6.50 1.9e-09 ***
## FirstAuthorFemale1 0.08287 0.07791 1.06 0.290
## LastAuthorFemale1 -0.09928 0.06646 -1.49 0.138
## Year1998 -0.58270 0.23284 -2.50 0.014 *
## Year1999 -0.22245 0.26592 -0.84 0.405
## Year2000 -0.16002 0.21300 -0.75 0.454
## Year2001 -0.07427 0.20285 -0.37 0.715
## Year2002 0.00364 0.23270 0.02 0.988
## Year2003 -0.09651 0.28533 -0.34 0.736
## Year2004 0.11264 0.18816 0.60 0.551
## Year2005 -0.28315 0.21353 -1.33 0.187
## Year2006 -0.03969 0.21841 -0.18 0.856
```

```

## Year2007          -0.10595    0.28886   -0.37    0.714
## Year2008          -0.09125    0.22893   -0.40    0.691
## Year2009          -0.15210    0.24749   -0.61    0.540
## Year2010          -0.17985    0.19854   -0.91    0.367
## Year2011          -0.20605    0.19771   -1.04    0.299
## Year2012          -0.22065    0.20771   -1.06    0.290
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.329
## Multiple R-squared:  0.12,   Adjusted R-squared:  -0.00324
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 128 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.146  0.856  0.944  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      7.19e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 111.1  1          10.54
## Year              111.1 15          1.17

```

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
## -1.1834 -0.2295 0.0117 0.1955 0.8760
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2734 0.1607 7.92 1.2e-12 ***
## FirstAuthorFemale1 0.0776 0.0773 1.00 0.3172
## Year1998 -0.5619 0.2012 -2.79 0.0061 **
## Year1999 -0.1756 0.2355 -0.75 0.4573
## Year2000 -0.1866 0.1792 -1.04 0.2997
## Year2001 -0.0299 0.1609 -0.19 0.8531
## Year2002 0.0310 0.1914 0.16 0.8716
## Year2003 -0.0848 0.2426 -0.35 0.7272
## Year2004 0.0630 0.1409 0.45 0.6556
## Year2005 -0.2771 0.1740 -1.59 0.1140
## Year2006 -0.0246 0.1810 -0.14 0.8919
## Year2007 -0.0900 0.2466 -0.37 0.7157
```

```

## Year2008          -0.0455      0.1937   -0.23   0.8146
## Year2009          -0.1481      0.2200   -0.67   0.5022
## Year2010          -0.1546      0.1598   -0.97   0.3353
## Year2011          -0.2039      0.1599   -1.28   0.2046
## Year2012          -0.2040      0.1700   -1.20   0.2325
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.324
## Multiple R-squared:  0.11,   Adjusted R-squared:  -0.00702
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 125 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.153  0.855  0.954  0.886  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      7.19e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##  nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
##  trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 83.45  1          9.135
## Year            83.45 15          1.159

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ 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 -2.36e-01 -2.08e-16  2.05e-01  8.85e-01
##

```

```

## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3997    0.1880   7.45 1.5e-11 ***
## LastAuthorFemale1 -0.0974    0.0664  -1.47  0.1453
## Year1998          -0.6058    0.2274  -2.66  0.0088 **
## Year1999          -0.2734    0.2427  -1.13  0.2621
## Year2000          -0.1991    0.2078  -0.96  0.3400
## Year2001          -0.1562    0.1881  -0.83  0.4080
## Year2002          -0.0326    0.2258  -0.14  0.8856
## Year2003          -0.1094    0.3158  -0.35  0.7296
## Year2004           0.1117    0.1880   0.59  0.5535
## Year2005          -0.3194    0.2089  -1.53  0.1288
## Year2006          -0.0685    0.2218  -0.31  0.7578
## Year2007          -0.1085    0.2937  -0.37  0.7124
## Year2008          -0.1430    0.2303  -0.62  0.5358
## Year2009          -0.2071    0.2321  -0.89  0.3740
## Year2010          -0.2084    0.1973  -1.06  0.2930
## Year2011          -0.2346    0.1974  -1.19  0.2368
## Year2012          -0.2621    0.2013  -1.30  0.1954
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.323
## Multiple R-squared:  0.114, Adjusted R-squared:  -0.00178
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 126 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0749 0.8430 0.9520 0.8860 0.9860 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      7.19e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 139"
## [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
##    2    7    6    2    9    4   13    4    9    4    5    9    7    6   16
## 2011 2012
##    9   14
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    1    4    1    6    3   12    3    8    4    2    4    5    4   11
## 2011 2012
##    7   10
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    0    4    1    4    2    9    2    7    4    2    4    5    3   11
## 2011 2012
##    7    9
## [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: 5.40255415050821"

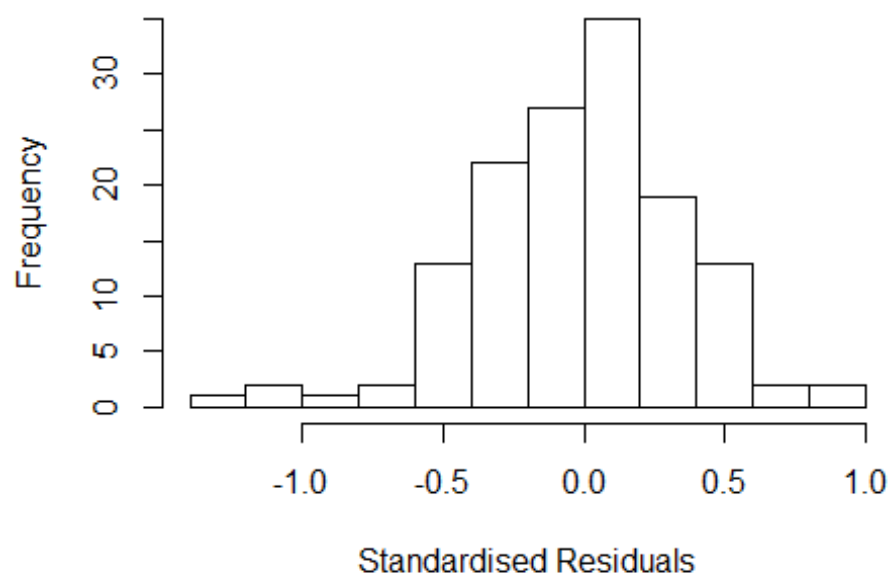
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 11, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.99069756244244"
## [1] "Male last author team size 2018 geometric mean: 6.42475583497184"

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

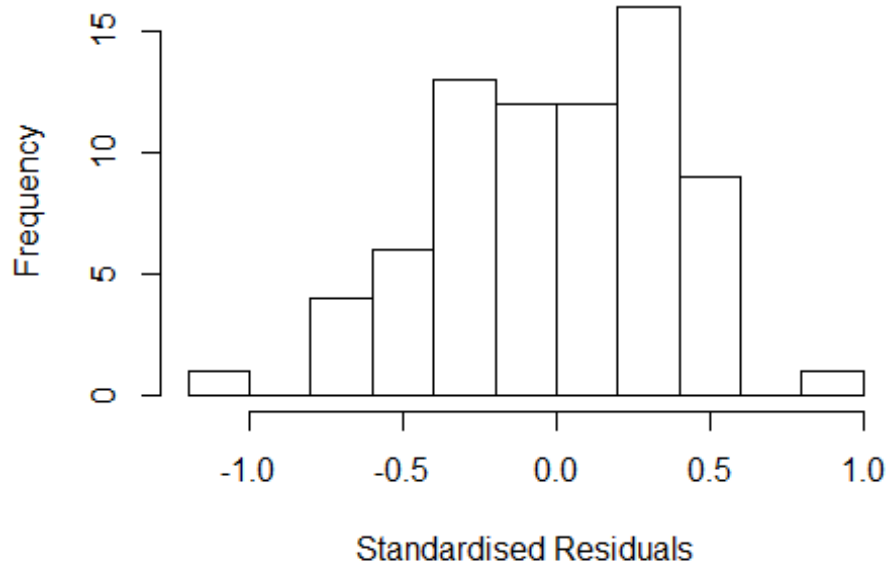
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2, 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  54.342  1          7.372
## LastAuthorFemale   5.069  1          2.252
## UniqueAuthors     29.229  4          1.525
## Year              1128.912 14          1.285
```


Residuals from first and last author and team size



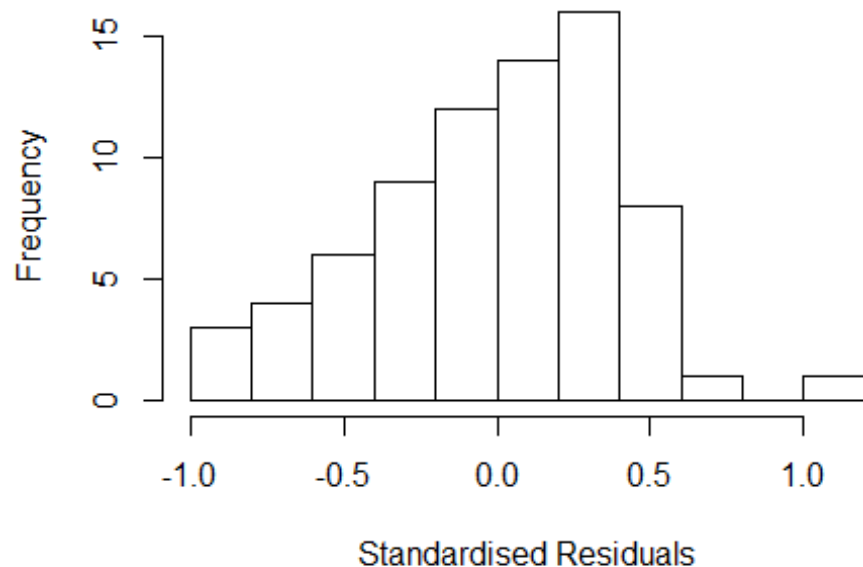
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.0340 -0.2876  0.0155  0.2676  0.9406
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0738     0.4283   2.51  0.0153 *
## FirstAuthorFemale1  0.1024     0.1263   0.81  0.4211
## LastAuthorFemale1  0.1812     0.1806   1.00  0.3201
## UniqueAuthors2     0.2519     0.3459   0.73  0.4698
## UniqueAuthors3     0.0660     0.4132   0.16  0.8738
## UniqueAuthors4     0.1262     0.3733   0.34  0.7366
## UniqueAuthors5     0.4386     0.3547   1.24  0.2216
## Year1999          -1.5124     0.1750 -8.64 1.1e-11 ***
## Year2000           0.1884     0.2618   0.72  0.4749
## Year2001           0.3308     0.5487   0.60  0.5492
```

```

## Year2002          -0.2332      0.2158    -1.08    0.2848
## Year2003          -0.8574      0.4545    -1.89    0.0647 .
## Year2004          -0.5810      0.2252    -2.58    0.0127 *
## Year2005          -0.2845      0.3426    -0.83    0.4100
## Year2006          -0.0183      0.2135    -0.09    0.9319
## Year2007          -0.2736      0.2476    -1.10    0.2743
## Year2008          -0.4752      0.2536    -1.87    0.0665 .
## Year2009          -0.5693      0.2356    -2.42    0.0192 *
## Year2010          -0.4784      0.3079    -1.55    0.1261
## Year2011          -0.7614      0.2878    -2.65    0.0107 *
## Year2012          -0.7144      0.2666    -2.68    0.0098 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.412
## Multiple R-squared:  0.431, Adjusted R-squared:  0.216
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 68 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.508  0.903  0.954  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      1.35e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 56.400 1          7.510
## LastAuthorFemale  1.696 1          1.302
## Year              93.980 14          1.176

```

Residuals from first and last author



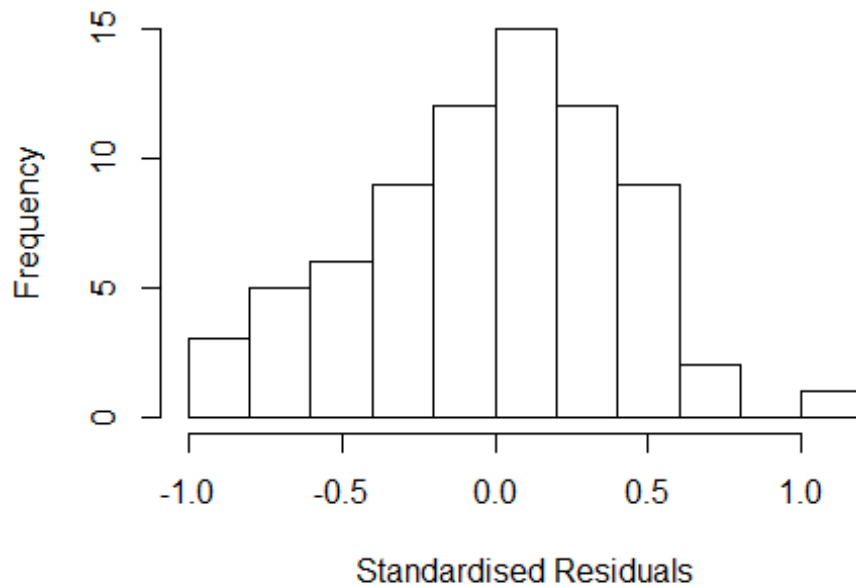
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9248 -0.3003 0.0335 0.2724 1.0623
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3823 0.1541 8.97 1.7e-12 ***
## FirstAuthorFemale1 0.0899 0.1288 0.70 0.4880
## LastAuthorFemale1 0.1167 0.1670 0.70 0.4875
## Year1999 -1.3823 0.1541 -8.97 1.7e-12 ***
## Year2000 0.0271 0.1977 0.14 0.8917
## Year2001 0.3047 0.3180 0.96 0.3419
## Year2002 -0.2447 0.2031 -1.20 0.2332
## Year2003 -0.8772 0.6413 -1.37 0.1767
## Year2004 -0.5986 0.2008 -2.98 0.0042 **
## Year2005 -0.3106 0.3025 -1.03 0.3090
## Year2006 0.1243 0.2014 0.62 0.5394
## Year2007 -0.2888 0.2538 -1.14 0.2598
```

```

## Year2008          -0.5484      0.2281   -2.40   0.0195 *
## Year2009          -0.5553      0.1851   -3.00   0.0040 **
## Year2010          -0.4574      0.2816   -1.62   0.1098
## Year2011          -0.7807      0.2830   -2.76   0.0078 **
## Year2012          -0.6956      0.2481   -2.80   0.0069 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.466
## Multiple R-squared:  0.341, Adjusted R-squared:  0.156
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 65 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.583  0.884  0.953  0.924  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      1.35e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 65.46 1      8.091
## Year              65.46 14      1.161

```

Residuals from first author



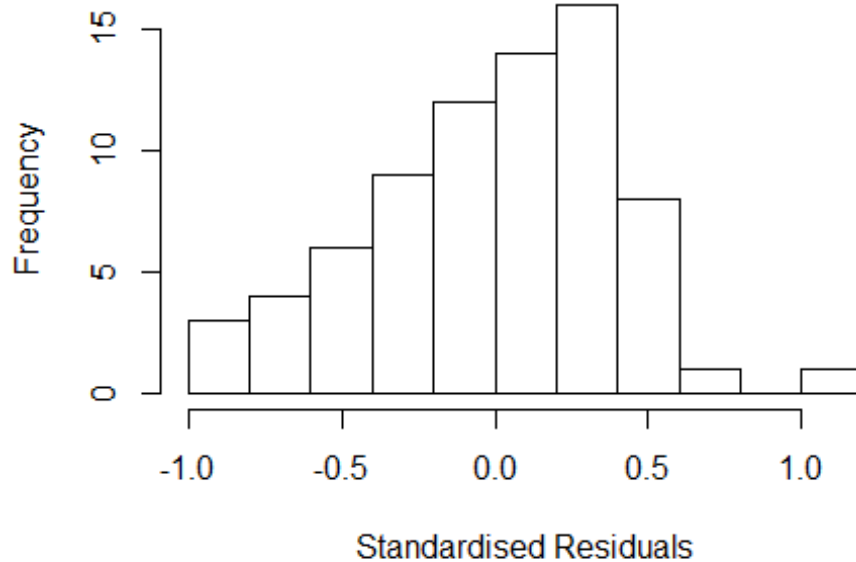
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.9486 -0.3217  0.0393  0.2770  1.0244
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.380      0.158   8.74 3.6e-12 ***
## FirstAuthorFemale1  0.104      0.139   0.75  0.4581
## Year1999         -1.380      0.158  -8.74 3.6e-12 ***
## Year2000           0.060      0.200   0.30  0.7646
## Year2001           0.307      0.324   0.95  0.3472
## Year2002          -0.237      0.212  -1.12  0.2688
## Year2003          -0.882      0.672  -1.31  0.1943
## Year2004          -0.575      0.203  -2.84  0.0062 **
## Year2005          -0.318      0.310  -1.03  0.3089
## Year2006           0.112      0.209   0.54  0.5928
## Year2007          -0.275      0.242  -1.13  0.2620
## Year2008          -0.556      0.233  -2.39  0.0202 *
```

```

## Year2009          -0.562      0.190   -2.96   0.0044 **
## Year2010          -0.431      0.281   -1.54   0.1299
## Year2011          -0.766      0.290   -2.64   0.0106 *
## Year2012          -0.665      0.254   -2.62   0.0111 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.456
## Multiple R-squared:  0.335, Adjusted R-squared:  0.163
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.593  0.887  0.951  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      1.35e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.816 1          1.348
## Year            1.816 14          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
## -0.9723 -0.2535  0.0395  0.2793  1.1047
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.40094    0.13639   10.27  1.1e-14 ***
## LastAuthorFemale1 0.13900    0.17699    0.79  0.4354
## Year1999       -1.40094    0.13639  -10.27  1.1e-14 ***
## Year2000         0.00257    0.18624    0.01  0.9890
## Year2001         0.28606    0.31243    0.92  0.3637
## Year2002        -0.24736    0.18886   -1.31  0.1954
## Year2003        -0.85094    0.56235   -1.51  0.1357
## Year2004        -0.57224    0.17097   -3.35  0.0014 **
## Year2005        -0.26304    0.26664   -0.99  0.3280
## Year2006         0.19556    0.13757    1.42  0.1605
## Year2007        -0.27205    0.23821   -1.14  0.2581
## Year2008        -0.50858    0.19025   -2.67  0.0097 **
```

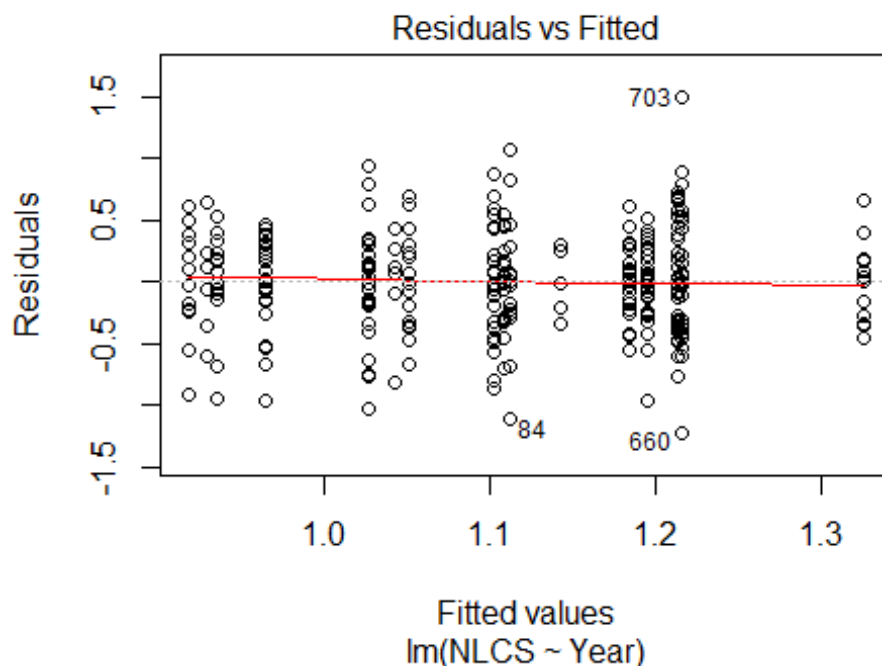
```

## Year2009          -0.51443      0.15230    -3.38    0.0013 **
## Year2010          -0.42861      0.28254    -1.52    0.1347
## Year2011          -0.74439      0.26571    -2.80    0.0069 **
## Year2012          -0.66599      0.21689    -3.07    0.0032 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.453
## Multiple R-squared:  0.337, Adjusted R-squared:  0.166
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 69 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.532  0.884  0.958  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           1.35e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01                5.00e-01
##   nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500          50          2          1          1000          200
##   trace.lev      mts      compute.rd
##           0          1000          0
##           psi          subsampling          cov
##           "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 74"
## [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
##    6   25   14   22   27   22   19   27   24   35   34   41   48   37   43
## 2011 2012
##   53   41
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    6    6   14   12    5   12   16   18   14   23   30   31   22   27
## 2011 2012
##   34   30
##

```



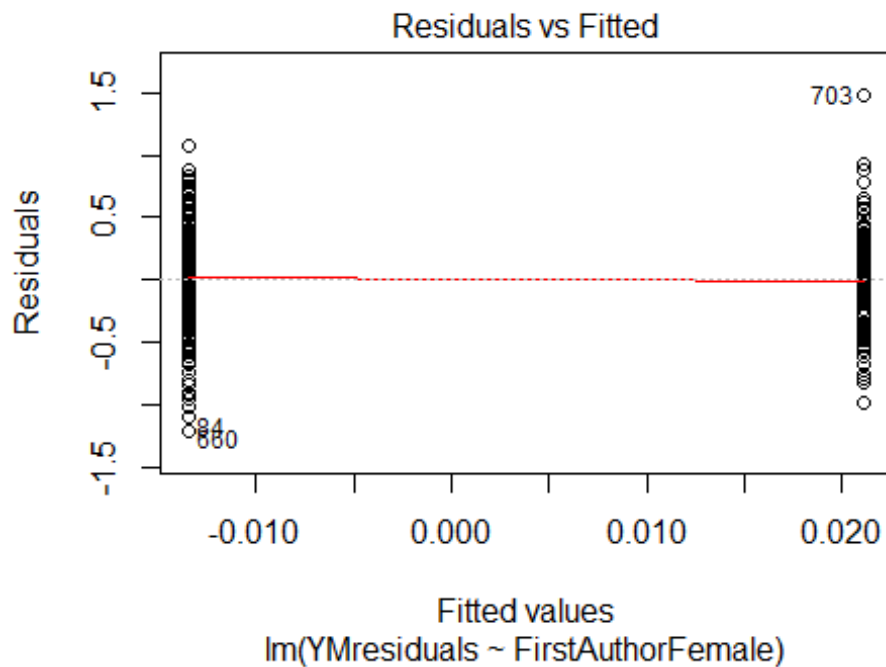
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    6    6   10    9    4    8   11   12   14   20   22   22   17   23
## 2011 2012
##   30   25
## [1] "Heteroscedasticity checks, confirming that there are problems with
##      these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 19, df = 15, p-value = 0.2
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.85, df = 1, p-value = 0.4
##
## [1] "Female first author team size 2018 geometric mean: 5.61721847212363"
## [1] "Male first author team size 2018 geometric mean: 4.64890738418337"
##
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
##
## Wilcoxon rank sum test with continuity correction
##
```

```
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 200, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6.98749725153495"
## [1] "Male last author team size 2018 geometric mean: 4.41422485172005"

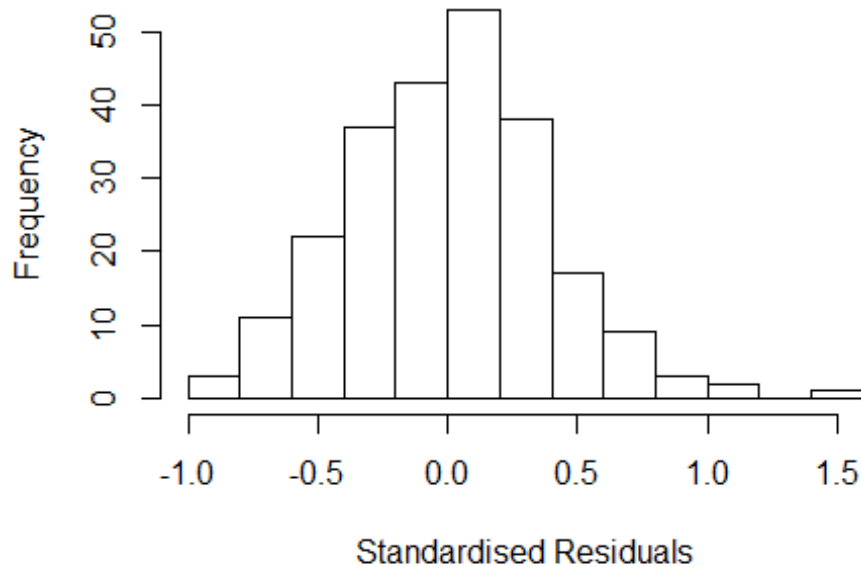
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 210, p-value = 0.04
## 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 | 1.380 | 1 | 1.175 |
| LastAuthorFemale | 1.260 | 1 | 1.123 |
| UniqueAuthors | 4.347 | 4 | 1.202 |
| Year | 5.195 | 15 | 1.056 |

Residuals from first and last author and team size



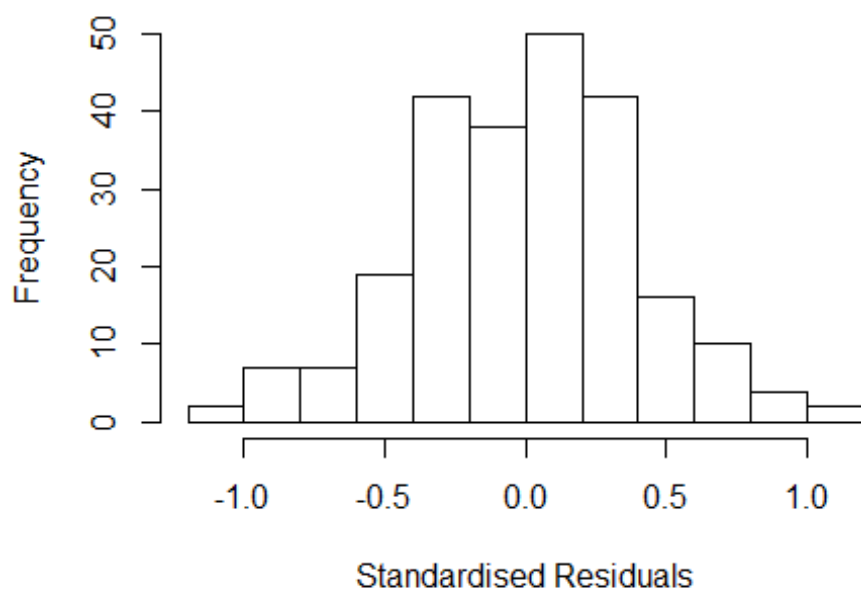
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9483 -0.2557 0.0157 0.2281 1.4010
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6781 0.2152 3.15 0.0019 **
## FirstAuthorFemale1 -0.0176 0.0572 -0.31 0.7582
## LastAuthorFemale1 0.0546 0.0563 0.97 0.3329
## UniqueAuthors2 0.3321 0.1531 2.17 0.0311 *
## UniqueAuthors3 0.3243 0.1491 2.18 0.0307 *
## UniqueAuthors4 0.3314 0.1504 2.20 0.0286 *
## UniqueAuthors5 0.4338 0.1432 3.03 0.0028 **
## Year1998 0.0684 0.2449 0.28 0.7802
## Year1999 0.0979 0.2690 0.36 0.7163
## Year2000 0.3155 0.2359 1.34 0.1825
```

```

## Year2001          0.2645      0.2830      0.93      0.3509
## Year2002          0.0323      0.2229      0.14      0.8849
## Year2003          0.0742      0.2212      0.34      0.7377
## Year2004         -0.1013      0.2328     -0.44      0.6640
## Year2005         -0.0693      0.2283     -0.30      0.7619
## Year2006          0.1387      0.2263      0.61      0.5407
## Year2007          0.0910      0.2142      0.42      0.6715
## Year2008          0.0491      0.2161      0.23      0.8203
## Year2009          0.1527      0.2175      0.70      0.4833
## Year2010         -0.0919      0.2117     -0.43      0.6647
## Year2011          0.1275      0.2191      0.58      0.5614
## Year2012          0.0590      0.2252      0.26      0.7936
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.387
## Multiple R-squared:  0.143, Adjusted R-squared:  0.0599
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 27 weights are ~= 1. The remaining 212 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.161  0.888  0.949  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      4.18e-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 1.356 1      1.164
## LastAuthorFemale  1.167 1      1.080
## Year              1.450 15      1.012

```

Residuals from first and last author



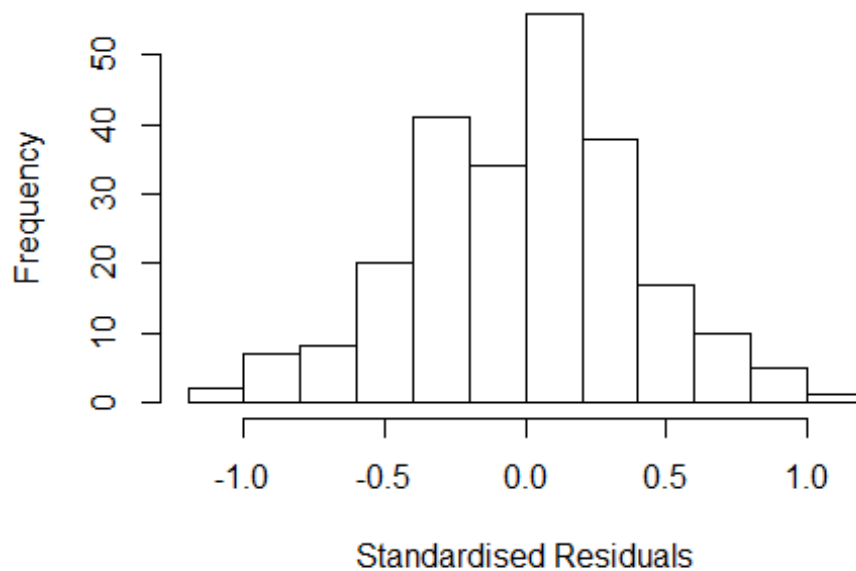
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1171 -0.2755 0.0115 0.2491 1.0773
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.92882 0.18242 5.09 7.6e-07 ***
## FirstAuthorFemale1 -0.00403 0.05951 -0.07 0.95
## LastAuthorFemale1 0.06430 0.05964 1.08 0.28
## Year1998 0.16172 0.23784 0.68 0.50
## Year1999 0.17088 0.28642 0.60 0.55
## Year2000 0.33239 0.20938 1.59 0.11
## Year2001 0.21757 0.23812 0.91 0.36
## Year2002 0.12359 0.21091 0.59 0.56
## Year2003 0.10671 0.20729 0.51 0.61
## Year2004 -0.02273 0.21376 -0.11 0.92
## Year2005 0.02902 0.20749 0.14 0.89
## Year2006 0.20591 0.22166 0.93 0.35
```

```

## Year2007          0.23581    0.19432    1.21    0.23
## Year2008          0.15018    0.20296    0.74    0.46
## Year2009          0.25672    0.19916    1.29    0.20
## Year2010          0.00126    0.20212    0.01    1.00
## Year2011          0.24628    0.20226    1.22    0.22
## Year2012          0.18830    0.20750    0.91    0.37
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.4
## Multiple R-squared:  0.0603, Adjusted R-squared:  -0.012
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 215 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.415  0.886  0.950  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      4.18e-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 1.323 1      1.150
## Year              1.323 15      1.009

```

Residuals from first author



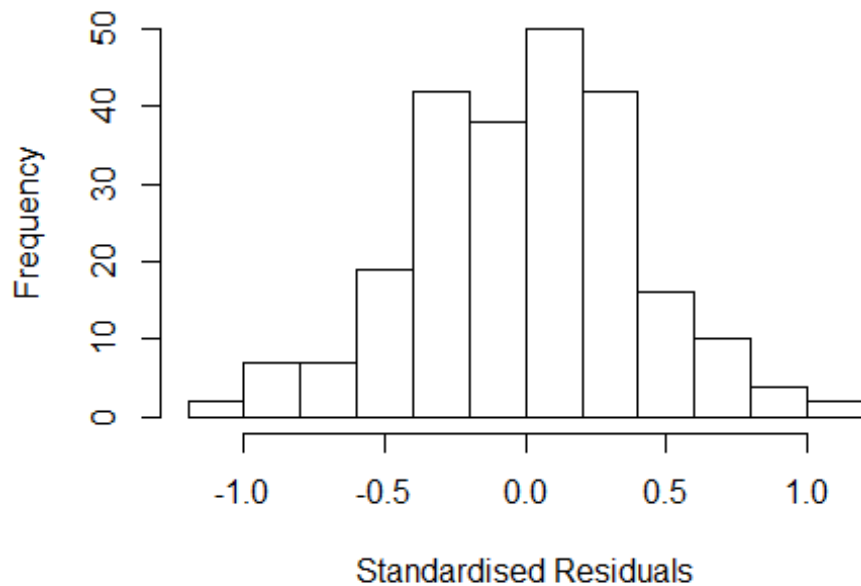
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1310 -0.2594 0.0354 0.2476 1.0745
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.92490 0.18038 5.13 6.4e-07 ***
## FirstAuthorFemale1 0.00904 0.05863 0.15 0.88
## Year1998 0.16606 0.23612 0.70 0.48
## Year1999 0.17762 0.28527 0.62 0.53
## Year2000 0.33761 0.20721 1.63 0.10
## Year2001 0.22150 0.23662 0.94 0.35
## Year2002 0.12971 0.20933 0.62 0.54
## Year2003 0.13626 0.20438 0.67 0.51
## Year2004 -0.00230 0.21093 -0.01 0.99
## Year2005 0.04838 0.20283 0.24 0.81
## Year2006 0.21765 0.22004 0.99 0.32
## Year2007 0.25265 0.19139 1.32 0.19
```

```

## Year2008          0.17097    0.20118    0.85    0.40
## Year2009          0.27241    0.19812    1.37    0.17
## Year2010          0.02042    0.19883    0.10    0.92
## Year2011          0.26327    0.20061    1.31    0.19
## Year2012          0.20610    0.20533    1.00    0.32
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.398
## Multiple R-squared:  0.0562, Adjusted R-squared:  -0.0118
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 221 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.400  0.886  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      4.18e-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 1.128 1          1.062
## Year            1.128 15          1.004

```


Residuals from last author



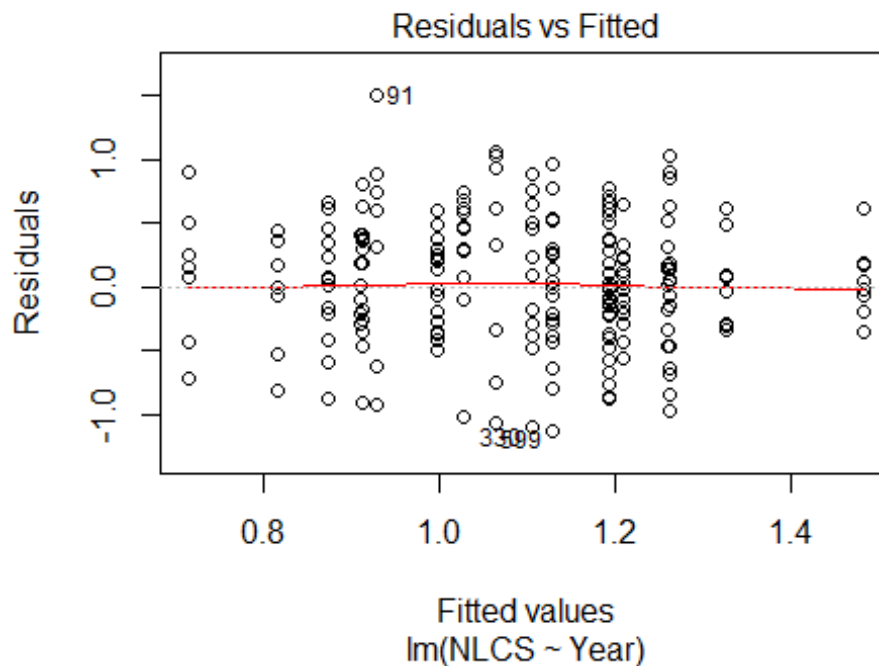
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.11573 -0.27433 0.00815 0.25133 1.08028
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.92759 0.18179 5.10 7.2e-07 ***
## LastAuthorFemale1 0.06323 0.05874 1.08 0.28
## Year1998 0.16355 0.23730 0.69 0.49
## Year1999 0.16913 0.27978 0.60 0.55
## Year2000 0.33198 0.20914 1.59 0.11
## Year2001 0.21882 0.23778 0.92 0.36
## Year2002 0.12279 0.20937 0.59 0.56
## Year2003 0.10726 0.20737 0.52 0.61
## Year2004 -0.02233 0.21373 -0.10 0.92
## Year2005 0.03027 0.20739 0.15 0.88
## Year2006 0.20593 0.22121 0.93 0.35
## Year2007 0.23625 0.19390 1.22 0.22
```

```

## Year2008          0.14978      0.20200      0.74      0.46
## Year2009          0.25703      0.19886      1.29      0.20
## Year2010          0.00159      0.20167      0.01      0.99
## Year2011          0.24538      0.20111      1.22      0.22
## Year2012          0.18814      0.20707      0.91      0.36
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.398
## Multiple R-squared:  0.0603, Adjusted R-squared:  -0.00745
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 215 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.411  0.883  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      4.18e-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: 239"
## [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
##   36   28   17   18   28   29   22   29   30   15   18   28   30   31   40
## 2011 2012
##   49   36
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    8   10    6   10   14    9   13   10   16    8    9   13   15   15   17
## 2011 2012
##   31   23

```

```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    7    5    7    9    9    7    7    9    8    9   10   10   12   12
## 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 = 41, df = 16, p-value = 5e-04
```



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

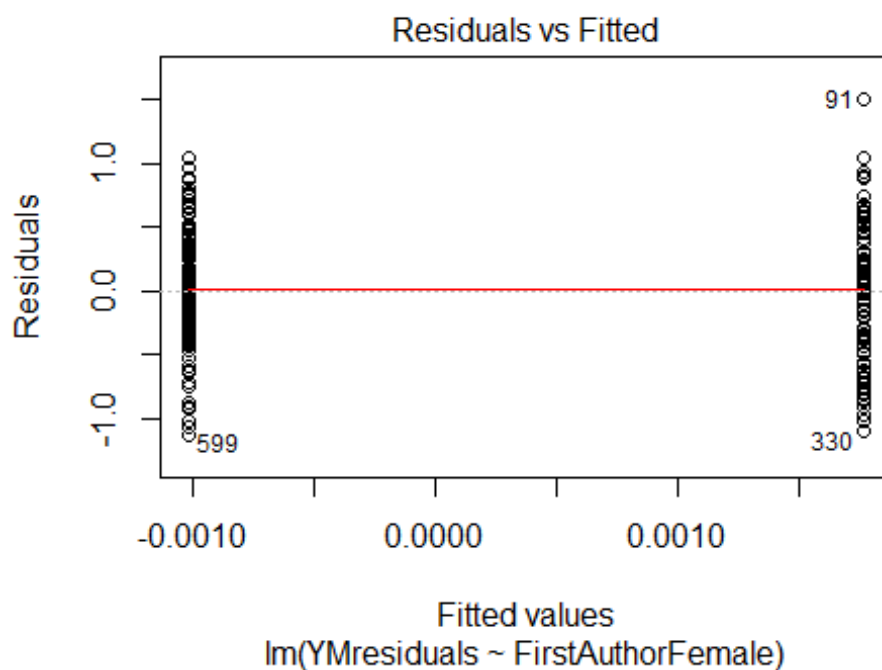
## [1] "Female first author team size 2018 geometric mean: 3.78967058674193"
## [1] "Male first author team size 2018 geometric mean: 4.16016764610381"

## Warning in wilcox.test.default(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 = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.7214587917035"
## [1] "Male last author team size 2018 geometric mean: 4.03296864170517"

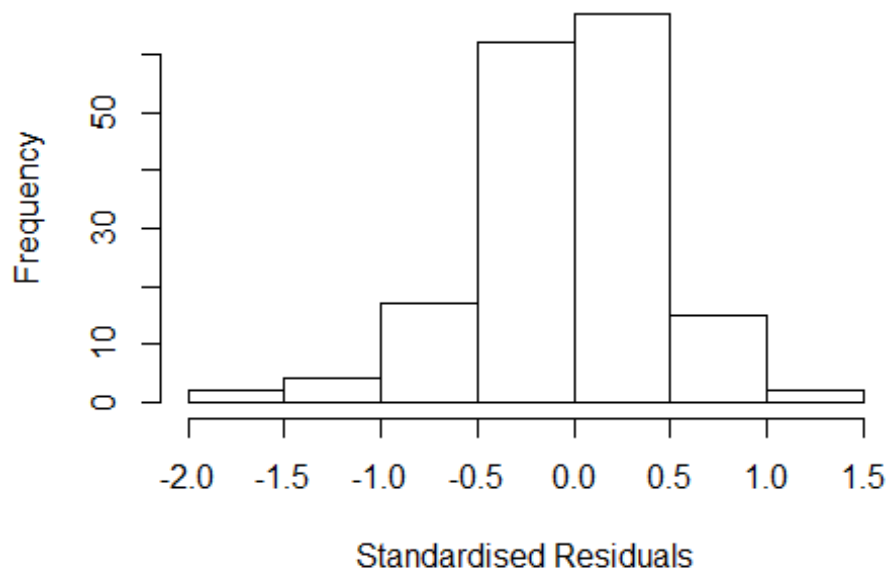
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 46, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
## Year as factors"
##
```

| | GVI | F | Df | GVI ^{1/(2*Df)} |
|-------------------|--------|----|----|-------------------------|
| FirstAuthorFemale | 3.255 | 1 | | 1.804 |
| LastAuthorFemale | 2.677 | 1 | | 1.636 |
| UniqueAuthors | 27.110 | 4 | | 1.511 |
| Year | 53.755 | 16 | | 1.133 |

Residuals from first and last author and team size



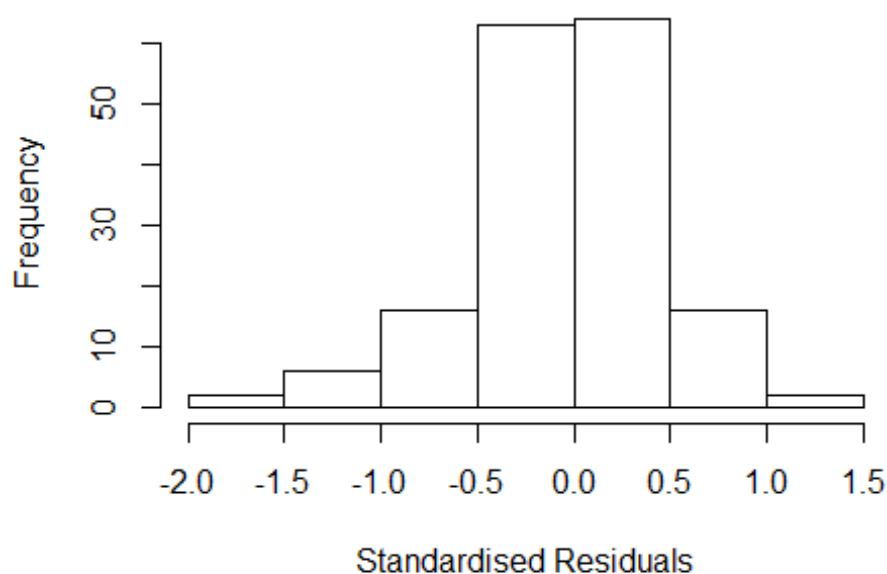
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.88e+00 -2.86e-01 -1.71e-05 2.97e-01 1.21e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6360 0.3135 2.03 0.044 *
## FirstAuthorFemale1 0.1936 0.1096 1.77 0.079 .
## LastAuthorFemale1 0.1049 0.1213 0.86 0.388
## UniqueAuthors2 0.2986 0.3169 0.94 0.348
## UniqueAuthors3 0.3310 0.2438 1.36 0.177
## UniqueAuthors4 0.4187 0.2329 1.80 0.074 .
## UniqueAuthors5 0.2674 0.2392 1.12 0.265
## Year1997 0.6359 0.4293 1.48 0.141
## Year1998 -0.0720 0.2235 -0.32 0.748
## Year1999 0.7485 0.3337 2.24 0.026 *
```

```

## Year2000          0.4334      0.2096      2.07      0.040 *
## Year2001          0.4679      0.2017      2.32      0.022 *
## Year2002          0.2609      0.1894      1.38      0.170
## Year2003          0.2598      0.2071      1.25      0.212
## Year2004         -0.0826      0.2299     -0.36      0.720
## Year2005         -0.2681      0.2415     -1.11      0.269
## Year2006          0.2212      0.2213      1.00      0.319
## Year2007          0.1191      0.2967      0.40      0.689
## Year2008          0.0954      0.2091      0.46      0.649
## Year2009         -0.0754      0.3004     -0.25      0.802
## Year2010         -0.0162      0.2147     -0.08      0.940
## Year2011          0.2124      0.1969      1.08      0.283
## Year2012          0.0742      0.2186      0.34      0.735
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.436
## Multiple R-squared:  0.243, Adjusted R-squared:  0.129
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 153 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0244 0.8630 0.9500 0.8930 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      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 1.989 1      1.410
## LastAuthorFemale  1.809 1      1.345
## Year              3.591 16      1.041

```

Residuals from first and last author



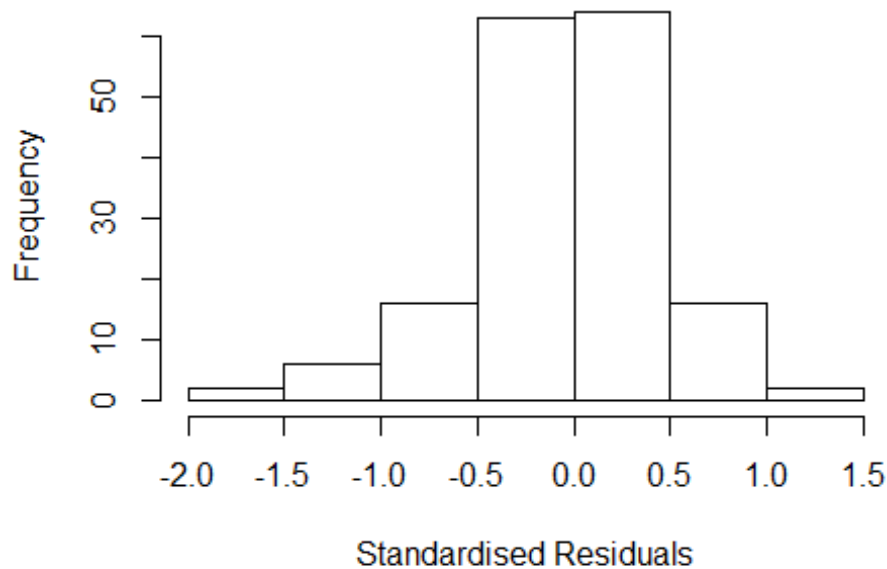
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.6854 -0.2728 -0.0124 0.2892 1.2060
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9139 0.1802 5.07 1.1e-06 ***
## FirstAuthorFemale1 0.1406 0.0898 1.57 0.119
## LastAuthorFemale1 0.0240 0.0976 0.25 0.806
## Year1997 0.6573 0.3656 1.80 0.074 .
## Year1998 -0.0124 0.2357 -0.05 0.958
## Year1999 0.6069 0.4234 1.43 0.154
## Year2000 0.4978 0.2257 2.21 0.029 *
## Year2001 0.4942 0.2106 2.35 0.020 *
## Year2002 0.3538 0.2023 1.75 0.082 .
## Year2003 0.3807 0.2175 1.75 0.082 .
## Year2004 0.0124 0.2408 0.05 0.959
## Year2005 -0.2038 0.2562 -0.80 0.428
```

```

## Year2006          0.2832      0.2334      1.21      0.227
## Year2007          0.2105      0.3396      0.62      0.536
## Year2008          0.1672      0.2173      0.77      0.443
## Year2009          0.0470      0.2896      0.16      0.871
## Year2010         -0.0164      0.2130     -0.08      0.939
## Year2011          0.2751      0.2212      1.24      0.216
## Year2012          0.1596      0.2291      0.70      0.487
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.448
## Multiple R-squared:  0.196, Adjusted R-squared:  0.0993
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 153 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.126  0.874  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      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 1.986 1      1.409
## Year              1.986 16      1.022

```


Residuals from first author



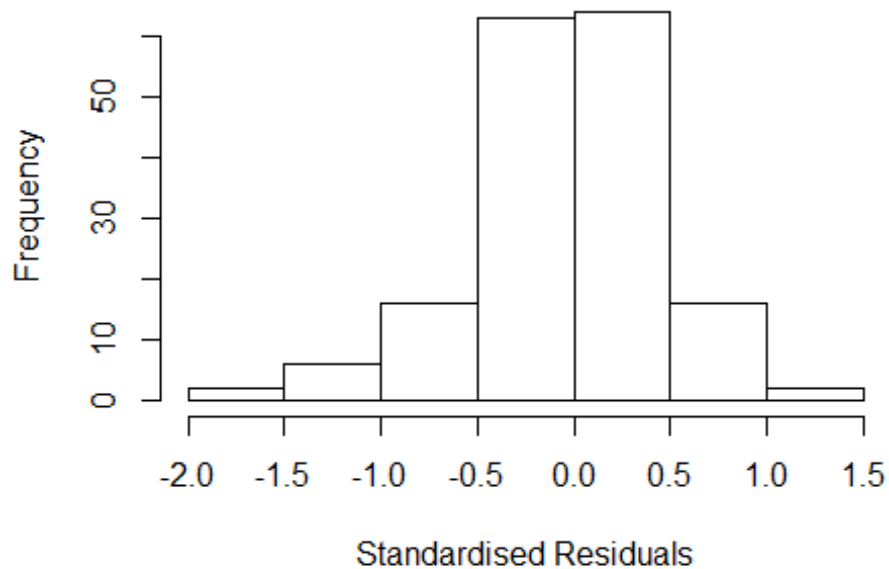
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.6848 -0.2841 -0.0118 0.2918 1.1970
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.91497 0.17859 5.12 9e-07 ***
## FirstAuthorFemale1 0.14229 0.08984 1.58 0.115
## Year1997 0.65958 0.36565 1.80 0.073 .
## Year1998 -0.00388 0.22248 -0.02 0.986
## Year1999 0.62755 0.39376 1.59 0.113
## Year2000 0.50422 0.21808 2.31 0.022 *
## Year2001 0.50111 0.19964 2.51 0.013 *
## Year2002 0.35531 0.19784 1.80 0.075 .
## Year2003 0.37851 0.21631 1.75 0.082 .
## Year2004 0.01447 0.23689 0.06 0.951
## Year2005 -0.20224 0.25243 -0.80 0.424
## Year2006 0.28944 0.22192 1.30 0.194
```

```

## Year2007          0.21629    0.34176    0.63    0.528
## Year2008          0.16493    0.21593    0.76    0.446
## Year2009          0.05503    0.27274    0.20    0.840
## Year2010         -0.01651    0.21127   -0.08    0.938
## Year2011          0.27977    0.21602    1.30    0.197
## Year2012          0.16501    0.22331    0.74    0.461
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.444
## Multiple R-squared:  0.198, Adjusted R-squared:  0.108
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 153 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.118  0.872  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      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 1.772 1          1.331
## Year            1.772 16          1.018

```

Residuals from last author



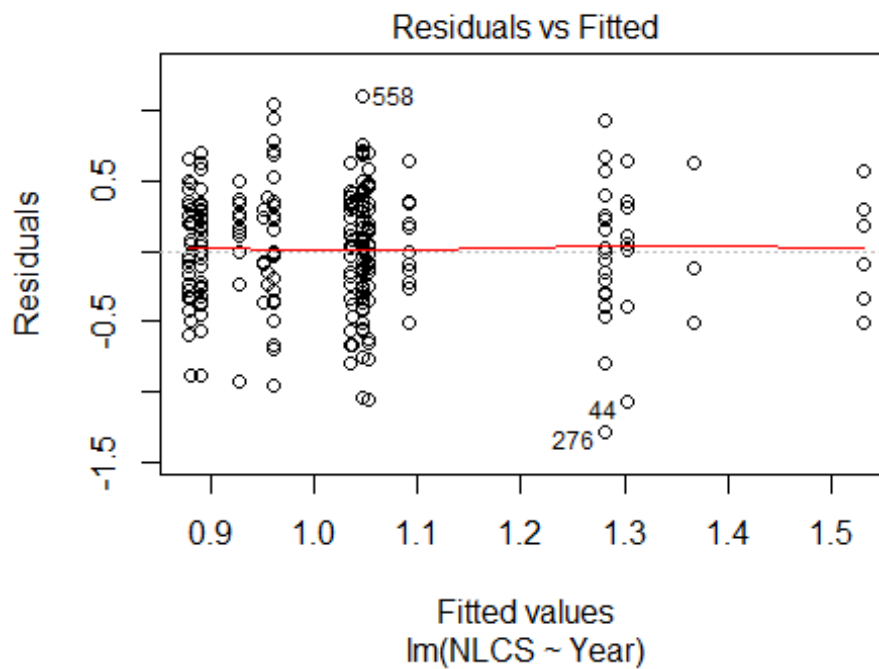
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.56802 -0.29734 -0.00187 0.28091 1.22362
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.95120 0.17549 5.42 2.3e-07 ***
## LastAuthorFemale1 0.03173 0.10202 0.31 0.7562
## Year1997 0.65857 0.35592 1.85 0.0662 .
## Year1998 0.00284 0.21283 0.01 0.9894
## Year1999 0.58509 0.43955 1.33 0.1852
## Year2000 0.53705 0.20231 2.65 0.0088 **
## Year2001 0.51167 0.20691 2.47 0.0145 *
## Year2002 0.39598 0.19345 2.05 0.0424 *
## Year2003 0.46174 0.20654 2.24 0.0268 *
## Year2004 0.02041 0.23583 0.09 0.9312
## Year2005 -0.10809 0.24388 -0.44 0.6583
## Year2006 0.34940 0.22139 1.58 0.1166
```

```

## Year2007      0.24332      0.33730      0.72      0.4718
## Year2008      0.20975      0.20340      1.03      0.3041
## Year2009      0.12218      0.28687      0.43      0.6708
## Year2010     -0.01112      0.20753     -0.05      0.9574
## Year2011      0.33617      0.20810      1.62      0.1083
## Year2012      0.22996      0.21623      1.06      0.2892
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.45
## Multiple R-squared:  0.173, Adjusted R-squared:  0.0801
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 154 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.199  0.883  0.956  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      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 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
##   13   13   11   16   14   31   19   23   21   27   39   36   46   48   41
## 2011 2012
##   54   45
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    7    3    8    3   11   13   12    9   14   19   20   20   30   27
## 2011 2012

```

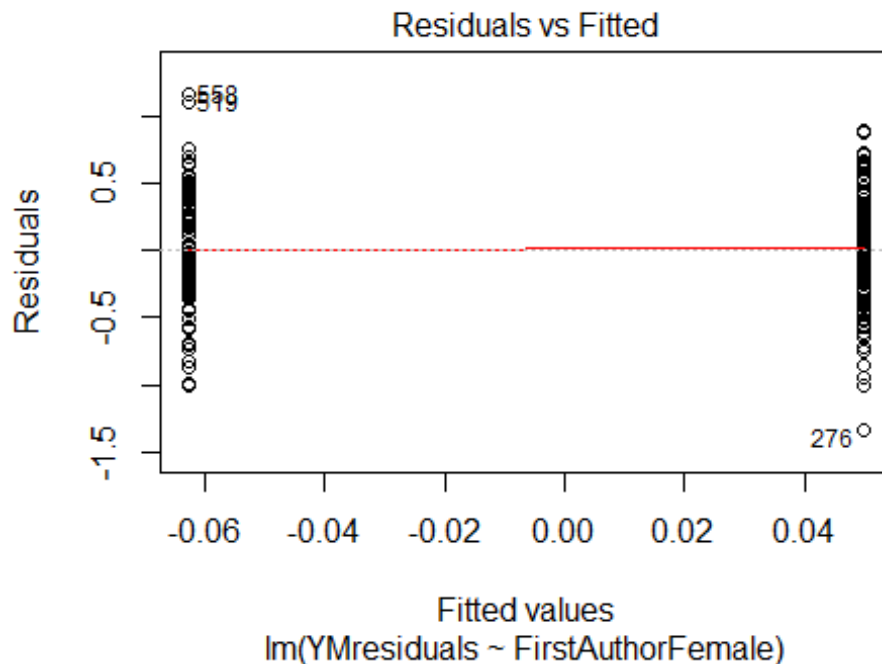
```
## 36 32
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3 7 2 7 3 10 10 10 8 8 14 13 17 23 23
## 2011 2012
## 35 29
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 15, df = 16, p-value = 0.5
```



```
##
## 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: 5.97803652483267"
## [1] "Male first author team size 2018 geometric mean: 5.17747565258689"
## Warning in wilcox.test.default(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 = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.27355472045695"
## [1] "Male last author team size 2018 geometric mean: 6.28486800745263"

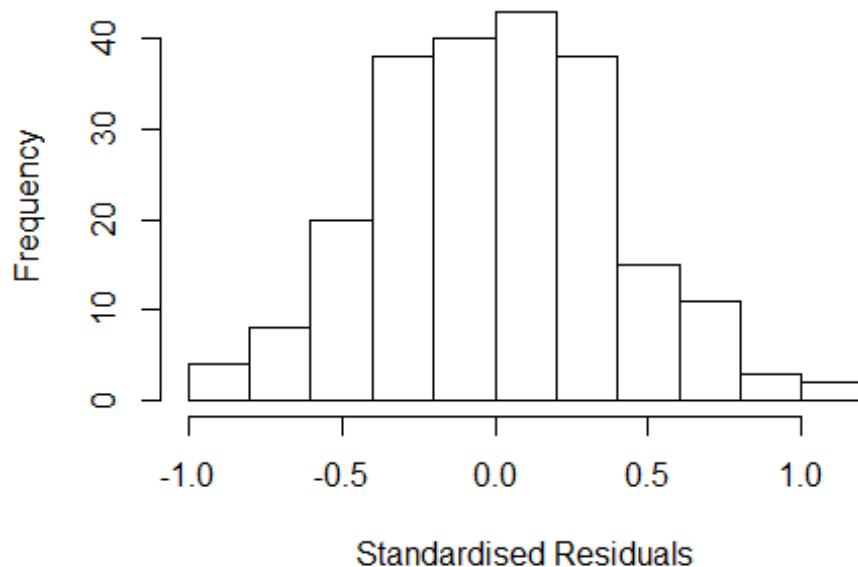
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## 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] "Regression 1: First author gender, last author gender, team size,
## Year as factors"
##
```

| | GVIF | Df | GVIF ^{1/(2*Df)} |
|-------------------|-------|----|--------------------------|
| FirstAuthorFemale | 1.794 | 1 | 1.339 |
| LastAuthorFemale | 2.067 | 1 | 1.438 |
| UniqueAuthors | 3.285 | 4 | 1.160 |
| Year | 6.899 | 16 | 1.062 |

Residuals from first and last author and team size



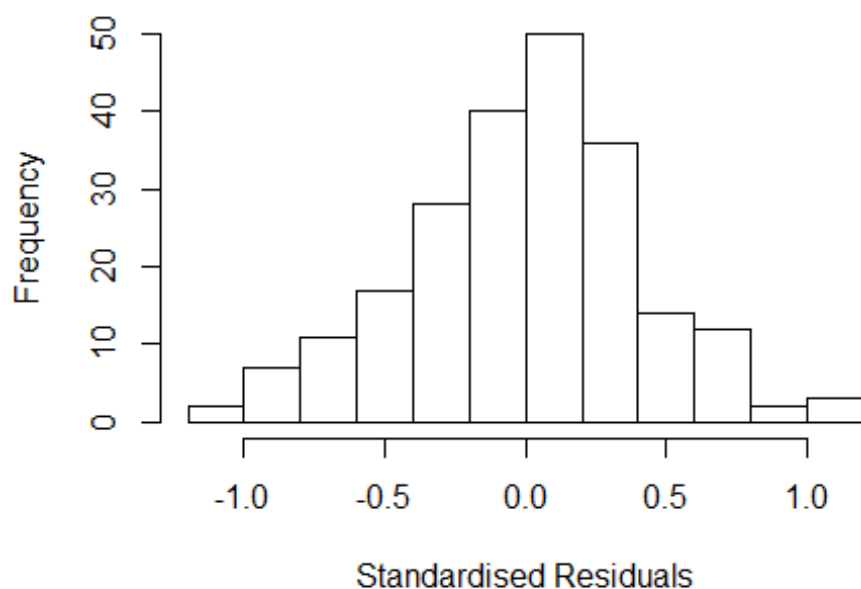
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.98273 -0.25706 0.00452 0.24132 1.14695
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5519 0.1673 3.30 0.0012 **
## FirstAuthorFemale1 0.0698 0.0651 1.07 0.2847
## LastAuthorFemale1 0.1965 0.0699 2.81 0.0054 **
## UniqueAuthors2 0.4575 0.1921 2.38 0.0182 *
## UniqueAuthors3 0.4331 0.1665 2.60 0.0100 **
## UniqueAuthors4 0.6023 0.1843 3.27 0.0013 **
## UniqueAuthors5 0.6306 0.1593 3.96 0.0001 ***
## Year1997 0.2833 0.2135 1.33 0.1860
## Year1998 0.3897 0.3237 1.20 0.2301
## Year1999 0.1748 0.1986 0.88 0.3798
```

```

## Year2000          -0.2461      0.2509   -0.98   0.3279
## Year2001          -0.2771      0.1944   -1.42   0.1557
## Year2002          -0.1283      0.2036   -0.63   0.5293
## Year2003          -0.1757      0.2287   -0.77   0.4430
## Year2004          -0.0735      0.2240   -0.33   0.7431
## Year2005          -0.0901      0.2118   -0.43   0.6710
## Year2006          -0.1280      0.1769   -0.72   0.4702
## Year2007           0.0132      0.2064    0.06   0.9491
## Year2008          -0.3304      0.1868   -1.77   0.0784 .
## Year2009          -0.1543      0.1787   -0.86   0.3888
## Year2010          -0.3031      0.2006   -1.51   0.1323
## Year2011          -0.1965      0.1791   -1.10   0.2740
## Year2012          -0.2696      0.1721   -1.57   0.1189
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.389
## Multiple R-squared:  0.264, Adjusted R-squared:  0.183
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 198 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.364  0.862  0.955  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      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 1.870 1      1.367
## LastAuthorFemale  1.822 1      1.350
## Year              2.362 16      1.027

```


Residuals from first and last author



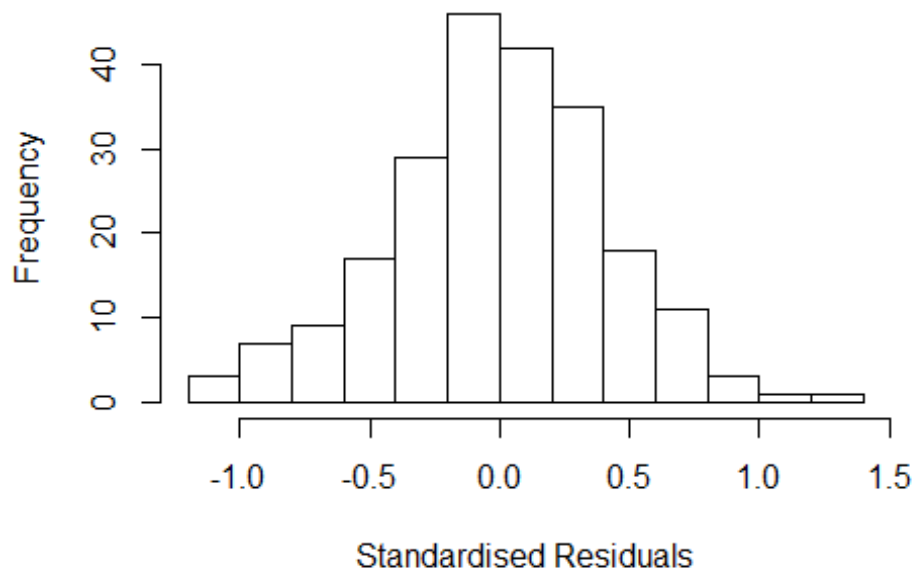
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0660 -0.2301 0.0231 0.2645 1.1935
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.90447 0.16756 5.40 1.9e-07 ***
## FirstAuthorFemale1 0.09744 0.06993 1.39 0.165
## LastAuthorFemale1 0.16278 0.07204 2.26 0.025 *
## Year1997 0.53775 0.21662 2.48 0.014 *
## Year1998 0.34170 0.97762 0.35 0.727
## Year1999 0.39953 0.24250 1.65 0.101
## Year2000 -0.01777 0.25909 -0.07 0.945
## Year2001 -0.08353 0.22158 -0.38 0.707
## Year2002 0.10835 0.22610 0.48 0.632
## Year2003 0.00874 0.24190 0.04 0.971
## Year2004 0.10895 0.22087 0.49 0.622
## Year2005 0.07551 0.20252 0.37 0.710
```

```

## Year2006          0.07234      0.19660      0.37      0.713
## Year2007          0.25863      0.21356      1.21      0.227
## Year2008         -0.12106      0.18462     -0.66      0.513
## Year2009          0.02322      0.18679      0.12      0.901
## Year2010         -0.09995      0.20926     -0.48      0.633
## Year2011          0.02630      0.18352      0.14      0.886
## Year2012         -0.07183      0.18290     -0.39      0.695
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.404
## Multiple R-squared:  0.159, Adjusted R-squared:  0.0845
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 197 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.363  0.859  0.950  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      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 1.472 1          1.213
## Year              1.472 16          1.012

```

Residuals from first author



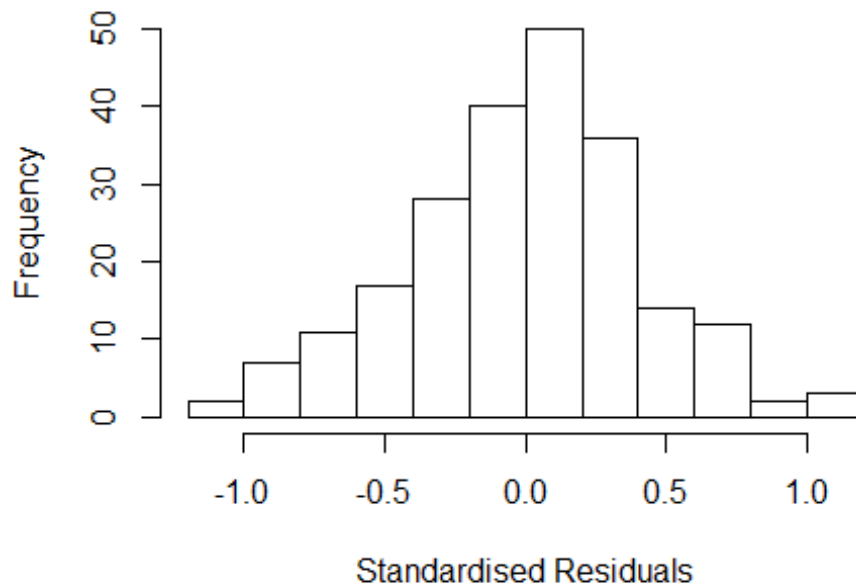
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.11139 -0.23624 0.00231 0.27291 1.21229
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.942254 0.139112 6.77 1.3e-10 ***
## FirstAuthorFemale1 0.144190 0.063531 2.27 0.024 *
## Year1997 0.486575 0.205586 2.37 0.019 *
## Year1998 0.338556 0.660914 0.51 0.609
## Year1999 0.407132 0.223716 1.82 0.070 .
## Year2000 0.003817 0.219677 0.02 0.986
## Year2001 -0.141401 0.186016 -0.76 0.448
## Year2002 0.133399 0.206050 0.65 0.518
## Year2003 -0.072993 0.215034 -0.34 0.735
## Year2004 0.096898 0.200816 0.48 0.630
## Year2005 0.073667 0.184932 0.40 0.691
## Year2006 0.028983 0.173352 0.17 0.867
```

```

## Year2007          0.235738    0.192791    1.22    0.223
## Year2008          -0.095702    0.166296   -0.58    0.566
## Year2009          -0.000594    0.162234    0.00    0.997
## Year2010          -0.107914    0.189568   -0.57    0.570
## Year2011          -0.009541    0.153744   -0.06    0.951
## Year2012          -0.077894    0.156918   -0.50    0.620
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.408
## Multiple R-squared:  0.139, Adjusted R-squared:  0.0668
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 199 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.357  0.857  0.951  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      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 1.505 1          1.227
## Year            1.505 16          1.013

```

Residuals from last author



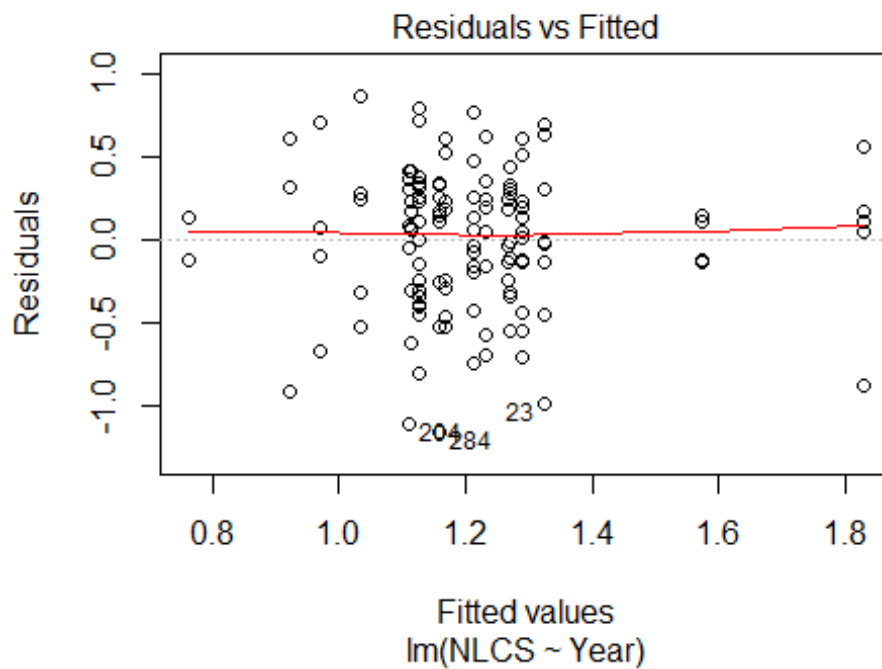
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1245 -0.2556 0.0246 0.2550 1.1310
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9256 0.1509 6.13 4.4e-09 ***
## LastAuthorFemale1 0.1945 0.0644 3.02 0.0028 **
## Year1997 0.5804 0.1952 2.97 0.0033 **
## Year1998 0.4021 1.0693 0.38 0.7073
## Year1999 0.4369 0.2159 2.02 0.0443 *
## Year2000 -0.0503 0.2527 -0.20 0.8423
## Year2001 -0.0378 0.2088 -0.18 0.8566
## Year2002 0.1075 0.2124 0.51 0.6133
## Year2003 0.0794 0.2159 0.37 0.7136
## Year2004 0.1323 0.2074 0.64 0.5244
## Year2005 0.0592 0.1856 0.32 0.7501
## Year2006 0.1161 0.1780 0.65 0.5149
```

```

## Year2007          0.2928      0.2021      1.45      0.1490
## Year2008          -0.1086      0.1663     -0.65      0.5146
## Year2009           0.0403      0.1706      0.24      0.8133
## Year2010          -0.0586      0.1976     -0.30      0.7670
## Year2011           0.0577      0.1674      0.34      0.7308
## Year2012          -0.0613      0.1675     -0.37      0.7148
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.403
## Multiple R-squared:  0.152, Adjusted R-squared:  0.0813
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 207 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.412  0.871  0.960  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      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 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
##   14   19   10   12   11   13    4    5    6   16   25   23   17   11   18
## 2011 2012
##   10   15
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    8    7    3    5    6    2    4    4    9   12   19    6    8   13
## 2011 2012

```

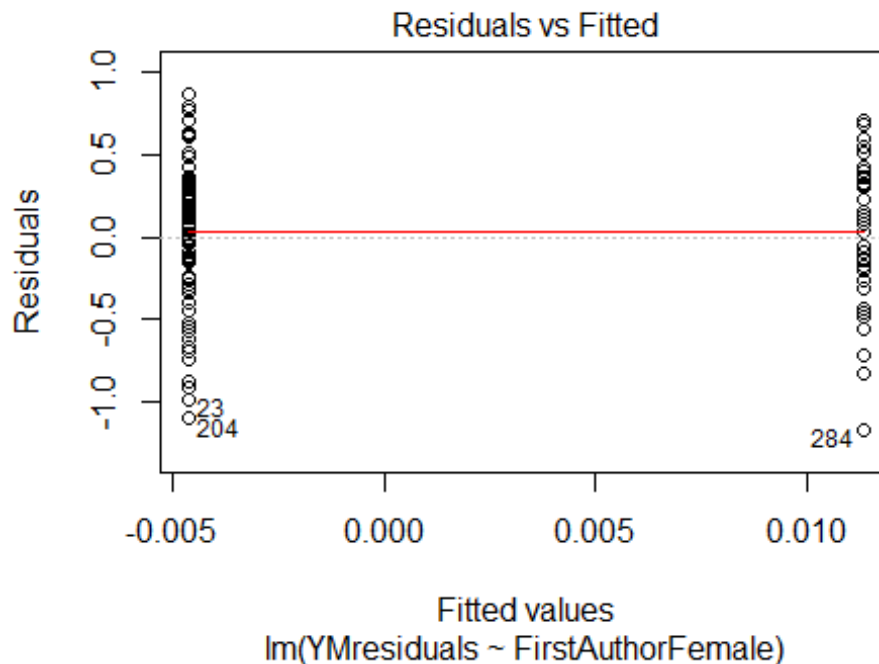
```
##      8    12
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    7    6    2    4    5    2    3    4    8   11   19    5    8   13
## 2011 2012
##    8    12
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 14, df = 16, p-value = 0.6
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.036, df = 1, p-value = 0.8
## [1] "Female first author team size 2018 geometric mean: 3.56208578803021"
## [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 = 5.5, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.56762134500816"
## [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 = 9.5, 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  3.590  1      1.895
## LastAuthorFemale  4.256  1      2.063
## UniqueAuthors    30.682  4      1.534
## Year              222.487 16      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 + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1250 -0.2484 0.0264 0.2489 1.0562
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.7544 0.2653 6.61 1.9e-09 ***
## FirstAuthorFemale1 -0.0293 0.1301 -0.23 0.8220
## LastAuthorFemale1 0.0249 0.1914 0.13 0.8968
## UniqueAuthors2 0.2526 0.2431 1.04 0.3013
## UniqueAuthors3 0.1205 0.2135 0.56 0.5737
## UniqueAuthors4 0.3175 0.2149 1.48 0.1427
## UniqueAuthors5 0.2376 0.2152 1.10 0.2723
## Year1997 -0.7411 0.4367 -1.70 0.0928 .
## Year1998 -0.8415 0.3275 -2.57 0.0117 *
## Year1999 -1.1212 0.8688 -1.29 0.1999
## Year2000 -0.7585 0.2965 -2.56 0.0120 *
## Year2001 -1.0590 0.4246 -2.49 0.0143 *
## Year2002 -1.2688 0.2838 -4.47 2.1e-05 ***
## Year2003 -0.4120 0.2992 -1.38 0.1717
## Year2004 -1.0145 0.3675 -2.76 0.0069 **
## Year2005 -0.6971 0.3136 -2.22 0.0285 *
## Year2006 -0.7146 0.2928 -2.44 0.0164 *
## Year2007 -0.8655 0.2983 -2.90 0.0046 **
## Year2008 -0.7710 0.3876 -1.99 0.0494 *
## Year2009 -0.7458 0.3394 -2.20 0.0303 *
## Year2010 -0.6548 0.2920 -2.24 0.0272 *
## Year2011 -0.8155 0.3302 -2.47 0.0152 *
## Year2012 -0.7628 0.2696 -2.83 0.0056 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.419
## Multiple R-squared: 0.215, Adjusted R-squared: 0.0403
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~1. The remaining 110 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.

```

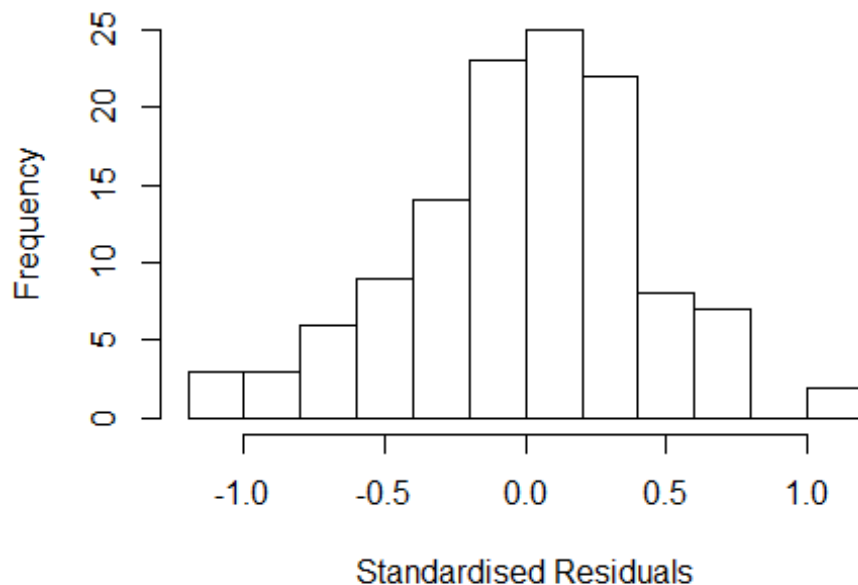
```
## 0.452 0.877 0.964 0.908 0.985 0.999
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          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"

## Warning in lmrob.fit(x, y, control, init = init, mf = mf): M-step did NOT
## converge. Returning unconverged SM-estimate

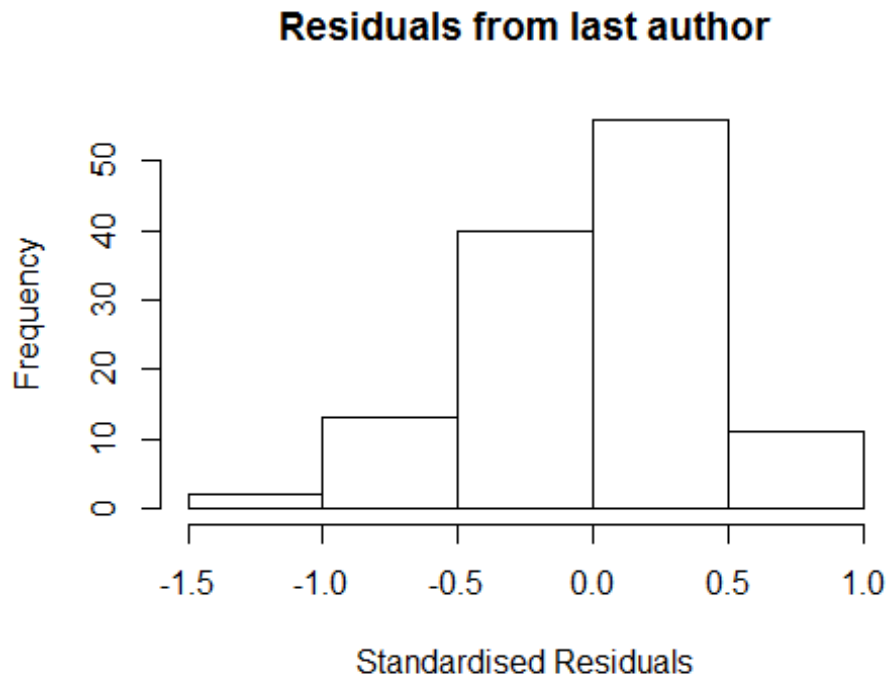
## [1] "Regression 3: First author gender, Year as factors"

## Warning in lmrob.fit(x, y, control, init = init, mf = mf): M-step did NOT
## converge. Returning unconverged SM-estimate
```

Residuals from first and last author and team size



```
## [1] "Regression 4: Last author gender, Year as factors"
##               GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 4.224 1      2.055
## Year             4.224 16      1.046
```



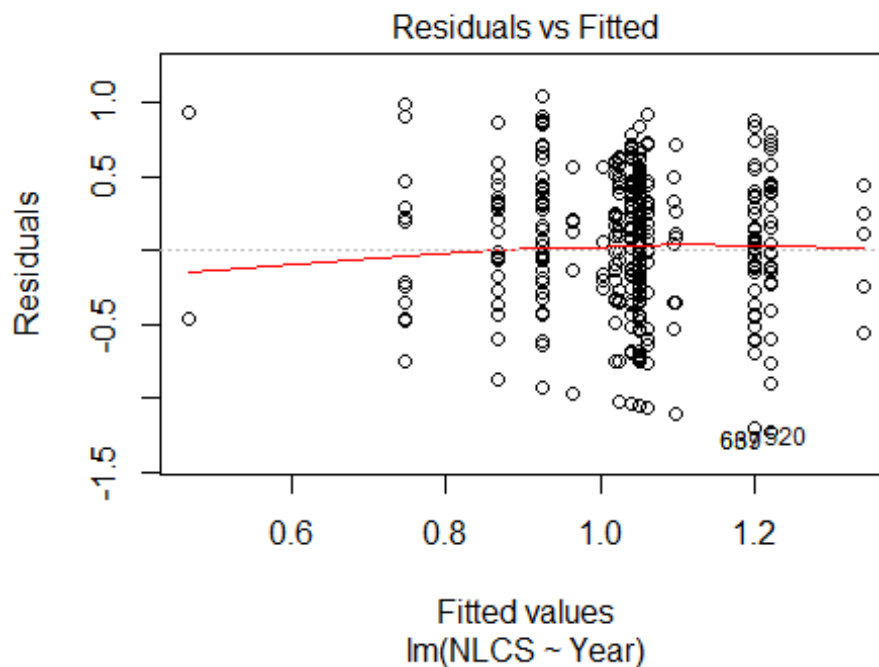
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2525 -0.2837  0.0317  0.2414  0.9664
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.8870     0.2193   8.60 8.7e-14 ***
## LastAuthorFemale1  0.0248     0.1762   0.14  0.8882
## Year1997        -0.6323     0.2966  -2.13  0.0353 *
## Year1998        -0.7720     0.2565  -3.01  0.0033 **
## Year1999        -1.1349     1.7433  -0.65  0.5165
## Year2000        -0.6682     0.2391  -2.79  0.0062 **
## Year2001        -0.9813     0.3865  -2.54  0.0126 *
```

```

## Year2002      -1.1384      0.2212      -5.15      1.3e-06 ***
## Year2003      -0.3689      0.2267      -1.63      0.1067
## Year2004      -0.9396      0.3790      -2.48      0.0148 *
## Year2005      -0.6566      0.2563      -2.56      0.0118 *
## Year2006      -0.6869      0.2516      -2.73      0.0074 **
## Year2007      -0.7701      0.2347      -3.28      0.0014 **
## Year2008      -0.6345      0.2943      -2.16      0.0334 *
## Year2009      -0.6345      0.2732      -2.32      0.0221 *
## Year2010      -0.5898      0.2420      -2.44      0.0165 *
## Year2011      -0.7328      0.2754      -2.66      0.0090 **
## Year2012      -0.6604      0.2269      -2.91      0.0044 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.418
## Multiple R-squared:  0.173, Adjusted R-squared:  0.0377
## Convergence in 35 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 110 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.349  0.872  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      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 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
##   19   23   19   20   36   36   54   43   50   54   51   77   57   42   74
## 2011 2012

```

```
## 79 73
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 5 5 6 3 7 5 21 15 23 25 26 36 38 31 47
## 2011 2012
## 59 52
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 5 5 4 3 5 5 17 10 19 22 23 31 36 26 42
## 2011 2012
## 54 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 = 16, df = 16, p-value = 0.5
```



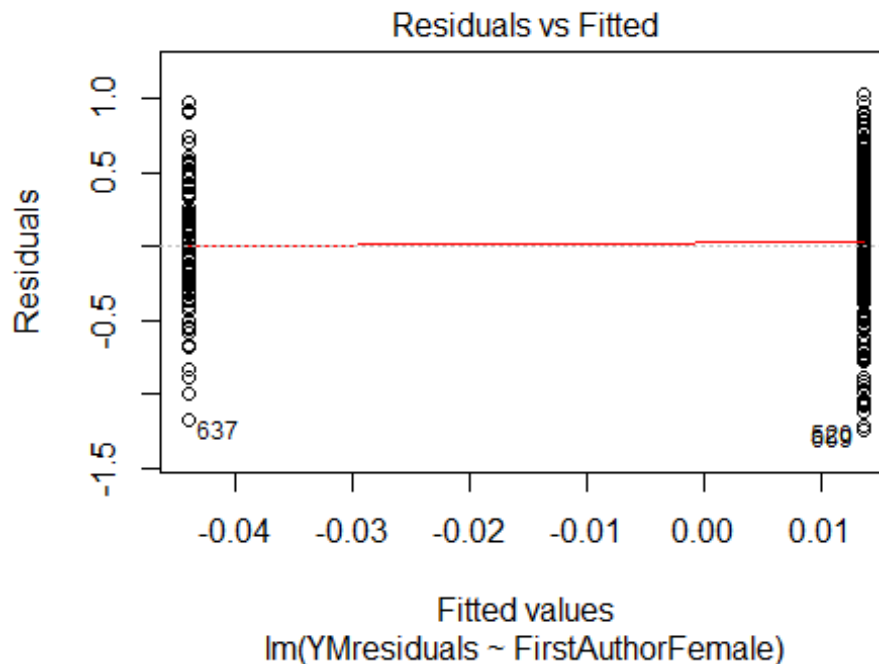
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2, df = 1, p-value = 0.2

## [1] "Female first author team size 2018 geometric mean: 3.84434415239624"
## [1] "Male first author team size 2018 geometric mean: 3.64272811235634"
```

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

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 440, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.98214415907465"
## [1] "Male last author team size 2018 geometric mean: 3.89907099703104"

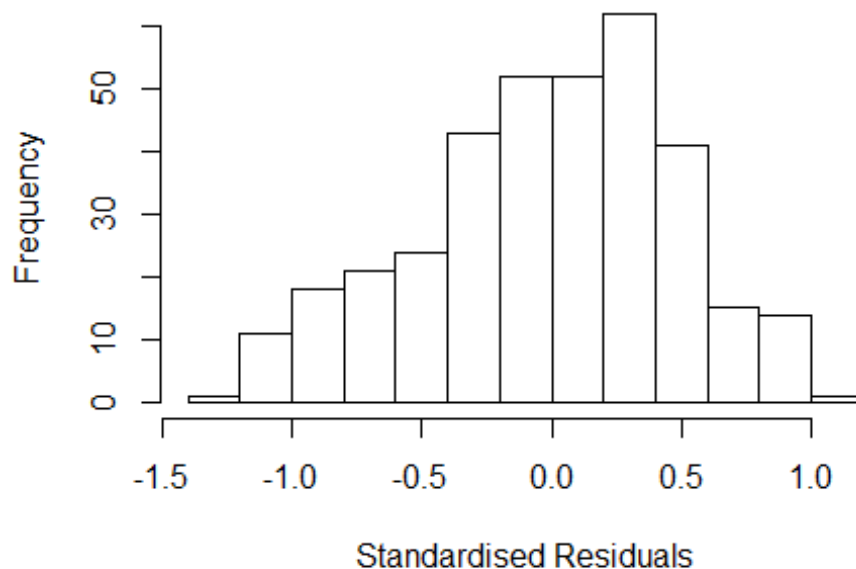
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 190, p-value = 0.1
## 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 2.085  1          1.444
```

| | | | |
|---------------------|-------|----|-------|
| ## LastAuthorFemale | 1.948 | 1 | 1.396 |
| ## UniqueAuthors | 2.990 | 4 | 1.147 |
| ## Year | 3.038 | 16 | 1.035 |

Residuals from first and last author and team size



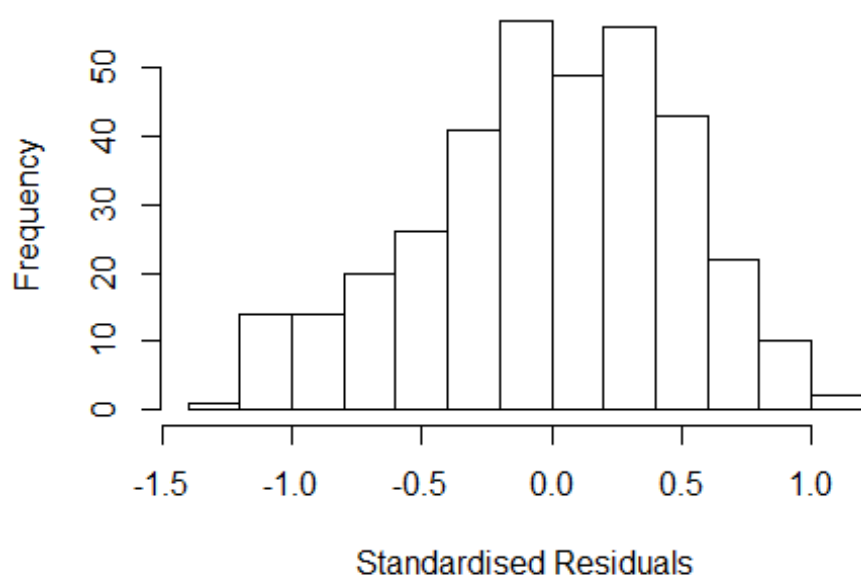
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.25124 -0.33149  0.00841  0.35631  1.03626
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.04259    0.23117   4.51   9e-06 ***
## FirstAuthorFemale1 -0.02821    0.08243  -0.34    0.73
## LastAuthorFemale1 -0.01923    0.09175  -0.21    0.83
## UniqueAuthors2    -0.00769    0.11952  -0.06    0.95
## UniqueAuthors3     0.09372    0.11745   0.80    0.43
## UniqueAuthors4     0.08045    0.11968   0.67    0.50
```

```

## UniqueAuthors5      0.13495      0.11115      1.21      0.23
## Year1997             -0.10182      0.26855     -0.38      0.70
## Year1998             -0.09801      0.36592     -0.27      0.79
## Year1999             -0.68535      0.52107     -1.32      0.19
## Year2000             -0.20626      0.32111     -0.64      0.52
## Year2001              0.30260      0.27567      1.10      0.27
## Year2002             -0.08087      0.23850     -0.34      0.73
## Year2003             -0.30505      0.28780     -1.06      0.29
## Year2004             -0.18353      0.24257     -0.76      0.45
## Year2005             -0.26256      0.24400     -1.08      0.28
## Year2006              0.00623      0.25074      0.02      0.98
## Year2007              0.03105      0.23128      0.13      0.89
## Year2008             -0.05927      0.23794     -0.25      0.80
## Year2009              0.07370      0.24855      0.30      0.77
## Year2010             -0.19827      0.24208     -0.82      0.41
## Year2011             -0.03152      0.22658     -0.14      0.89
## Year2012             -0.06070      0.23598     -0.26      0.80
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.514
## Multiple R-squared:  0.0649, Adjusted R-squared:  0.00298
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 325 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.532  0.882  0.953  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      2.82e-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 1.836 1      1.355
## LastAuthorFemale  1.780 1      1.334
## Year              1.300 16      1.008

```


Residuals from first and last author



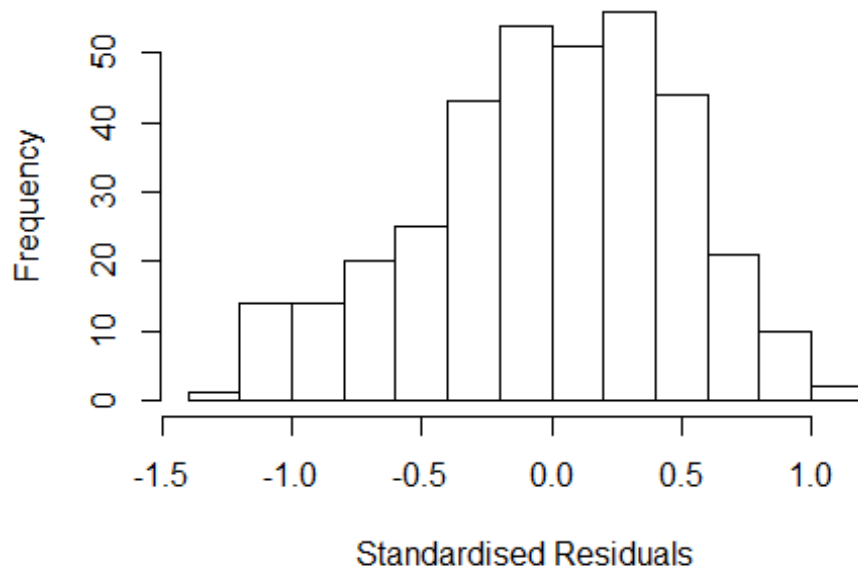
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2187 -0.3407 0.0303 0.3445 1.1003
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0983 0.1931 5.69 2.8e-08 ***
## FirstAuthorFemale1 -0.0524 0.0791 -0.66 0.51
## LastAuthorFemale1 -0.0416 0.0897 -0.46 0.64
## Year1997 -0.0669 0.2418 -0.28 0.78
## Year1998 -0.0628 0.3688 -0.17 0.86
## Year1999 -0.7461 0.5355 -1.39 0.16
## Year2000 -0.1345 0.3028 -0.44 0.66
## Year2001 0.2752 0.2444 1.13 0.26
## Year2002 -0.0619 0.2150 -0.29 0.77
## Year2003 -0.2537 0.2700 -0.94 0.35
## Year2004 -0.1567 0.2224 -0.70 0.48
## Year2005 -0.2400 0.2203 -1.09 0.28
```

```

## Year2006          0.0353      0.2299      0.15      0.88
## Year2007          0.0661      0.2076      0.32      0.75
## Year2008         -0.0285      0.2146     -0.13      0.89
## Year2009          0.1204      0.2233      0.54      0.59
## Year2010         -0.1576      0.2211     -0.71      0.48
## Year2011          0.0155      0.2028      0.08      0.94
## Year2012         -0.0100      0.2115     -0.05      0.96
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.503
## Multiple R-squared:  0.0586, Adjusted R-squared:  0.00818
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 33 weights are ~= 1. The remaining 322 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.537  0.870  0.953  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      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 1.142 1          1.069
## Year              1.142 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
## -1.2157 -0.3411 0.0361 0.3445 1.1198
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.09826 0.19313 5.69 2.8e-08 ***
## FirstAuthorFemale1 -0.07023 0.06322 -1.11 0.27
## Year1997 -0.06564 0.24467 -0.27 0.79
## Year1998 -0.05791 0.37194 -0.16 0.88
## Year1999 -0.74786 0.54255 -1.38 0.17
## Year2000 -0.12998 0.30355 -0.43 0.67
## Year2001 0.27504 0.24566 1.12 0.26
## Year2002 -0.05894 0.21397 -0.28 0.78
## Year2003 -0.26209 0.26799 -0.98 0.33
## Year2004 -0.16046 0.22218 -0.72 0.47
## Year2005 -0.23808 0.22048 -1.08 0.28
## Year2006 0.03101 0.23111 0.13 0.89
```

```

## Year2007          0.06262    0.20769    0.30    0.76
## Year2008         -0.02964    0.21469   -0.14    0.89
## Year2009          0.11747    0.22330    0.53    0.60
## Year2010         -0.16137    0.22049   -0.73    0.46
## Year2011          0.00963    0.20276    0.05    0.96
## Year2012         -0.01438    0.21141   -0.07    0.95
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.502
## Multiple R-squared:  0.0583, Adjusted R-squared:  0.0108
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 325 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.537  0.872  0.953  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      2.82e-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 1.103 1      1.050
## Year      1.103 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
## -1.2077 -0.3436  0.0249  0.3476  1.0482

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.09827    0.19320   5.68 2.8e-08 ***
## LastAuthorFemale1 -0.07071    0.07217  -0.98   0.33
## Year1997        -0.09118    0.23526  -0.39   0.70
## Year1998        -0.07582    0.35935  -0.21   0.83
## Year1999        -0.74646    0.51477  -1.45   0.15
## Year2000        -0.14703    0.29891  -0.49   0.62
## Year2001         0.26022    0.24898   1.05   0.30
## Year2002        -0.07473    0.21420  -0.35   0.73
## Year2003        -0.25735    0.27222  -0.95   0.35
## Year2004        -0.16094    0.22222  -0.72   0.47
## Year2005        -0.25062    0.21977  -1.14   0.25
## Year2006         0.02937    0.22947   0.13   0.90
## Year2007         0.05680    0.20755   0.27   0.78
## Year2008        -0.03855    0.21353  -0.18   0.86
## Year2009         0.10941    0.22310   0.49   0.62
## Year2010        -0.16215    0.22213  -0.73   0.47
## Year2011         0.00528    0.20248   0.03   0.98
## Year2012        -0.01246    0.21175  -0.06   0.95
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.501
## Multiple R-squared:  0.0569, Adjusted R-squared:  0.00929
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 326 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.541  0.869  0.954  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      2.82e-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: 355"
## [1] ""

```

```

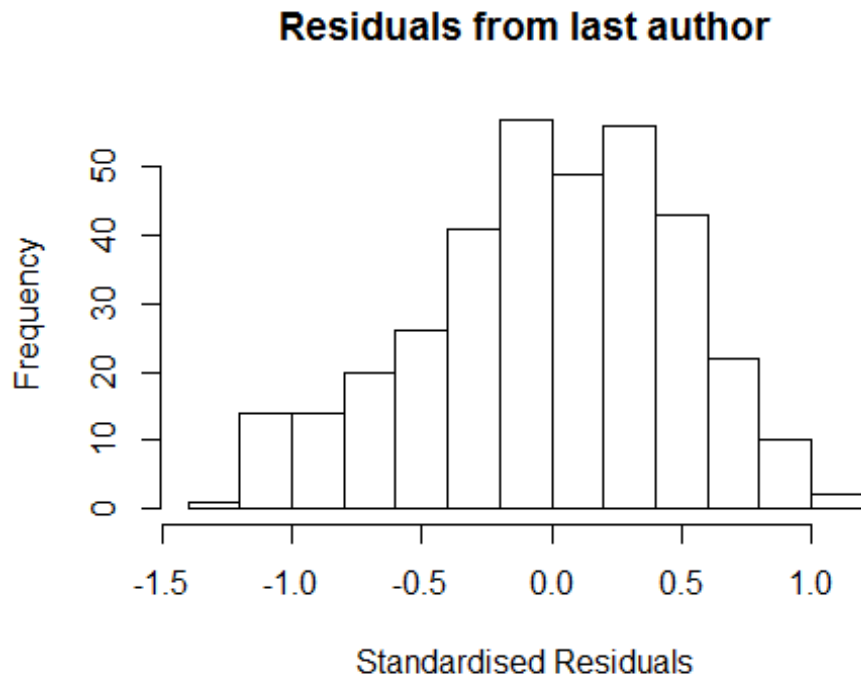
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2733"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 11 14 5 12 15 24 17 24 18 18 27 19 26 13 23
## 2011 2012
## 19 16
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 0 1 1 6 4 7 12 11 7 7 14 7 12 4 12
## 2011 2012
## 7 6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 0 1 1 3 3 4 7 8 6 5 9 5 10 3 8
## 2011 2012
## 6 5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.28225473667665"
## [1] "Male first author team size 2018 geometric mean: 3.93597934253086"

## Warning in wilcox.test.default(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 = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.91901897169873"
## [1] "Male last author team size 2018 geometric mean: 3.30975091964687"

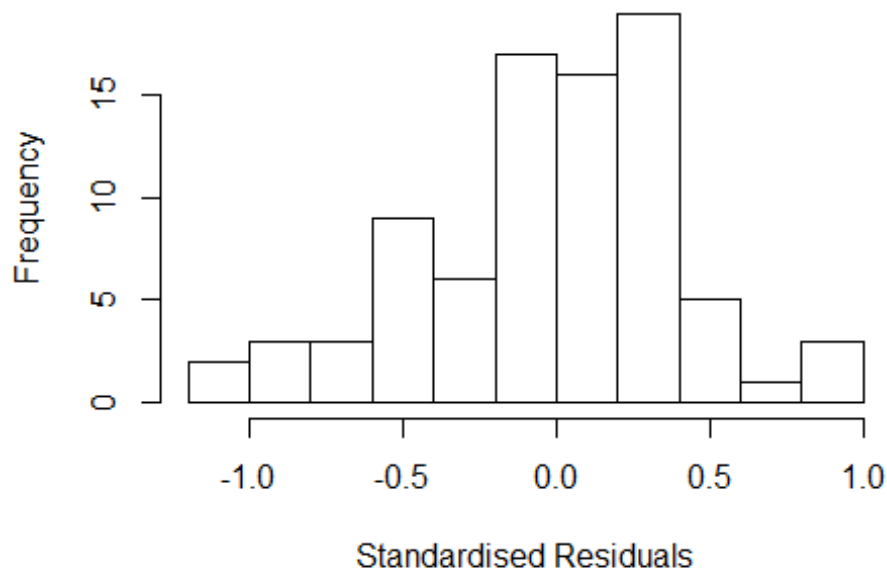
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 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 3.289e+00 1      1.814e+00
## LastAuthorFemale  1.272e+13 1      3.567e+06
## UniqueAuthors    4.397e+14 4      6.767e+01
## Year              2.706e+15 15     3.269e+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
## -1.151 -0.272 0.021 0.253 0.946
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.7936 0.1980 9.06 6.0e-13 ***
## FirstAuthorFemale1 -0.0401 0.1179 -0.34 0.73468
## LastAuthorFemale1 0.2079 0.1707 1.22 0.22790
## UniqueAuthors2 -0.0565 0.2297 -0.25 0.80638
## UniqueAuthors3 0.0248 0.1815 0.14 0.89153
## UniqueAuthors4 -0.0663 0.1867 -0.35 0.72380
## UniqueAuthors5 -0.0806 0.1980 -0.41 0.68546
## Year1998 -0.7279 0.2730 -2.67 0.00977 **
## Year1999 -0.3853 0.2321 -1.66 0.10195
## Year2000 -0.5003 0.2862 -1.75 0.08535 .
```



```

## Year2001          -0.8601      0.4012    -2.14    0.03599 *
## Year2002          -0.7437      0.1420    -5.24    2.1e-06 ***
## Year2003          -0.5608      0.2261    -2.48    0.01584 *
## Year2004          -0.6674      0.1883    -3.54    0.00076 ***
## Year2005          -1.2411      0.3219    -3.86    0.00028 ***
## Year2006          -0.7688      0.2062    -3.73    0.00042 ***
## Year2007          -0.8775      0.1612    -5.45    9.4e-07 ***
## Year2008          -0.6911      0.1788    -3.86    0.00027 ***
## Year2009          -0.5746      0.2330    -2.47    0.01646 *
## Year2010          -0.8745      0.2028    -4.31    5.9e-05 ***
## Year2011          -0.7009      0.1877    -3.73    0.00041 ***
## Year2012          -0.7259      0.2250    -3.23    0.00201 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.446
## Multiple R-squared:  0.184, Adjusted R-squared:  -0.0925
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 72 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.485  0.891  0.955  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           1.19e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01             5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500           50           2           1           1000           200
## trace.lev      mts      compute.rd
##           0           1000           0
##           psi             subsampling             cov
##           "bisquare"             "nonsingular"             ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.905e+00  1      1.704e+00
## LastAuthorFemale  1.532e+13  1      3.914e+06
## Year              3.957e+13 15      2.840e+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)
##

```

```

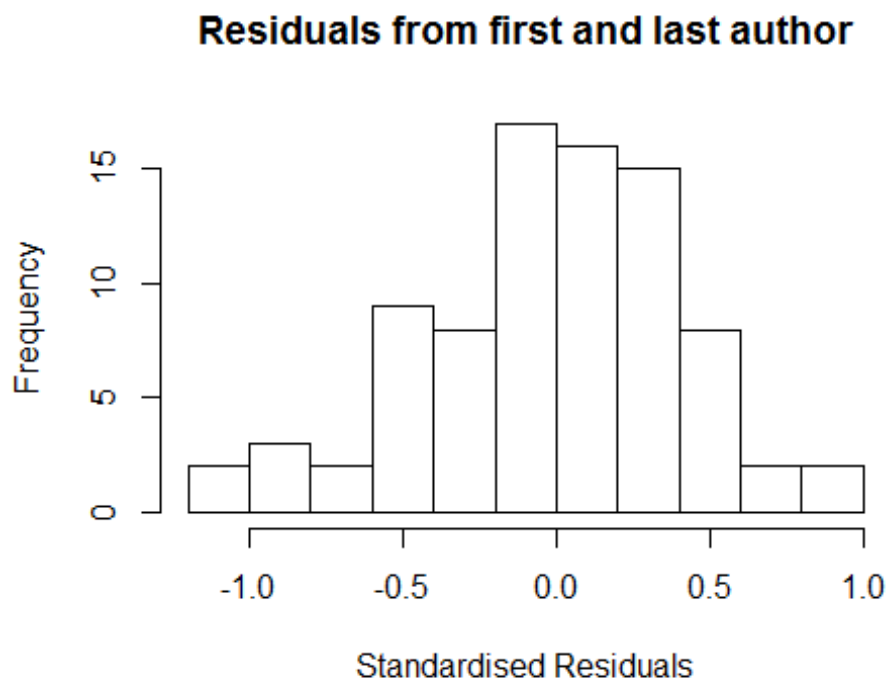
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1228 -0.2493  0.0114  0.2431  0.9307
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.71e+00   1.54e-08  1.11e+08 < 2e-16 ***
## FirstAuthorFemale1 -6.41e-02   1.17e-01 -5.50e-01  0.58526
## LastAuthorFemale1  1.97e-01   1.72e-01  1.14e+00  0.25678
## Year1998        -6.93e-01   1.72e-01 -4.02e+00  0.00015 ***
## Year1999        -3.39e-01   1.67e-01 -2.03e+00  0.04638 *
## Year2000        -4.88e-01   2.40e-01 -2.03e+00  0.04599 *
## Year2001        -8.21e-01   3.52e-01 -2.33e+00  0.02269 *
## Year2002        -6.67e-01   5.70e-02 -1.17e+01 < 2e-16 ***
## Year2003        -5.22e-01   2.41e-01 -2.17e+00  0.03387 *
## Year2004        -5.90e-01   1.89e-01 -3.13e+00  0.00261 **
## Year2005        -1.23e+00   2.75e-01 -4.46e+00  3.2e-05 ***
## Year2006        -7.32e-01   1.70e-01 -4.30e+00  5.7e-05 ***
## Year2007        -8.31e-01   1.45e-01 -5.73e+00  2.7e-07 ***
## Year2008        -6.63e-01   1.34e-01 -4.96e+00  5.3e-06 ***
## Year2009        -4.81e-01   2.03e-01 -2.38e+00  0.02043 *
## Year2010        -7.95e-01   1.76e-01 -4.51e+00  2.7e-05 ***
## Year2011        -6.61e-01   1.61e-01 -4.10e+00  0.00012 ***
## Year2012        -6.64e-01   1.89e-01 -3.52e+00  0.00080 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.406
## Multiple R-squared:  0.188, Adjusted R-squared:  -0.0205
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 71 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.424  0.873  0.946  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      1.19e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01        5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500            50          2            1            1000      200
##      trace.lev      mts      compute.rd

```

```
##           0           1000           0
##           psi         subsampling         cov
##           "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"

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

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



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

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
```

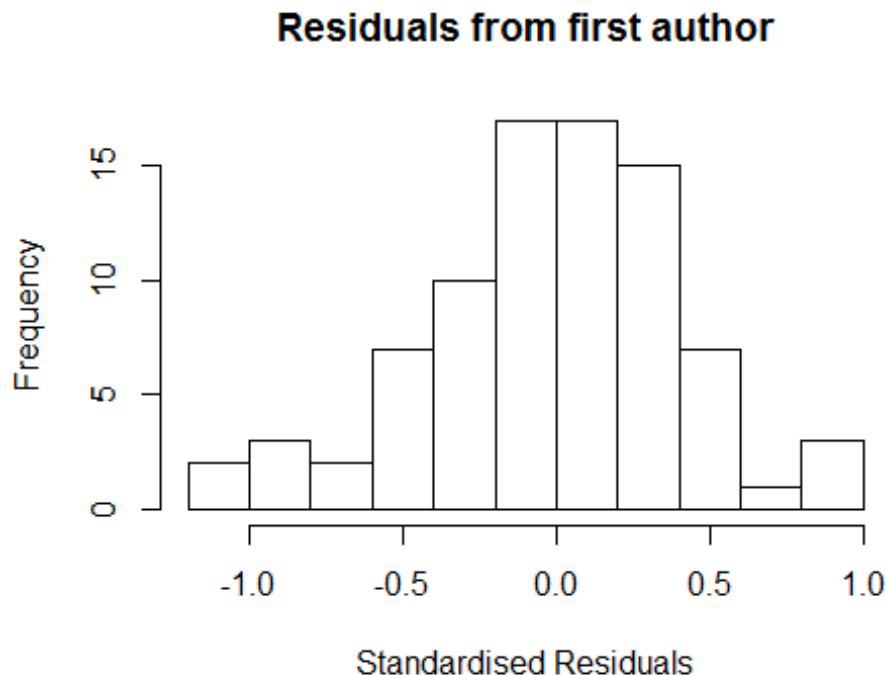
```

## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.141 -0.245  0.024  0.247  0.930
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.7130     0.0000      Inf < 2e-16 ***
## FirstAuthorFemale1 -0.0224     0.1113     -0.20  0.84099
## Year1998          -0.4960     0.0000     -Inf < 2e-16 ***
## Year1999          -0.3395     0.1671     -2.03  0.04614 *
## Year2000          -0.4878     0.2396     -2.04  0.04576 *
## Year2001          -0.8444     0.3738     -2.26  0.02715 *
## Year2002          -0.6734     0.0573    -11.75 < 2e-16 ***
## Year2003          -0.5306     0.2378     -2.23  0.02902 *
## Year2004          -0.5723     0.1843     -3.10  0.00279 **
## Year2005          -1.2250     0.2744     -4.47  3.1e-05 ***
## Year2006          -0.7413     0.1717     -4.32  5.3e-05 ***
## Year2007          -0.7748     0.1379     -5.62  4.0e-07 ***
## Year2008          -0.6479     0.1265     -5.12  2.8e-06 ***
## Year2009          -0.5063     0.2022     -2.50  0.01474 *
## Year2010          -0.7858     0.1859     -4.23  7.4e-05 ***
## Year2011          -0.6500     0.1725     -3.77  0.00035 ***
## Year2012          -0.6760     0.1882     -3.59  0.00062 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.407
## Multiple R-squared:  0.179, Adjusted R-squared:  -0.0171
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  10 weights are ~= 1. The remaining 74 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.413  0.874  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          1.19e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01          5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500           50           2           1          1000           200
##      trace.lev      mts      compute.rd
##           0          1000           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 3.532e+13 1      5.943e+06
## Year              3.532e+13 15      2.829e+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
## -1.0852 -0.2425  0.0227  0.2553  0.9276
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.7130     0.0000    Inf < 2e-16 ***
## LastAuthorFemale1  0.1634     0.1604     1.02  0.31187
```

```

## Year1998      -0.6594      0.1604      -4.11      0.00011 ***
## Year1999      -0.3397      0.1667      -2.04      0.04548 *
## Year2000      -0.4869      0.2384      -2.04      0.04505 *
## Year2001      -0.8553      0.3732      -2.29      0.02506 *
## Year2002      -0.6767      0.0551     -12.27      < 2e-16 ***
## Year2003      -0.5372      0.2307      -2.33      0.02290 *
## Year2004      -0.6278      0.1668      -3.76      0.00036 ***
## Year2005      -1.2226      0.2729      -4.48      3.0e-05 ***
## Year2006      -0.7463      0.1708      -4.37      4.4e-05 ***
## Year2007      -0.8463      0.1365      -6.20      4.0e-08 ***
## Year2008      -0.6662      0.1339      -4.98      4.8e-06 ***
## Year2009      -0.5198      0.2012      -2.58      0.01198 *
## Year2010      -0.8163      0.1747      -4.67      1.5e-05 ***
## Year2011      -0.6910      0.1409      -4.90      6.3e-06 ***
## Year2012      -0.6858      0.1839      -3.73      0.00040 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.414
## Multiple R-squared:  0.183, Adjusted R-squared:  -0.0127
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 73 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.473  0.876  0.957   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           1.19e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01           5.00e-01
##   nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500           50           2           1           1000           200
##   trace.lev      mts      compute.rd
##           0           1000           0
##           psi           subsampling           cov
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  84"
## [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]"

```

```

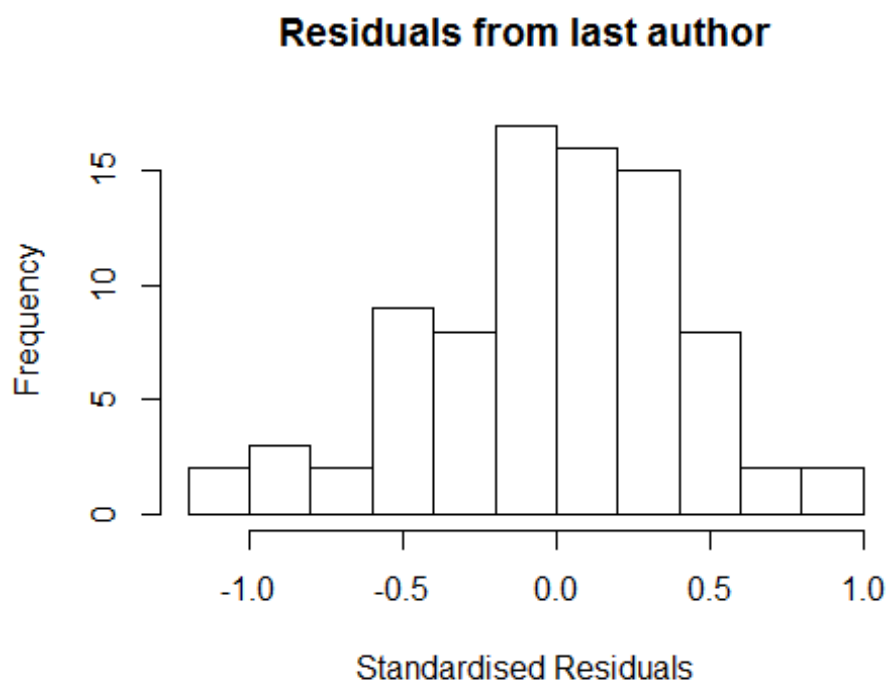
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 10 12 5 2 2 16 3 22 17 13 18 35 19 22 23
## 2011 2012
## 25 15
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3 3 1 2 1 3 0 9 11 6 11 22 10 11 18
## 2011 2012
## 20 11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2 2 1 2 1 2 0 7 10 5 11 20 9 9 18
## 2011 2012
## 19 11
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.1365646722804"
## [1] "Male first author team size 2018 geometric mean: 4.45628831179955"

## Warning in wilcox.test.default(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 = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.86825415074099"
## [1] "Male last author team size 2018 geometric mean: 3.54347545128651"

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

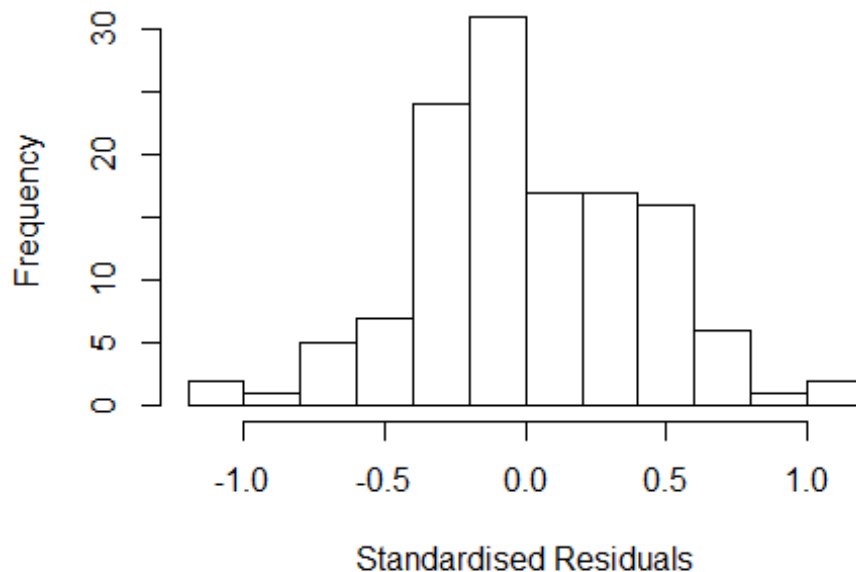
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 44, 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 | 6.567e+13 | 1 | 8.104e+06 |
| ## LastAuthorFemale | 4.771e+00 | 1 | 2.184e+00 |
| ## UniqueAuthors | 1.104e+16 | 4 | 1.012e+02 |
| ## Year | 5.571e+16 | 15 | 3.616e+00 |

Residuals from first and last author and team size



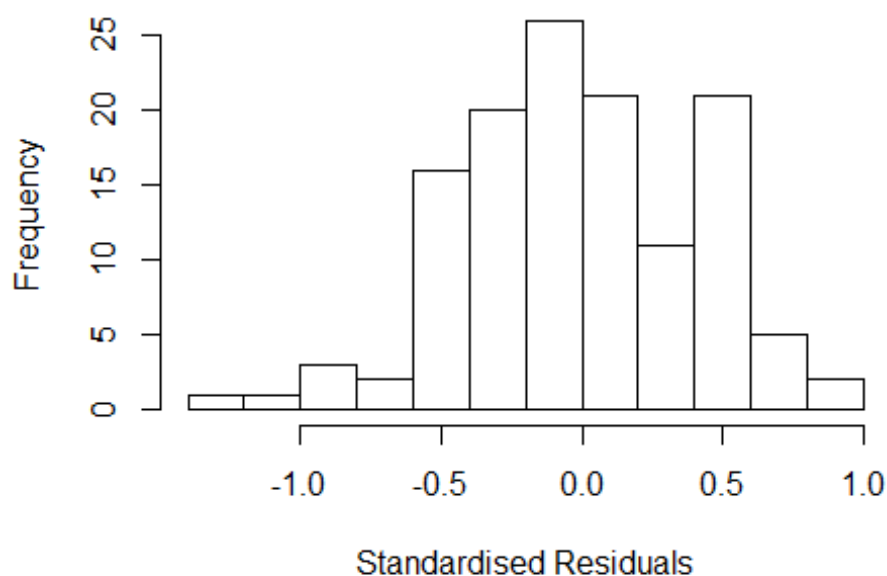
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1047 -0.2600 -0.0404  0.2951  1.1636
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.12493    0.25744   4.37 2.9e-05 ***
## FirstAuthorFemale1 -0.23237    0.08356  -2.78  0.0064 **
## LastAuthorFemale1  0.00319    0.08103   0.04  0.9686
## UniqueAuthors2   -0.03888    0.25656  -0.15  0.8798
## UniqueAuthors3    0.09539    0.27708   0.34  0.7313
## UniqueAuthors4    0.20506    0.24995   0.82  0.4138
## UniqueAuthors5    0.30234    0.26803   1.13  0.2618
## Year1997         -0.37984    0.45161  -0.84  0.4022
## Year1998          0.26225    0.30089   0.87  0.3854
## Year1999         -0.09513    0.19594  -0.49  0.6283
```

```

## Year2000          0.00753    0.14746    0.05    0.9594
## Year2001          0.20840    0.49451    0.42    0.6743
## Year2003          0.01868    0.34825    0.05    0.9573
## Year2004         -0.11777    0.14891   -0.79    0.4308
## Year2005         -0.03365    0.22425   -0.15    0.8810
## Year2006         -0.08668    0.16606   -0.52    0.6028
## Year2007         -0.17858    0.13507   -1.32    0.1890
## Year2008         -0.07293    0.17483   -0.42    0.6774
## Year2009         -0.51104    0.12368   -4.13    7.2e-05 ***
## Year2010         -0.22705    0.16039   -1.42    0.1598
## Year2011         -0.10062    0.16164   -0.62    0.5349
## Year2012          0.13410    0.15658    0.86    0.3937
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.428
## Multiple R-squared:  0.239, Adjusted R-squared:  0.0891
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 116 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.439  0.887  0.952  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      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 5.329e+13 1      7.300e+06
## LastAuthorFemale  2.607e+01 1      5.106e+00
## Year              1.372e+15 15      3.196e+00

```

Residuals from first and last author



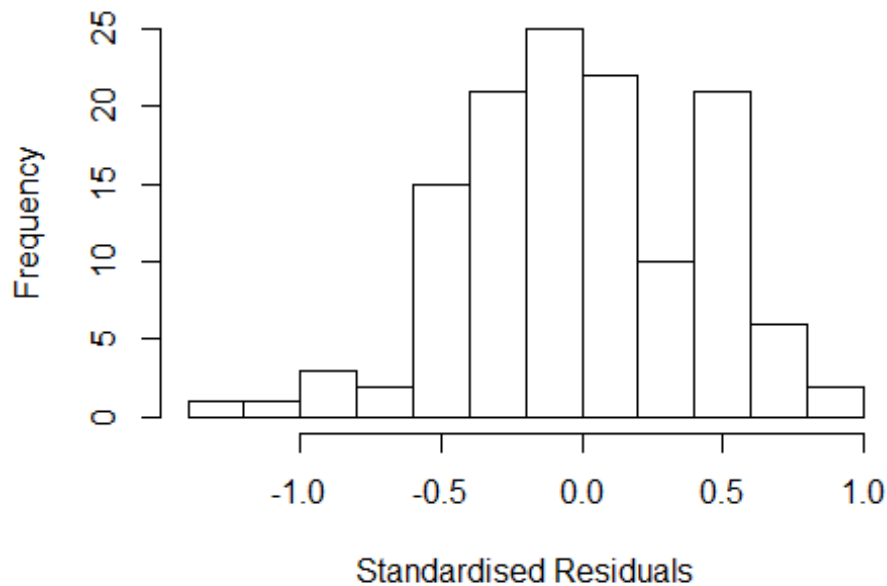
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.340 -0.267 -0.025 0.348 0.938
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.14357 0.04639 24.65 < 2e-16 ***
## FirstAuthorFemale1 -0.21314 0.08610 -2.48 0.01482 *
## LastAuthorFemale1 0.01311 0.08236 0.16 0.87384
## Year1997 -0.41305 0.33268 -1.24 0.21701
## Year1998 0.21446 0.09492 2.26 0.02581 *
## Year1999 -0.06607 0.10835 -0.61 0.54326
## Year2000 0.28132 0.09412 2.99 0.00345 **
## Year2001 0.20345 0.36222 0.56 0.57548
## Year2003 0.19664 0.24116 0.82 0.41659
## Year2004 0.06830 0.12665 0.54 0.59077
## Year2005 0.17654 0.20084 0.88 0.38129
## Year2006 0.04334 0.14419 0.30 0.76430
```

```

## Year2007          -0.06046    0.09915   -0.61  0.54322
## Year2008          0.08059    0.15799    0.51  0.61101
## Year2009         -0.44500    0.11820   -3.76  0.00027 ***
## Year2010         -0.08214    0.12692   -0.65  0.51884
## Year2011          0.00176    0.13384    0.01  0.98955
## Year2012          0.33660    0.11825    2.85  0.00526 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.489
## Multiple R-squared:  0.171, Adjusted R-squared:  0.044
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 119 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.432  0.908  0.954  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      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 5.506e+14 1      2.346e+07
## Year              5.506e+14 15      3.100e+00

```

Residuals from first author



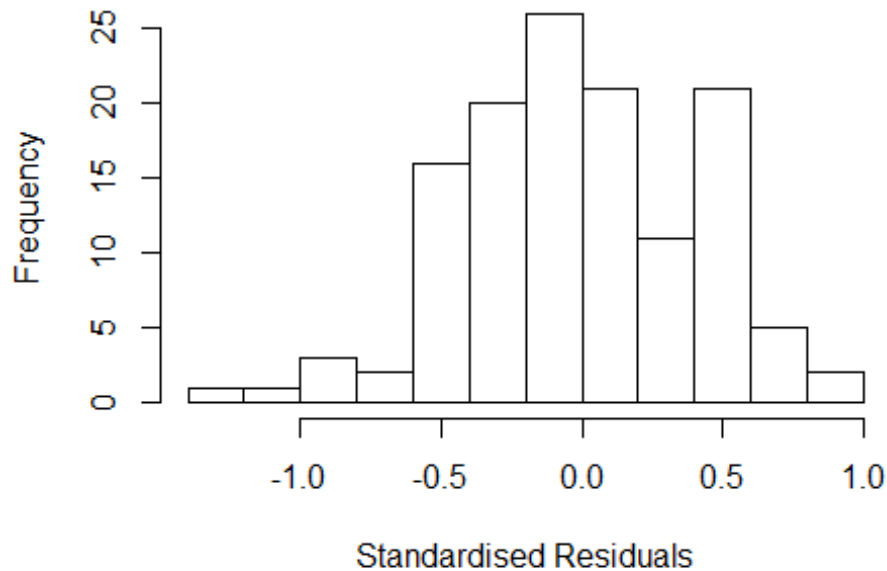
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3491 -0.2595 -0.0254 0.3481 0.9401
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.14262 0.04728 24.17 < 2e-16 ***
## FirstAuthorFemale1 -0.21124 0.08747 -2.41 0.01736 *
## Year1997 -0.40650 0.33703 -1.21 0.23031
## Year1998 0.22662 0.04728 4.79 5.1e-06 ***
## Year1999 -0.06512 0.10876 -0.60 0.55055
## Year2000 0.29538 0.04728 6.25 7.7e-09 ***
## Year2001 0.21000 0.35301 0.59 0.55312
## Year2003 0.20649 0.23566 0.88 0.38279
## Year2004 0.07512 0.11681 0.64 0.52148
## Year2005 0.18404 0.18789 0.98 0.32945
## Year2006 0.05240 0.13432 0.39 0.69717
## Year2007 -0.05553 0.09277 -0.60 0.55063
```

```

## Year2008          0.08434    0.15599    0.54  0.58980
## Year2009         -0.43914    0.11628   -3.78  0.00026 ***
## Year2010         -0.07448    0.11689   -0.64  0.52531
## Year2011          0.00787    0.12747    0.06  0.95087
## Year2012          0.33949    0.11748    2.89  0.00463 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.484
## Multiple R-squared:  0.172, Adjusted R-squared:  0.0533
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 120 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.417  0.907   0.954   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      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 3.888 1          1.972
## Year            3.888 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
## -1.2172 -0.2890 -0.0375 0.3033 0.9563
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.03700 0.09382 11.05 <2e-16 ***
## LastAuthorFemale1 -0.02161 0.08519 -0.25 0.8002
## Year1997 -0.39569 0.26292 -1.51 0.1351
## Year1998 0.14261 0.12673 1.13 0.2629
## Year1999 0.04050 0.13558 0.30 0.7657
## Year2000 0.42261 0.12673 3.33 0.0012 **
## Year2001 0.22081 0.25708 0.86 0.3922
## Year2003 0.18018 0.24185 0.74 0.4578
## Year2004 0.08305 0.17042 0.49 0.6270
## Year2005 0.12130 0.24144 0.50 0.6164
## Year2006 0.00786 0.17216 0.05 0.9637
## Year2007 -0.03627 0.13343 -0.27 0.7863
```

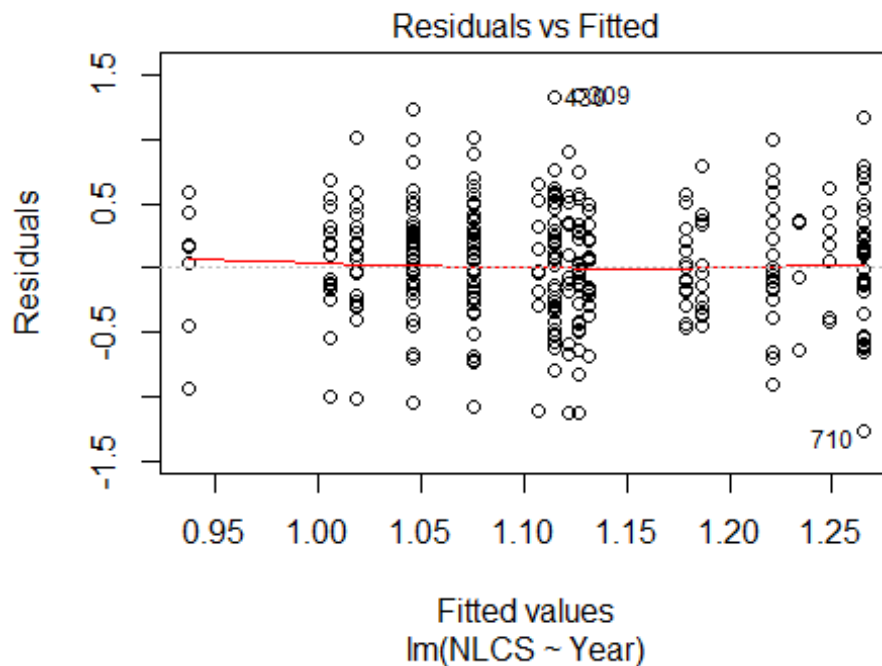
```

## Year2008          0.02322      0.20064      0.12      0.9081
## Year2009         -0.39364      0.15059     -2.61      0.0102 *
## Year2010         -0.09891      0.15956     -0.62      0.5366
## Year2011          0.00569      0.16599      0.03      0.9727
## Year2012          0.32150      0.15741      2.04      0.0435 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.497
## Multiple R-squared:  0.129, Adjusted R-squared:  0.00453
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 116 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.527  0.895  0.963  0.926  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      7.75e-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 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
##   11   21   34   17   31   30   25   24   28   23   45   41   70   44   66
## 2011 2012
##   68   53
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4   10    7    7   14    9   12   12   17   10   27   24   40   21   52
## 2011 2012
##   49   38

```



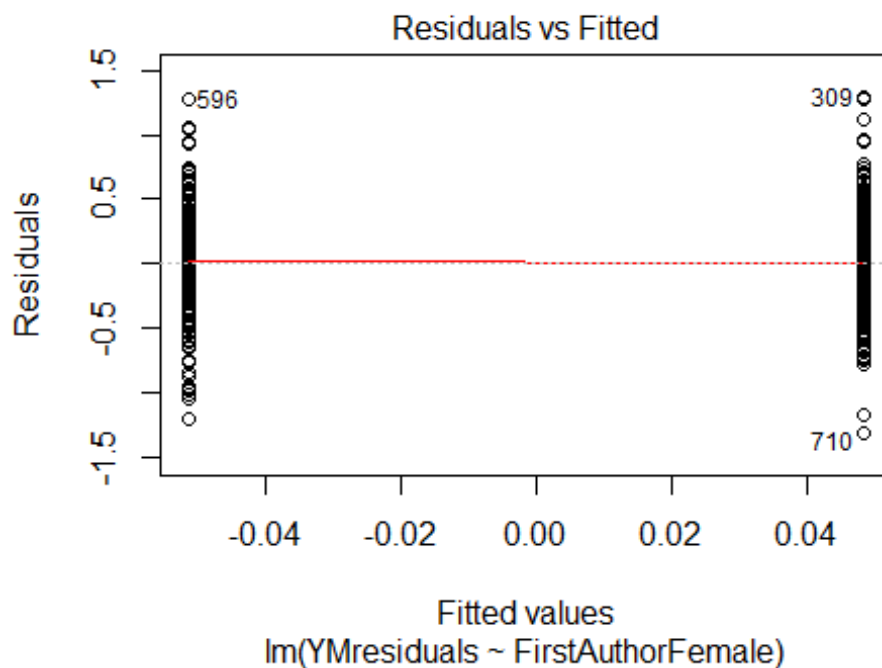
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4   10    7    7   14    7    9    9   16    9   22   18   35   19   45
## 2011 2012
##   42   32
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity 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 = 0.0074, df = 1, p-value = 0.9
## [1] "Female first author team size 2018 geometric mean: 4.05824093463222"
## [1] "Male first author team size 2018 geometric mean: 3.10450802591602"
## Warning in wilcox.test.default(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] "Female last author team size 2018 geometric mean: 3.32401094260808"
## [1] "Male last author team size 2018 geometric mean: 3.91081040974378"

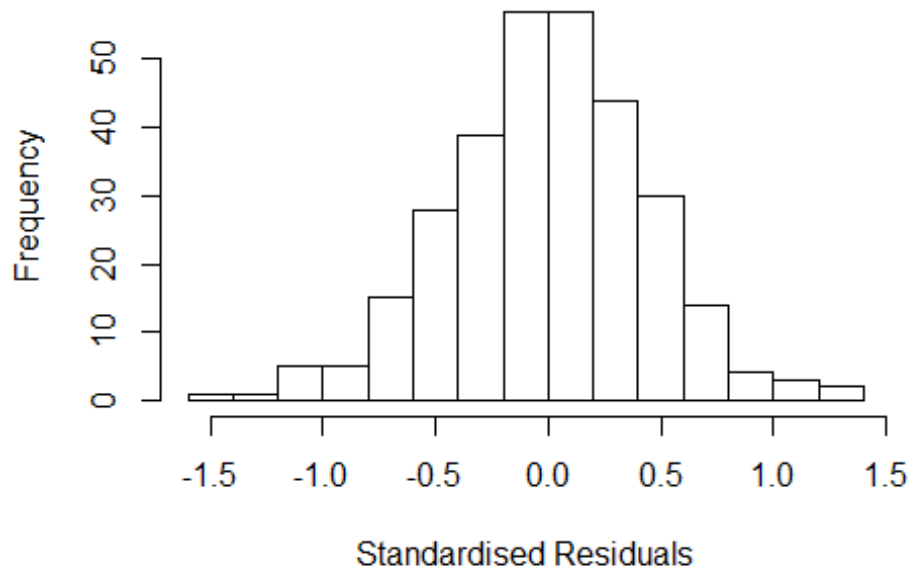
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## 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 | 1.208 | 1 | 1.099 |
| LastAuthorFemale | 1.504 | 1 | 1.226 |
| UniqueAuthors | 3.402 | 4 | 1.165 |
| Year | 4.635 | 16 | 1.049 |

Residuals from first and last author and team size



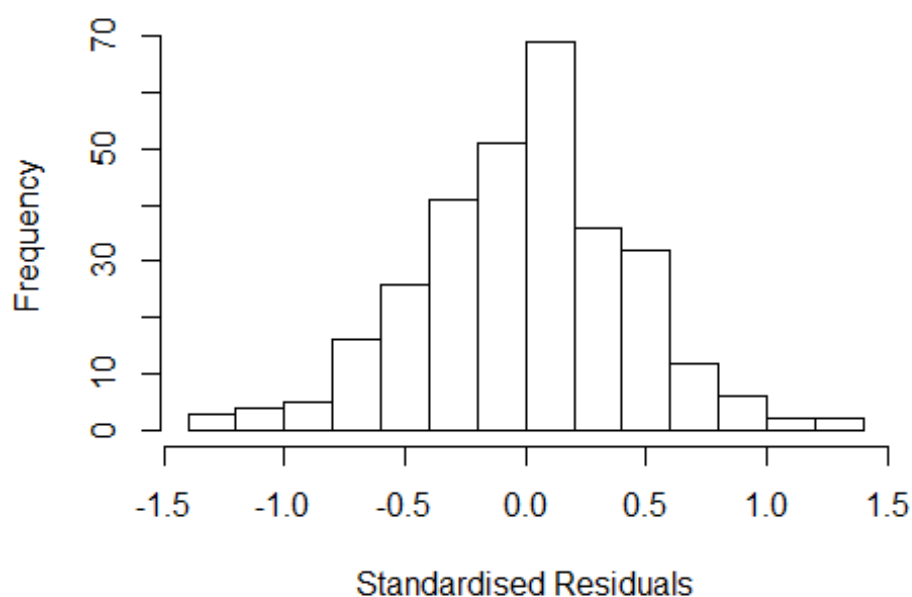
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.40542 -0.27206 0.00557 0.27900 1.39077
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.22870 0.30224 4.07 6.2e-05 ***
## FirstAuthorFemale1 0.07511 0.05226 1.44 0.152
## LastAuthorFemale1 -0.15618 0.06815 -2.29 0.023 *
## UniqueAuthors2 -0.05202 0.19442 -0.27 0.789
## UniqueAuthors3 0.03170 0.19300 0.16 0.870
## UniqueAuthors4 -0.04854 0.19286 -0.25 0.801
## UniqueAuthors5 0.11966 0.19160 0.62 0.533
## Year1997 0.00525 0.27603 0.02 0.985
## Year1998 -0.24131 0.30612 -0.79 0.431
## Year1999 -0.14476 0.26917 -0.54 0.591
```

```

## Year2000      -0.07767    0.25243   -0.31    0.759
## Year2001      0.06741    0.26375    0.26    0.798
## Year2002     -0.16071    0.25117   -0.64    0.523
## Year2003     -0.14260    0.33992   -0.42    0.675
## Year2004     -0.21579    0.25654   -0.84    0.401
## Year2005     -0.18253    0.27506   -0.66    0.507
## Year2006     -0.07040    0.25515   -0.28    0.783
## Year2007     -0.22853    0.24304   -0.94    0.348
## Year2008     -0.15648    0.24498   -0.64    0.523
## Year2009     -0.10218    0.26193   -0.39    0.697
## Year2010     -0.15385    0.24216   -0.64    0.526
## Year2011     -0.16754    0.24222   -0.69    0.490
## Year2012      0.06991    0.24767    0.28    0.778
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.422
## Multiple R-squared:  0.0823, Adjusted R-squared:  0.0107
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 36 weights are ~= 1. The remaining 269 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.244  0.870  0.947  0.898  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      3.28e-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 1.213 1      1.102
## LastAuthorFemale  1.442 1      1.201
## Year              1.637 16      1.016

```

Residuals from first and last author



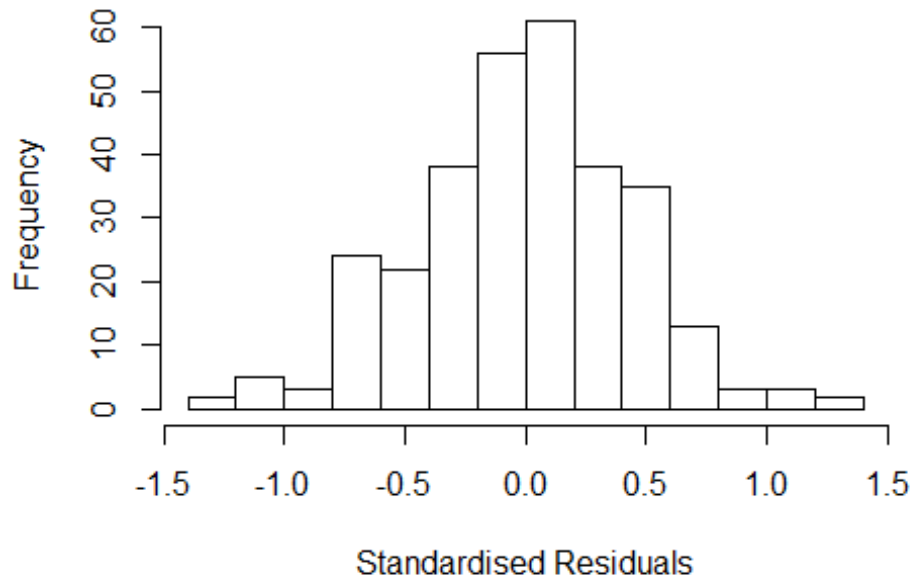
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3874 -0.2815 0.0127 0.2694 1.3904
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2820 0.2629 4.88 1.8e-06 ***
## FirstAuthorFemale1 0.0598 0.0528 1.13 0.259
## LastAuthorFemale1 -0.1608 0.0672 -2.39 0.017 *
## Year1997 -0.0385 0.3056 -0.13 0.900
## Year1998 -0.2967 0.3143 -0.94 0.346
## Year1999 -0.1622 0.2962 -0.55 0.584
## Year2000 -0.1300 0.2775 -0.47 0.640
## Year2001 0.0404 0.2914 0.14 0.890
## Year2002 -0.2006 0.2841 -0.71 0.481
## Year2003 -0.0700 0.3754 -0.19 0.852
## Year2004 -0.2717 0.2854 -0.95 0.342
## Year2005 -0.1860 0.2942 -0.63 0.528
```

```

## Year2006          -0.1034      0.2841   -0.36    0.716
## Year2007          -0.2563      0.2731   -0.94    0.349
## Year2008          -0.1809      0.2743   -0.66    0.510
## Year2009          -0.0951      0.2930   -0.32    0.746
## Year2010          -0.1927      0.2720   -0.71    0.479
## Year2011          -0.1881      0.2738   -0.69    0.492
## Year2012           0.0457      0.2781    0.16    0.870
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.424
## Multiple R-squared:  0.0619, Adjusted R-squared:  0.00291
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 31 weights are ~= 1. The remaining 274 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.261  0.868  0.949  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      3.28e-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
## -1.29412 -0.26635 0.00817 0.27241 1.28582
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.242804 0.231638 5.37 1.7e-07 ***
## FirstAuthorFemale1 0.041289 0.053216 0.78 0.44
## Year1997 -0.015504 0.267669 -0.06 0.95
## Year1998 -0.274483 0.300825 -0.91 0.36
## Year1999 -0.114551 0.266805 -0.43 0.67
## Year2000 -0.092824 0.247598 -0.37 0.71
## Year2001 0.000367 0.271735 0.00 1.00
## Year2002 -0.154457 0.253277 -0.61 0.54
## Year2003 -0.133511 0.346706 -0.39 0.70
## Year2004 -0.266524 0.255407 -1.04 0.30
## Year2005 -0.166993 0.255966 -0.65 0.51
## Year2006 -0.101911 0.258267 -0.39 0.69
```

```

## Year2007          -0.224682    0.240950   -0.93     0.35
## Year2008          -0.172006    0.245813   -0.70     0.48
## Year2009          -0.073274    0.264002   -0.28     0.78
## Year2010          -0.187623    0.242653   -0.77     0.44
## Year2011          -0.210506    0.244081   -0.86     0.39
## Year2012           0.010023    0.250866    0.04     0.97
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.432
## Multiple R-squared:  0.037, Adjusted R-squared:  -0.0201
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 280 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.350  0.865   0.956   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      3.28e-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 1.435 1      1.198
## Year              1.435 16      1.011
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.3537 -0.2972  0.0151  0.2633  1.4055

```



```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3014    0.2604    5.00  1e-06 ***
## LastAuthorFemale1 -0.1517    0.0673   -2.26  0.025 *
## Year1997          -0.0360    0.2985   -0.12  0.904
## Year1998          -0.2858    0.3122   -0.92  0.361
## Year1999          -0.1521    0.2904   -0.52  0.601
## Year2000          -0.1160    0.2732   -0.42  0.672
## Year2001           0.0499    0.2862    0.17  0.862
## Year2002          -0.1988    0.2787   -0.71  0.476
## Year2003          -0.0590    0.3849   -0.15  0.878
## Year2004          -0.2531    0.2789   -0.91  0.365
## Year2005          -0.1957    0.2915   -0.67  0.502
## Year2006          -0.0872    0.2786   -0.31  0.755
## Year2007          -0.2483    0.2679   -0.93  0.355
## Year2008          -0.1727    0.2701   -0.64  0.523
## Year2009          -0.0874    0.2897   -0.30  0.763
## Year2010          -0.1886    0.2681   -0.70  0.482
## Year2011          -0.1781    0.2694   -0.66  0.509
## Year2012           0.0523    0.2755    0.19  0.849
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.422
## Multiple R-squared:  0.0571, Adjusted R-squared:  0.0013
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 280 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.246  0.864  0.955  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      3.28e-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: 305"
## [1] ""

```

```

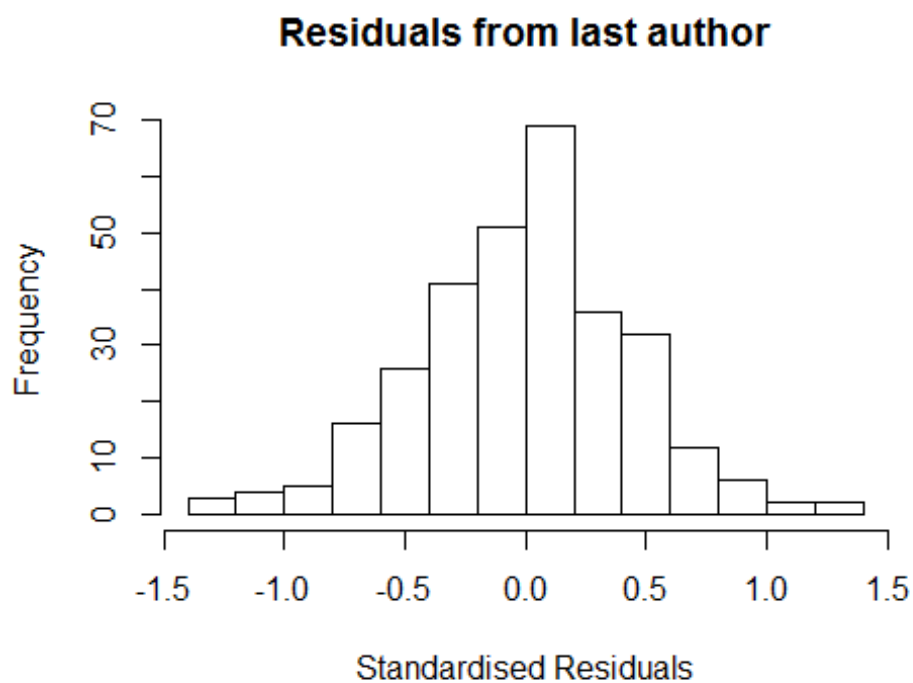
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2736"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 10 14 10 12 14 18 11 14 9 9 21 17 15 17 24
## 2011 2012
## 31 26
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1 3 5 2 7 5 5 7 6 7 10 12 13 11 12
## 2011 2012
## 18 18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1 2 5 2 6 5 4 4 4 5 7 12 8 9 10
## 2011 2012
## 15 16
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.37577213814232"
## [1] "Male first author team size 2018 geometric mean: 4.58233707506153"

## Warning in wilcox.test.default(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] "Female last author team size 2018 geometric mean: 4.70689796458981"
## [1] "Male last author team size 2018 geometric mean: 4.28682719924364"

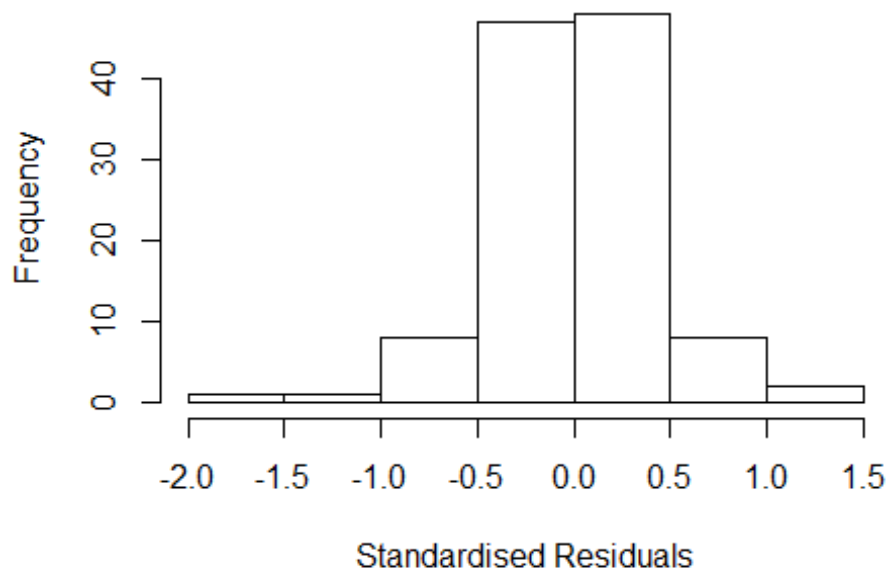
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 73, p-value = 0.5
## 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  5.745  1      2.397
## LastAuthorFemale   2.365  1      1.538
## UniqueAuthors    475.322  4      2.161
## Year              1949.906 16      1.267
```

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
## -1.5594 -0.2062  0.0152  0.2268  1.1918
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    5.99e-01   4.31e-08  1.39e+07 < 2e-16 ***
## FirstAuthorFemale1 -8.42e-02   1.10e-01 -7.60e-01  0.44658
## LastAuthorFemale1 -5.68e-02   1.03e-01 -5.50e-01  0.58433
## UniqueAuthors2    -1.06e-01   1.90e-01 -5.60e-01  0.57677
## UniqueAuthors3    -6.74e-02   1.50e-01 -4.50e-01  0.65547
## UniqueAuthors4    -4.34e-02   2.03e-01 -2.10e-01  0.83129
## UniqueAuthors5     2.02e-01   1.21e-01  1.67e+00  0.09869 .
## Year1997          1.51e+00   2.21e-01  6.83e+00  8.9e-10 ***
## Year1998         -2.86e-01   1.22e-01 -2.35e+00  0.02070 *
## Year1999          4.30e-01   1.23e-01  3.48e+00  0.00076 ***
```

```

## Year2000          6.63e-01    2.89e-01    2.29e+00    0.02417 *
## Year2001          2.77e-01    2.70e-01    1.03e+00    0.30770
## Year2002          9.27e-01    1.60e-01    5.79e+00    9.9e-08 ***
## Year2003          7.58e-01    2.07e-01    3.66e+00    0.00042 ***
## Year2004          7.93e-01    2.26e-01    3.51e+00    0.00069 ***
## Year2005          6.63e-01    1.80e-01    3.69e+00    0.00038 ***
## Year2006          6.11e-01    1.79e-01    3.42e+00    0.00095 ***
## Year2007          3.96e-01    1.93e-01    2.05e+00    0.04348 *
## Year2008          7.02e-01    1.55e-01    4.53e+00    1.8e-05 ***
## Year2009          6.98e-01    1.98e-01    3.53e+00    0.00065 ***
## Year2010          6.91e-01    1.47e-01    4.71e+00    8.6e-06 ***
## Year2011          7.05e-01    1.74e-01    4.05e+00    0.00011 ***
## Year2012          5.78e-01    1.62e-01    3.58e+00    0.00056 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.353
## Multiple R-squared:  0.426, Adjusted R-squared:  0.289
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 103 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0124 0.8640 0.9550 0.8850 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          8.70e-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 3.446 1          1.856
## LastAuthorFemale  2.123 1          1.457
## Year              6.809 16          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)

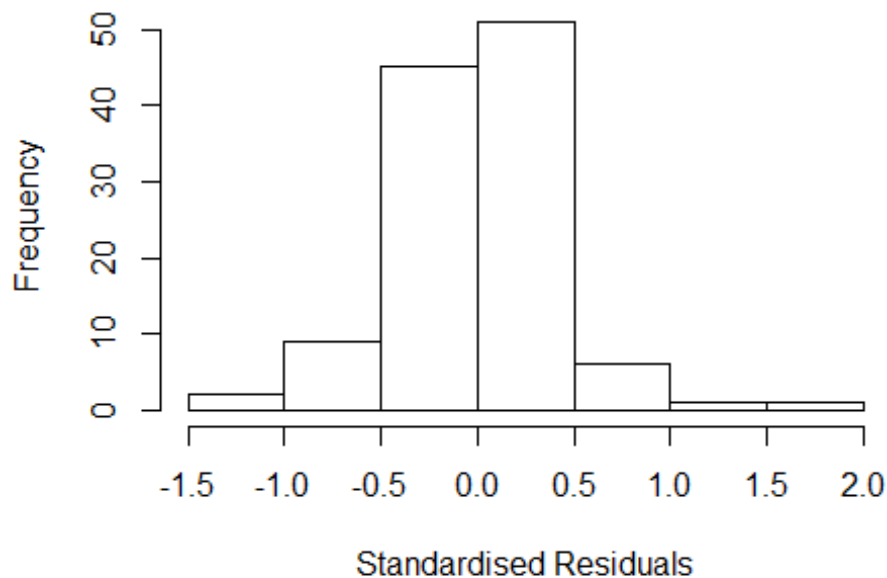
```

```
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.315 -0.239  0.014  0.259  1.842
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    5.99e-01   3.24e-08  1.85e+07 < 2e-16 ***
## FirstAuthorFemale1 -1.35e-02   1.00e-01 -1.30e-01  0.89328
## LastAuthorFemale1 -8.20e-02   1.07e-01 -7.60e-01  0.44623
## Year1997        -1.11e-01   5.21e-02 -2.12e+00  0.03622 *
## Year1998        -3.18e-01   1.06e-01 -3.00e+00  0.00346 **
## Year1999         4.08e-01   5.41e-02  7.54e+00  2.6e-11 ***
## Year2000         6.98e-01   2.23e-01  3.13e+00  0.00230 **
## Year2001         2.02e-01   2.19e-01  9.20e-01  0.35818
## Year2002         8.54e-01   8.27e-02  1.03e+01 < 2e-16 ***
## Year2003         7.16e-01   2.14e-01  3.35e+00  0.00116 **
## Year2004         7.42e-01   1.64e-01  4.52e+00  1.8e-05 ***
## Year2005         7.00e-01   2.04e-01  3.44e+00  0.00087 ***
## Year2006         6.52e-01   1.47e-01  4.45e+00  2.3e-05 ***
## Year2007         3.97e-01   1.33e-01  2.99e+00  0.00352 **
## Year2008         7.56e-01   1.14e-01  6.62e+00  2.1e-09 ***
## Year2009         7.24e-01   1.31e-01  5.54e+00  2.6e-07 ***
## Year2010         7.82e-01   1.39e-01  5.61e+00  2.0e-07 ***
## Year2011         7.29e-01   1.46e-01  4.98e+00  2.8e-06 ***
## Year2012         6.91e-01   1.21e-01  5.72e+00  1.2e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.405
## Multiple R-squared:  0.338, Adjusted R-squared:  0.214
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 102 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0034  0.8910  0.9530  0.9000  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      8.70e-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"

## 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 3.279 1      1.811
## Year              3.279 16      1.038

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
```

```

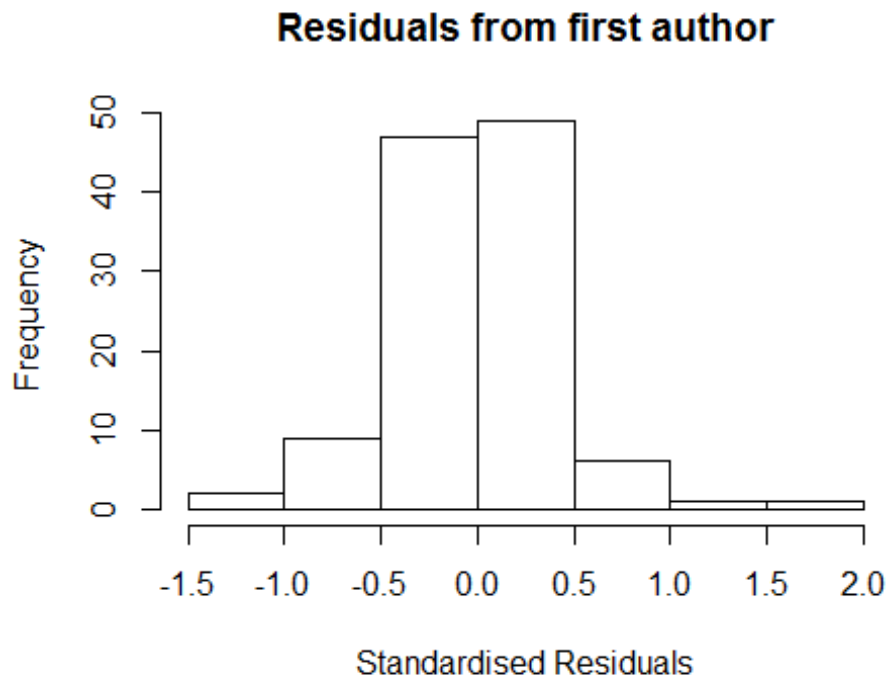
##      Min      1Q  Median      3Q      Max
## -1.295 -0.249  0.000  0.246  1.845
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.5990     0.0000      Inf < 2e-16 ***
## FirstAuthorFemale1 -0.0239     0.1015    -0.24  0.81424
## Year1997        -0.1140     0.0333    -3.43  0.00090 ***
## Year1998        -0.3503     0.0985    -3.55  0.00059 ***
## Year1999         0.4080     0.0541     7.54  2.5e-11 ***
## Year2000         0.6976     0.2236     3.12  0.00238 **
## Year2001         0.1935     0.2078     0.93  0.35422
## Year2002         0.8540     0.0828    10.32 < 2e-16 ***
## Year2003         0.7061     0.2004     3.52  0.00065 ***
## Year2004         0.7307     0.1701     4.30  4.1e-05 ***
## Year2005         0.7084     0.2025     3.50  0.00071 ***
## Year2006         0.6324     0.1338     4.73  7.7e-06 ***
## Year2007         0.3725     0.1177     3.16  0.00208 **
## Year2008         0.7512     0.1181     6.36  6.6e-09 ***
## Year2009         0.7100     0.1358     5.23  9.8e-07 ***
## Year2010         0.7474     0.1204     6.21  1.3e-08 ***
## Year2011         0.7196     0.1517     4.74  7.2e-06 ***
## Year2012         0.6803     0.1247     5.46  3.7e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.336, Adjusted R-squared:  0.219
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~ = 1. The remaining 100 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0016 0.8880 0.9570 0.8970 0.9810 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier          eps.x
##      1.00e-07          1.00e-07      8.70e-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"

## 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 2.999 1          1.732
## Year             2.999 16         1.035

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.32006 -0.24133  0.00894  0.25859  1.84055
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.5990    0.0000    Inf < 2e-16 ***
## LastAuthorFemale1 -0.0838    0.1084   -0.77  0.44103
```

```

## Year1997      -0.1095      0.0578      -1.89      0.06119 .
## Year1998      -0.3174      0.1064      -2.98      0.00360 **
## Year1999       0.4080      0.0541       7.54      2.5e-11 ***
## Year2000       0.6980      0.2226       3.14      0.00227 **
## Year2001       0.2005      0.2180       0.92      0.35999
## Year2002       0.8540      0.0827      10.32      < 2e-16 ***
## Year2003       0.7030      0.1875       3.75      0.00030 ***
## Year2004       0.7286      0.1307       5.58      2.2e-07 ***
## Year2005       0.6884      0.1808       3.81      0.00025 ***
## Year2006       0.6465      0.1345       4.81      5.6e-06 ***
## Year2007       0.3904      0.1184       3.30      0.00137 **
## Year2008       0.7510      0.1110       6.76      1.0e-09 ***
## Year2009       0.7172      0.1259       5.70      1.3e-07 ***
## Year2010       0.7774      0.1319       5.89      5.5e-08 ***
## Year2011       0.7211      0.1268       5.69      1.4e-07 ***
## Year2012       0.6867      0.1277       5.38      5.2e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.406
## Multiple R-squared:  0.337, Adjusted R-squared:  0.221
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 101 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.004  0.888  0.952  0.900  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           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 2737"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last

```

```

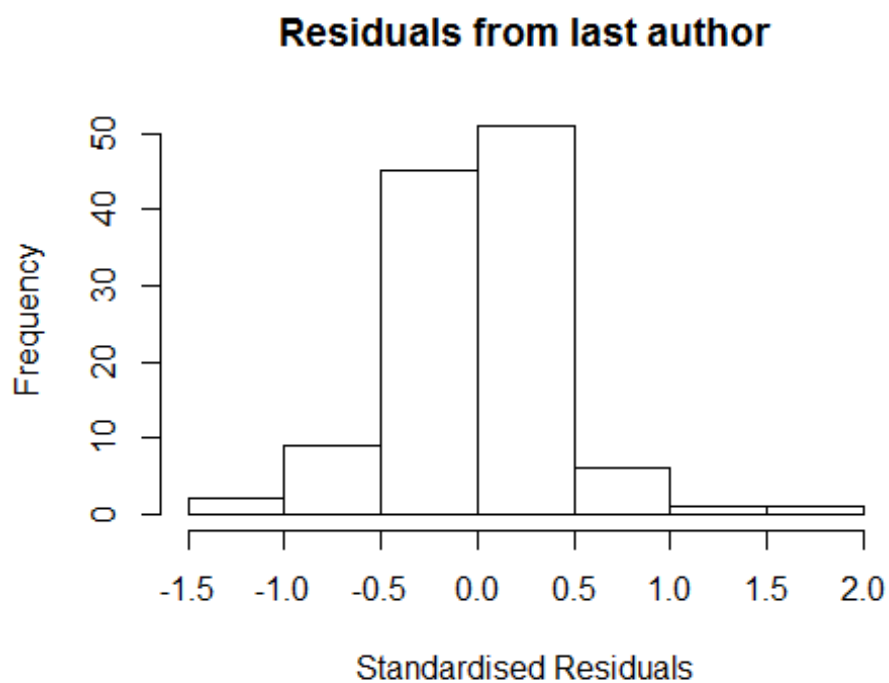
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    2    8    6   11    4    7    6    6    5    1    5   12   20   19
## 2011 2012
##   30   13
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    6    3    6    1    3    2    5    4    1    3    5   15   10
## 2011 2012
##   23   13
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    6    3    4    1    2    2    3    4    1    3    5   11    8
## 2011 2012
##   19   12
## [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: 3.92851190016586"

## Warning in wilcox.test.default(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 = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.09534502215844"
## [1] "Male last author team size 2018 geometric mean: 4.17033027746852"

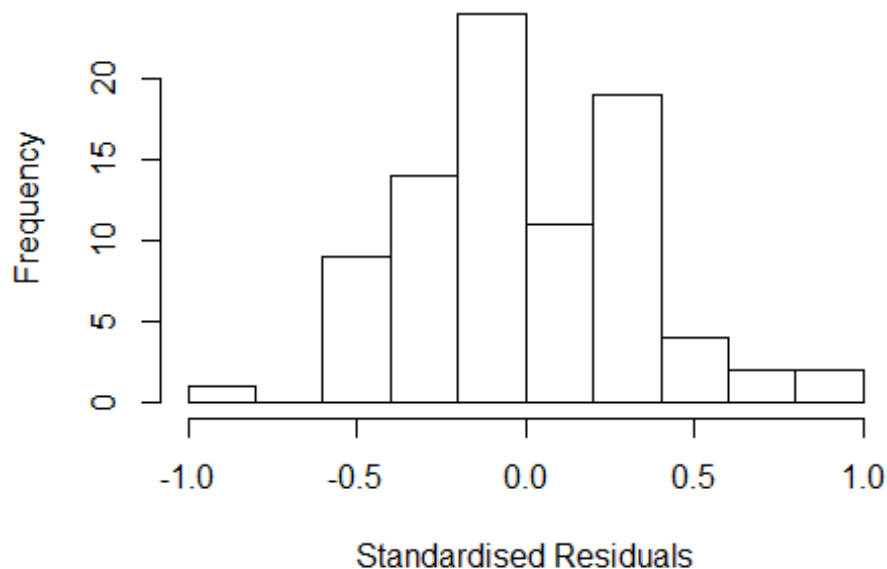
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 27, 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 -4.641e+13  1          NaN
## LastAuthorFemale  5.719e+00  1          2.392
## UniqueAuthors    -3.302e+29  4          NaN
## Year              -1.241e+30 16          NaN
```

Residuals from first and last author and team size



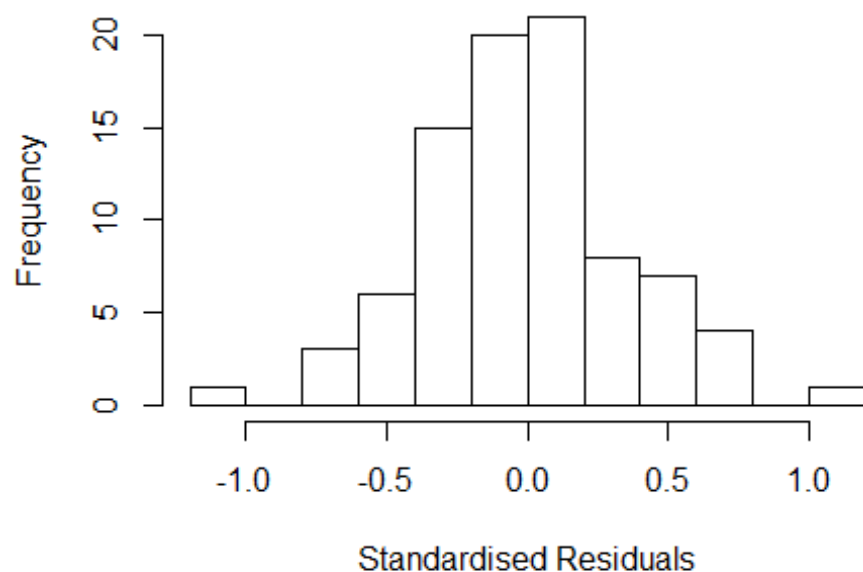
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.93885 -0.22455 -0.00288 0.23616 0.93873
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.36583 0.53782 2.54 0.01358 *
## FirstAuthorFemale1 -0.00495 0.09622 -0.05 0.95914
## LastAuthorFemale1 0.10803 0.11454 0.94 0.34923
## UniqueAuthors2 0.02212 0.51874 0.04 0.96613
## UniqueAuthors3 0.04224 0.52524 0.08 0.93616
## UniqueAuthors4 0.32437 0.54610 0.59 0.55465
## UniqueAuthors5 0.37255 0.50553 0.74 0.46389
## Year1997 -0.80712 0.13596 -5.94 1.4e-07 ***
## Year1998 -0.68999 0.13200 -5.23 2.1e-06 ***
## Year1999 -0.15472 0.54244 -0.29 0.77641
```

```

## Year2000      -0.27546    0.23584   -1.17  0.24720
## Year2001      -0.24239    0.15627   -1.55  0.12590
## Year2002      -0.77122    0.28842   -2.67  0.00954 **
## Year2003      -0.86174    0.18183   -4.74  1.3e-05 ***
## Year2004      -0.62627    0.19181   -3.27  0.00177 **
## Year2005      -0.50398    0.26536   -1.90  0.06212 .
## Year2006      -0.19339    0.15627   -1.24  0.22050
## Year2007       0.03025    0.20667    0.15  0.88410
## Year2008      -0.58841    0.10846   -5.43  9.8e-07 ***
## Year2009      -0.74815    0.13773   -5.43  9.6e-07 ***
## Year2010      -0.35119    0.15096   -2.33  0.02323 *
## Year2011      -0.42698    0.13915   -3.07  0.00317 **
## Year2012      -0.64094    0.15589   -4.11  0.00012 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.387
## Multiple R-squared:  0.346, Adjusted R-squared:  0.118
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 74 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.536  0.908  0.955  0.923  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      1.16e-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 2.095e+15 1 4.578e+07
## LastAuthorFemale 6.392e+00 1 2.528e+00
## Year -1.616e+16 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
## -1.07e+00 -2.77e-01  6.25e-16  1.87e-01  1.11e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.37e+00   1.04e-01  1.31e+01 < 2e-16 ***
## FirstAuthorFemale1  1.49e-02   1.04e-01  1.40e-01  0.887
## LastAuthorFemale1  7.48e-02   1.32e-01  5.70e-01  0.572
## Year1997         -7.87e-01   2.75e-08 -2.86e+07 < 2e-16 ***
## Year1998         -5.99e-01   1.39e-01 -4.32e+00 5.3e-05 ***
## Year1999          1.62e-01   3.03e-01  5.30e-01  0.595
## Year2000         -1.04e-01   1.24e-01 -8.40e-01  0.405
## Year2001          1.28e-01   1.04e-01  1.23e+00  0.223
## Year2002         -4.32e-01   2.77e-01 -1.56e+00  0.124
## Year2003         -5.33e-01   1.05e-01 -5.06e+00 3.5e-06 ***
## Year2004         -2.62e-01   1.49e-01 -1.76e+00  0.083 .
## Year2005         -3.16e-01   2.70e-01 -1.17e+00  0.246
```

```

## Year2006          1.77e-01   1.04e-01   1.70e+00   0.094 .
## Year2007          2.82e-01   2.64e-01   1.06e+00   0.291
## Year2008         -3.64e-01   1.49e-01  -2.44e+00   0.017 *
## Year2009         -5.18e-01   1.05e-01  -4.94e+00  5.4e-06 ***
## Year2010         -2.09e-01   1.62e-01  -1.29e+00   0.201
## Year2011         -2.95e-01   1.59e-01  -1.86e+00   0.067 .
## Year2012         -3.30e-01   1.50e-01  -2.20e+00   0.032 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.412
## Multiple R-squared:  0.245, Adjusted R-squared:  0.0421
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 74 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.448  0.909  0.954  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      1.16e-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 1.081e+13 1      3.288e+06
## Year              1.081e+13 16      2.555e+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

```



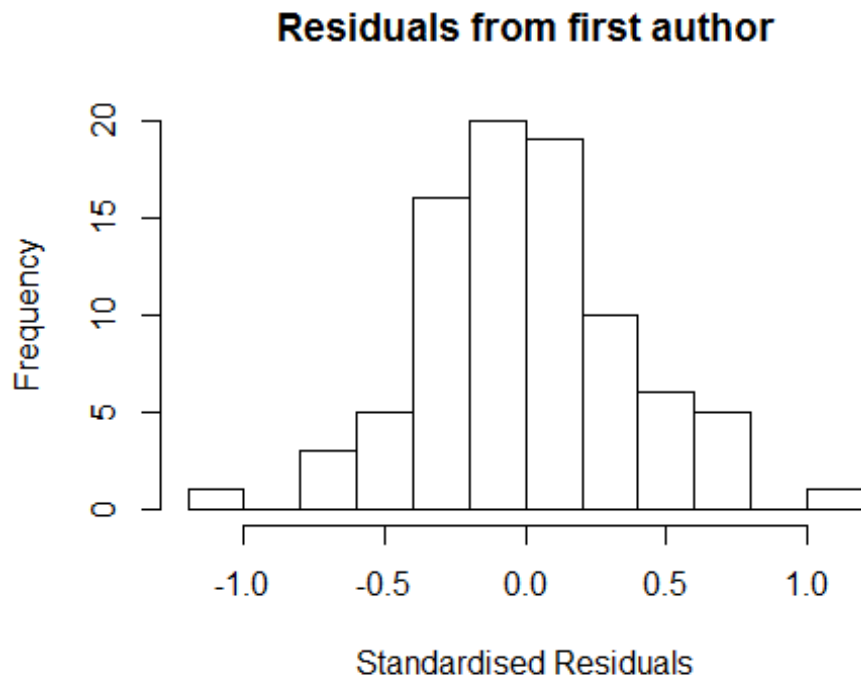
```

## -1.07e+00 -2.79e-01 1.67e-16 2.16e-01 1.08e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.34e+00   9.52e-02  1.41e+01 < 2e-16 ***
## FirstAuthorFemale1 3.88e-02   9.52e-02  4.10e-01  0.684
## Year1997       -7.87e-01   4.25e-08 -1.85e+07 < 2e-16 ***
## Year1998       -5.61e-01   1.14e-01 -4.93e+00 5.5e-06 ***
## Year1999        1.80e-01   3.03e-01  5.90e-01  0.555
## Year2000       -7.29e-02   1.18e-01 -6.20e-01  0.539
## Year2001        1.52e-01   9.52e-02  1.60e+00  0.115
## Year2002       -3.71e-01   2.87e-01 -1.29e+00  0.201
## Year2003       -4.84e-01   6.91e-02 -6.99e+00 1.5e-09 ***
## Year2004       -2.46e-01   1.50e-01 -1.64e+00  0.105
## Year2005       -2.92e-01   2.59e-01 -1.13e+00  0.263
## Year2006        2.01e-01   9.52e-02  2.11e+00  0.039 *
## Year2007        2.99e-01   2.66e-01  1.13e+00  0.264
## Year2008       -3.50e-01   1.48e-01 -2.36e+00  0.021 *
## Year2009       -5.00e-01   1.01e-01 -4.97e+00 4.7e-06 ***
## Year2010       -1.57e-01   1.35e-01 -1.17e+00  0.247
## Year2011       -2.70e-01   1.51e-01 -1.78e+00  0.079 .
## Year2012       -2.99e-01   1.36e-01 -2.20e+00  0.031 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.43
## Multiple R-squared:  0.238, Adjusted R-squared:  0.0477
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 78 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.506 0.919 0.960 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           1.16e-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"

```

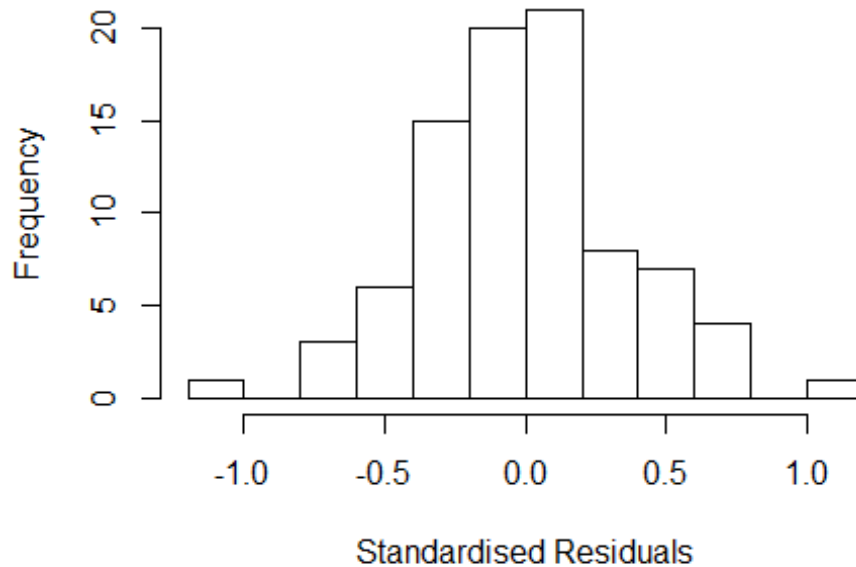
```
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

## 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



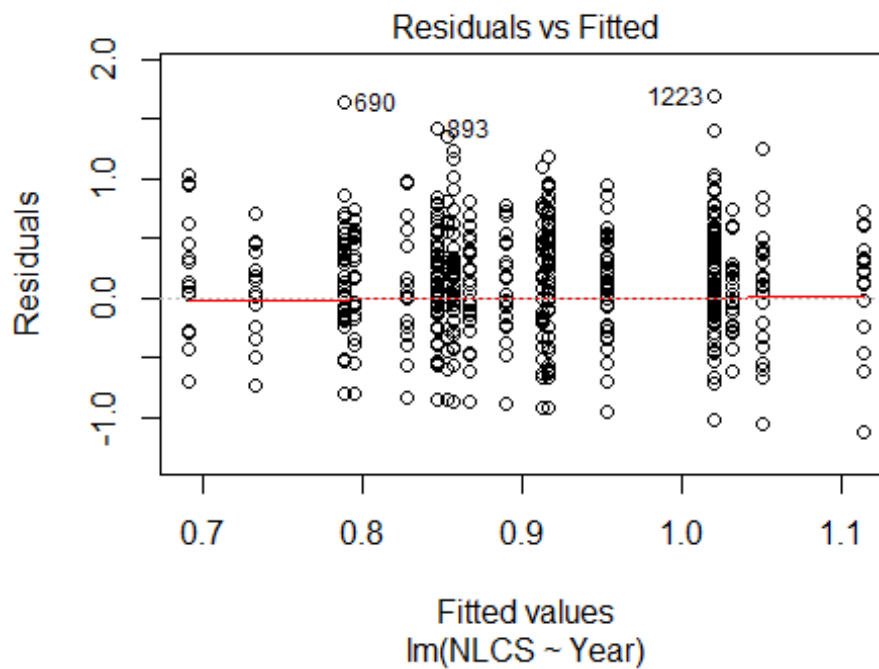
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.08e+00 -2.74e-01 -3.82e-16 1.93e-01 1.12e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.38e+00 2.96e-08 4.67e+07 < 2e-16 ***
## LastAuthorFemale1 8.31e-02 1.19e-01 7.00e-01 0.4859
## Year1997 -7.87e-01 0.00e+00 -Inf < 2e-16 ***
## Year1998 -6.09e-01 1.22e-01 -5.00e+00 4.3e-06 ***
## Year1999 1.51e-01 2.82e-01 5.40e-01 0.5928
## Year2000 -1.14e-01 9.32e-02 -1.22e+00 0.2258
## Year2001 1.13e-01 2.77e-08 4.08e+06 < 2e-16 ***
## Year2002 -4.52e-01 2.40e-01 -1.88e+00 0.0645 .
## Year2003 -5.45e-01 6.85e-02 -7.95e+00 2.8e-11 ***
## Year2004 -2.72e-01 1.35e-01 -2.01e+00 0.0484 *
## Year2005 -3.31e-01 2.48e-01 -1.33e+00 0.1869
## Year2006 1.62e-01 2.22e-08 7.28e+06 < 2e-16 ***
```

```

## Year2007          2.71e-01    2.47e-01    1.10e+00    0.2759
## Year2008          -3.73e-01    1.35e-01   -2.76e+00    0.0074 **
## Year2009          -5.23e-01    9.51e-02   -5.50e+00    6.2e-07 ***
## Year2010          -2.26e-01    1.20e-01   -1.88e+00    0.0638 .
## Year2011          -3.07e-01    1.26e-01   -2.44e+00    0.0173 *
## Year2012          -3.40e-01    1.29e-01   -2.63e+00    0.0105 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.406
## Multiple R-squared:  0.246, Adjusted R-squared:  0.0577
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 75 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.425  0.906   0.952   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      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 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
##   34   46   43   48   37   49   36   21   48   48   50   58   61   77   87
## 2011 2012
##   94  124
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   19   25   22   27   19   25   24   16   39   27   40   45   47   56   70
## 2011 2012

```

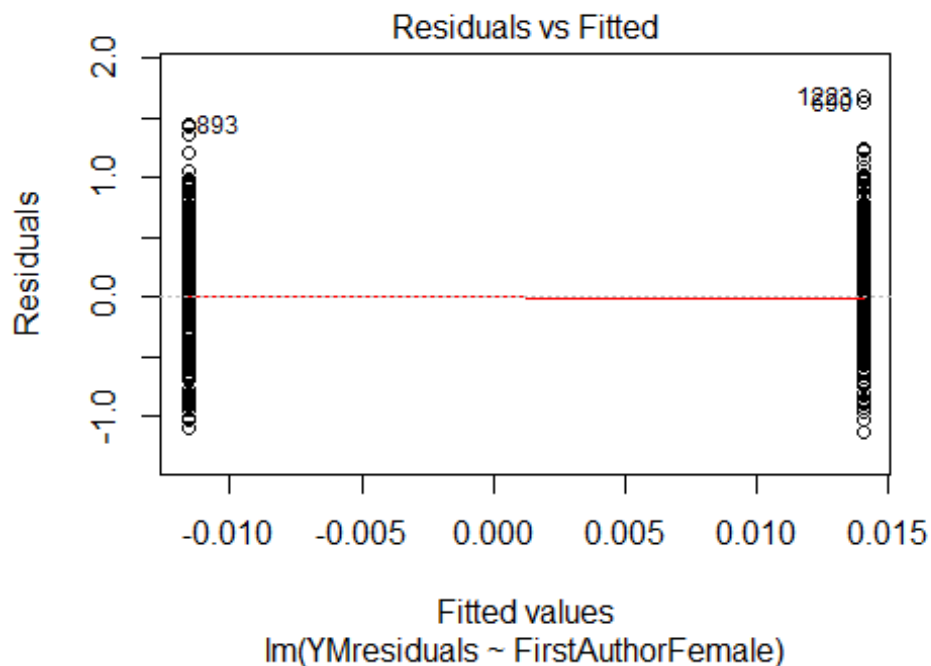
```
## 71 100
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 17 22 19 24 17 23 20 15 31 20 30 38 40 51 61
## 2011 2012
## 66 84
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 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: 4.30589025547354"
## [1] "Male first author team size 2018 geometric mean: 3.23781445507615"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 640, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.8212648761683"
## [1] "Male last author team size 2018 geometric mean: 3.76638352222714"

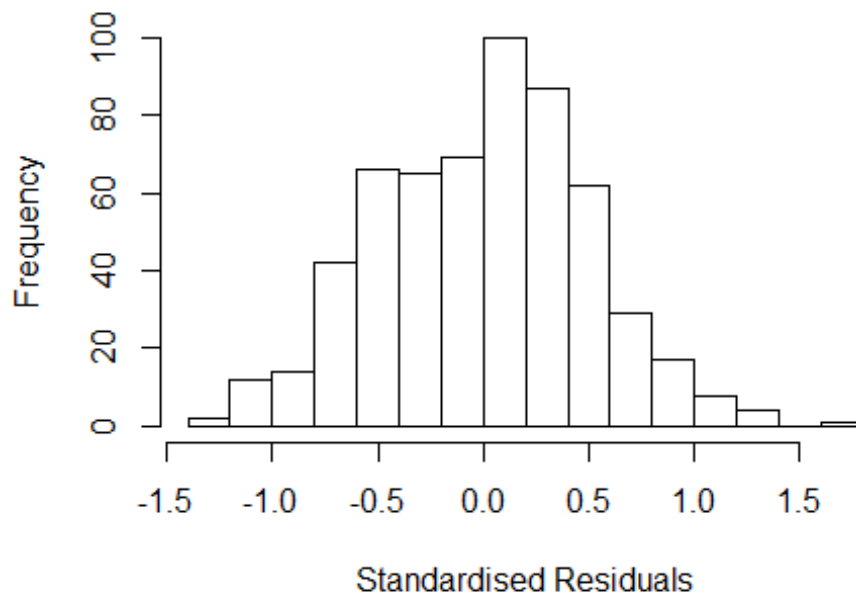
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 510, 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 | 1.232 | 1 | 1.110 |
| LastAuthorFemale | 1.177 | 1 | 1.085 |
| UniqueAuthors | 1.833 | 4 | 1.079 |
| Year | 1.895 | 16 | 1.020 |

Residuals from first and last author and team size



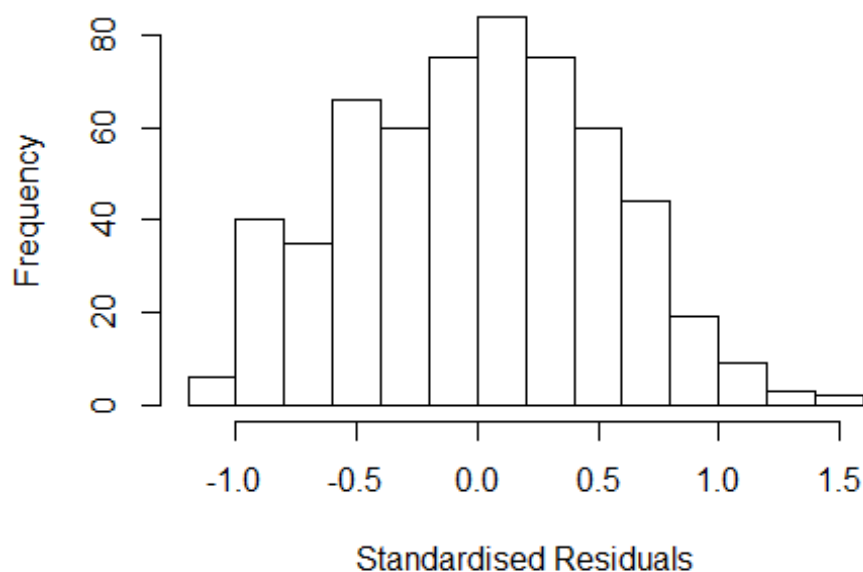
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2830 -0.3607  0.0403  0.3536  1.6453
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.54813    0.11648   4.71 3.2e-06 ***
## FirstAuthorFemale1 -0.03703    0.04703  -0.79  0.431
## LastAuthorFemale1 -0.00727    0.05235  -0.14  0.890
## UniqueAuthors2     0.15511    0.06967   2.23  0.026 *
## UniqueAuthors3     0.14682    0.07292   2.01  0.045 *
## UniqueAuthors4     0.37502    0.09341   4.01 6.8e-05 ***
## UniqueAuthors5     0.46130    0.07110   6.49 1.9e-10 ***
## Year1997          0.04051    0.16274   0.25  0.803
## Year1998         -0.08434    0.16683  -0.51  0.613
## Year1999          0.14840    0.14625   1.01  0.311
```

```

## Year2000          0.31786    0.19459    1.63    0.103
## Year2001          0.28160    0.14091    2.00    0.046 *
## Year2002          0.03400    0.16656    0.20    0.838
## Year2003         -0.02082    0.15776   -0.13    0.895
## Year2004          0.06136    0.13732    0.45    0.655
## Year2005          0.18610    0.16740    1.11    0.267
## Year2006          0.18049    0.14489    1.25    0.213
## Year2007          0.27306    0.13659    2.00    0.046 *
## Year2008          0.03157    0.14061    0.22    0.822
## Year2009          0.02368    0.13790    0.17    0.864
## Year2010          0.07962    0.13684    0.58    0.561
## Year2011          0.13356    0.14705    0.91    0.364
## Year2012          0.20830    0.12738    1.64    0.103
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.515
## Multiple R-squared:  0.131, Adjusted R-squared:  0.0965
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 53 weights are ~= 1. The remaining 525 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.286  0.882  0.948  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      1.73e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.107 1      1.052
## LastAuthorFemale  1.112 1      1.055
## Year              1.181 16      1.005

```


Residuals from first and last author



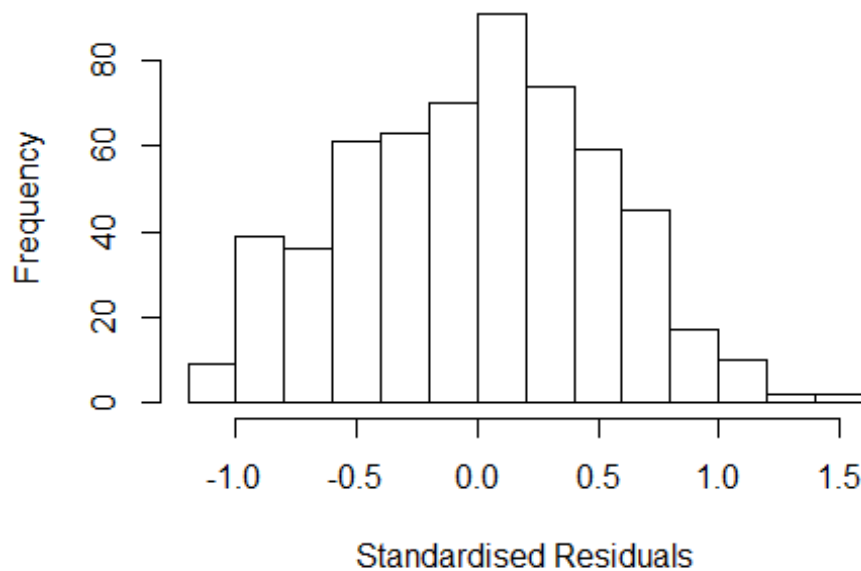
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0880 -0.4037 0.0184 0.3750 1.4474
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.76389 0.13228 5.77 1.3e-08 ***
## FirstAuthorFemale1 -0.00452 0.04753 -0.10 0.924
## LastAuthorFemale1 -0.03626 0.05333 -0.68 0.497
## Year1997 0.04282 0.17884 0.24 0.811
## Year1998 -0.13927 0.18978 -0.73 0.463
## Year1999 0.11528 0.16286 0.71 0.479
## Year2000 0.32415 0.18605 1.74 0.082 .
## Year2001 0.29797 0.14904 2.00 0.046 *
## Year2002 0.03026 0.18424 0.16 0.870
## Year2003 -0.03900 0.17115 -0.23 0.820
## Year2004 0.07442 0.16212 0.46 0.646
## Year2005 0.21080 0.18183 1.16 0.247
```

```

## Year2006          0.16795    0.16315    1.03    0.304
## Year2007          0.20675    0.15381    1.34    0.179
## Year2008         -0.03101    0.15634   -0.20    0.843
## Year2009          0.08984    0.15427    0.58    0.561
## Year2010          0.09289    0.14905    0.62    0.533
## Year2011          0.15938    0.16023    0.99    0.320
## Year2012          0.25499    0.14596    1.75    0.081 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.566
## Multiple R-squared:  0.0392, Adjusted R-squared:  0.00828
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 50 weights are ~= 1. The remaining 528 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.493  0.884  0.950  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      1.73e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
##   nResample    max.it    best.r.s    k.fast.s    k.max maxit.scale
##      500         50         2         1      1000         200
##   trace.lev    mts    compute.rd
##      0        1000         0
##           psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## 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



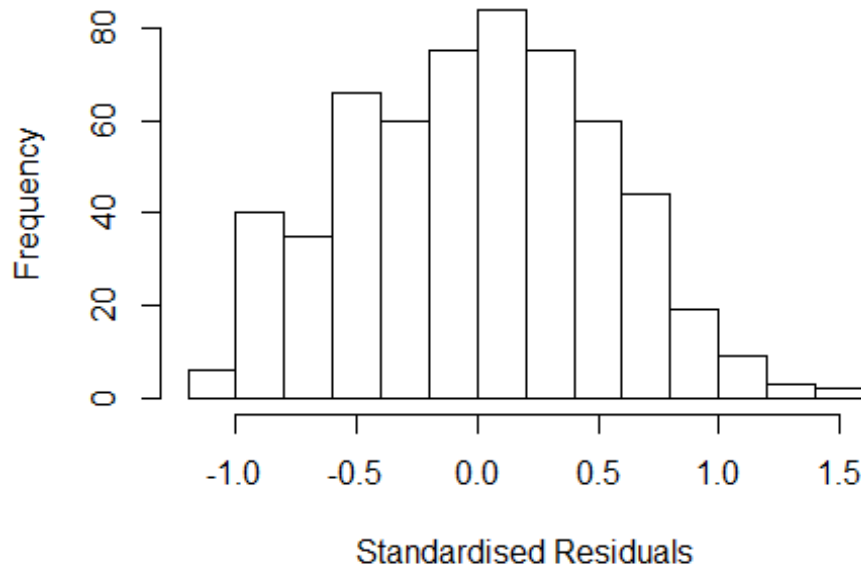
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.085 -0.409 0.016 0.376 1.428
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.75998 0.13139 5.78 1.2e-08 ***
## FirstAuthorFemale1 -0.00977 0.04717 -0.21 0.836
## Year1997 0.04249 0.17807 0.24 0.811
## Year1998 -0.14395 0.18955 -0.76 0.448
## Year1999 0.11383 0.16296 0.70 0.485
## Year2000 0.32549 0.18617 1.75 0.081 .
## Year2001 0.29158 0.14767 1.97 0.049 *
## Year2002 0.02994 0.18289 0.16 0.870
## Year2003 -0.03500 0.17107 -0.20 0.838
## Year2004 0.06829 0.16119 0.42 0.672
## Year2005 0.20905 0.18145 1.15 0.250
## Year2006 0.16673 0.16339 1.02 0.308
```

```

## Year2007          0.20112      0.15244      1.32      0.188
## Year2008          -0.03591     0.15558     -0.23      0.818
## Year2009          0.08435     0.15331      0.55      0.582
## Year2010          0.08543     0.14774      0.58      0.563
## Year2011          0.15011     0.15867      0.95      0.345
## Year2012          0.25036     0.14542      1.72      0.086 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.565
## Multiple R-squared:  0.0385, Adjusted R-squared:  0.00928
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 49 weights are ~= 1. The remaining 529 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.503  0.883  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      1.73e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.093 1          1.046
## Year              1.093 16          1.003

```

Residuals from last author



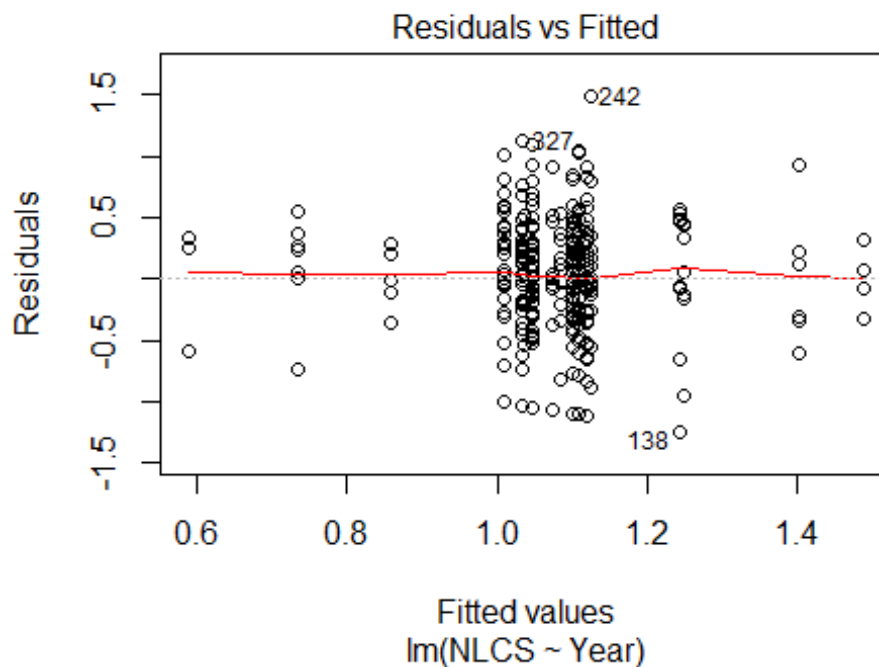
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0864 -0.4055 0.0196 0.3751 1.4504
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7622 0.1314 5.80 1.1e-08 ***
## LastAuthorFemale1 -0.0371 0.0529 -0.70 0.484
## Year1997 0.0437 0.1789 0.24 0.807
## Year1998 -0.1388 0.1898 -0.73 0.465
## Year1999 0.1152 0.1627 0.71 0.479
## Year2000 0.3242 0.1861 1.74 0.082 .
## Year2001 0.2977 0.1491 2.00 0.046 *
## Year2002 0.0298 0.1842 0.16 0.872
## Year2003 -0.0393 0.1711 -0.23 0.818
## Year2004 0.0738 0.1617 0.46 0.648
## Year2005 0.2106 0.1817 1.16 0.247
## Year2006 0.1678 0.1631 1.03 0.304
```

```

## Year2007          0.2068      0.1537      1.35      0.179
## Year2008         -0.0309      0.1563     -0.20      0.843
## Year2009          0.0899      0.1542      0.58      0.560
## Year2010          0.0927      0.1489      0.62      0.534
## Year2011          0.1592      0.1601      0.99      0.320
## Year2012          0.2545      0.1456      1.75      0.081 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.566
## Multiple R-squared:  0.0392, Adjusted R-squared:  0.01
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 50 weights are ~= 1. The remaining 528 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.491  0.885   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      1.73e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 578"
## [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
##   11   13   11   21   19   24   12   18   23   24   22   49   48   55   70
## 2011 2012
##   68   69
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    6    5    7    4    8    6    8   12   16   10   35   32   36   43
## 2011 2012

```

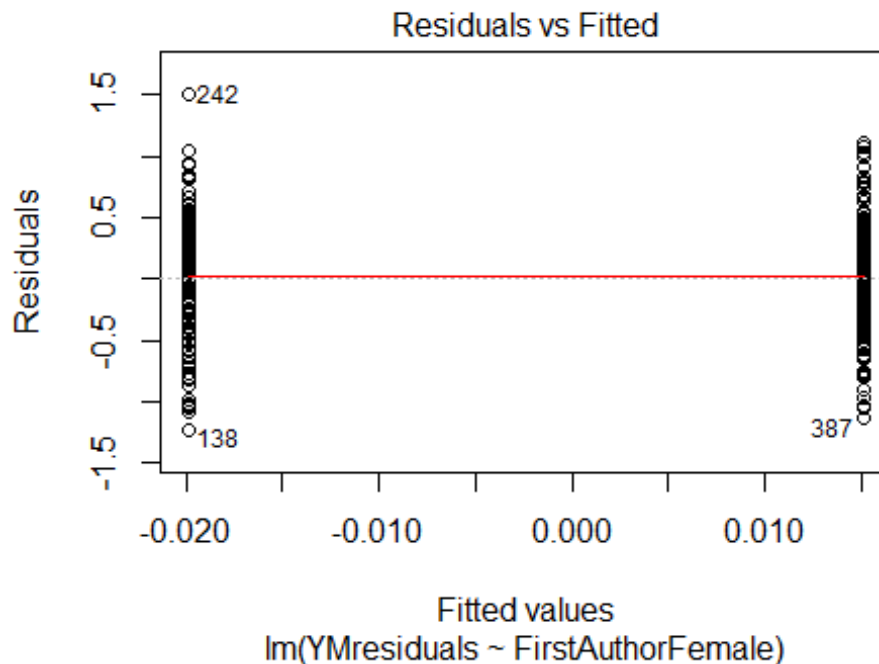
```
## 46 43
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3 4 4 5 2 8 3 7 7 14 7 31 28 27 38
## 2011 2012
## 36 41
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 12, df = 16, p-value = 0.8
```



```
##
## 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: 2.81269119147023"
## [1] "Male first author team size 2018 geometric mean: 2.25139026595515"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

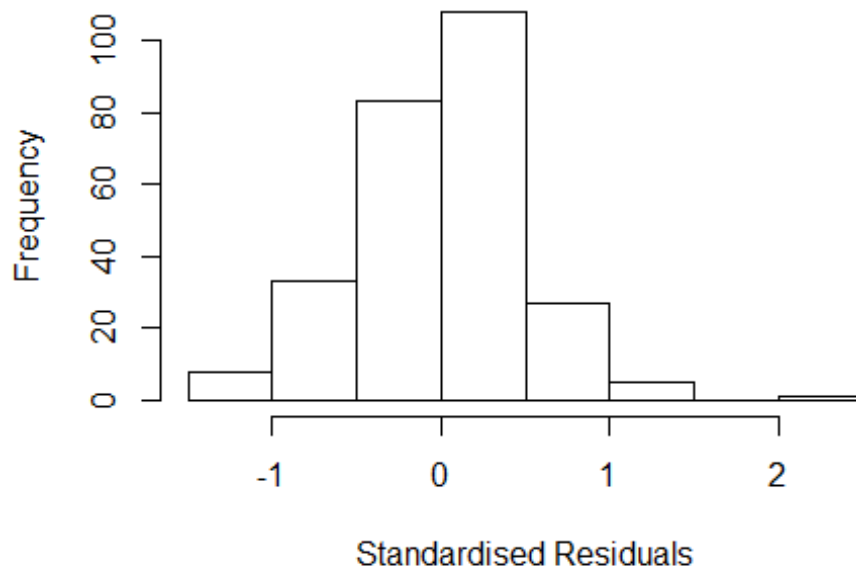
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 350, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.71870536205653"
## [1] "Male last author team size 2018 geometric mean: 2.5572158810111"

## Warning in wilcox.test.default(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 = 0.9
## 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 1.541  1      1.241
## LastAuthorFemale  1.371  1      1.171
## UniqueAuthors    4.123  4      1.194
## Year              5.547 16      1.055
```


Residuals from first and last author and team size



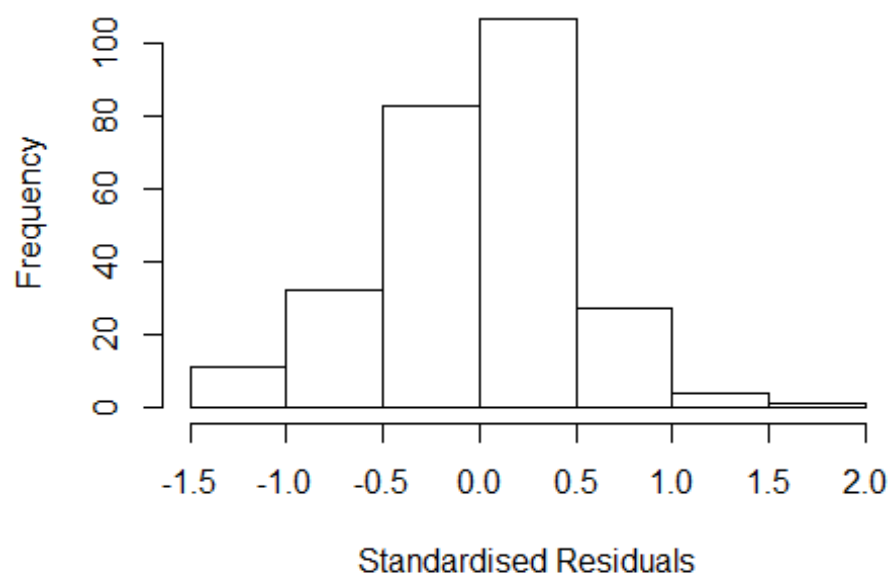
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.167 -0.317 0.033 0.302 2.078
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.2027 0.3078 0.66 0.51096
## FirstAuthorFemale1 0.0478 0.0707 0.68 0.49996
## LastAuthorFemale1 0.0571 0.0656 0.87 0.38534
## UniqueAuthors2 0.3341 0.1227 2.72 0.00695 **
## UniqueAuthors3 0.4020 0.1214 3.31 0.00107 **
## UniqueAuthors4 0.4676 0.1204 3.88 0.00013 ***
## UniqueAuthors5 0.3529 0.1195 2.95 0.00345 **
## Year1997 0.5395 0.3752 1.44 0.15175
## Year1998 0.5769 0.3281 1.76 0.07991 .
## Year1999 0.5759 0.3561 1.62 0.10710
```

```

## Year2000          1.1790      0.3924      3.00  0.00294 **
## Year2001          0.8784      0.3405      2.58  0.01049 *
## Year2002          0.2987      0.3290      0.91  0.36484
## Year2003          0.2450      0.3342      0.73  0.46424
## Year2004          0.4880      0.3041      1.60  0.10987
## Year2005          0.3275      0.3108      1.05  0.29299
## Year2006          0.6087      0.3939      1.55  0.12362
## Year2007          0.4346      0.2944      1.48  0.14123
## Year2008          0.5584      0.3059      1.83  0.06918 .
## Year2009          0.4986      0.3027      1.65  0.10084
## Year2010          0.5063      0.2982      1.70  0.09088 .
## Year2011          0.5202      0.2924      1.78  0.07655 .
## Year2012          0.4263      0.2891      1.47  0.14156
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.476
## Multiple R-squared:  0.144, Adjusted R-squared:  0.0665
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 242 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0171 0.8780 0.9570 0.9060 0.9850 0.9990
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          3.77e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.533 1          1.238
## LastAuthorFemale  1.192 1          1.092
## Year              1.663 16          1.016

```

Residuals from first and last author



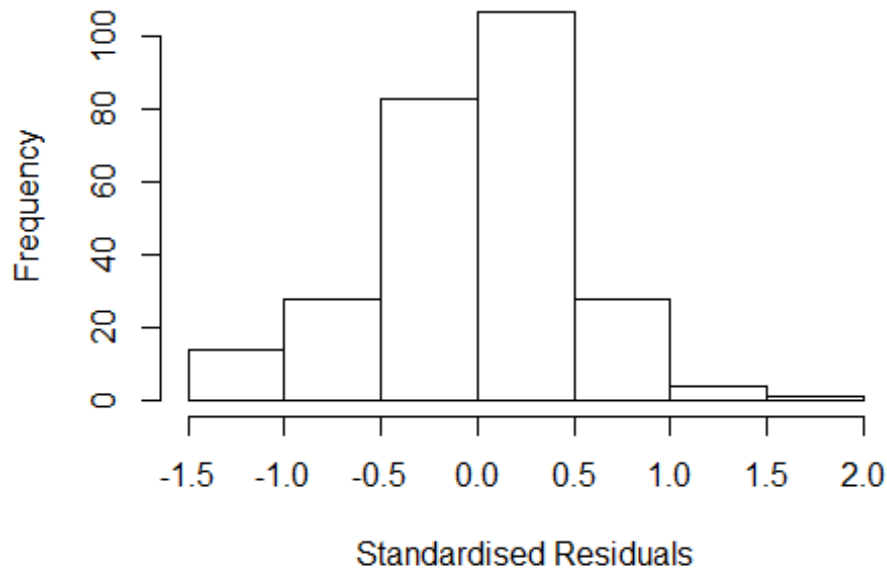
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2893 -0.3296 0.0347 0.3180 1.6966
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5731 0.2756 2.08 0.039 *
## FirstAuthorFemale1 0.0668 0.0738 0.91 0.366
## LastAuthorFemale1 0.0379 0.0654 0.58 0.563
## Year1997 0.5229 0.3161 1.65 0.099 .
## Year1998 0.2067 0.3001 0.69 0.492
## Year1999 0.6336 0.3444 1.84 0.067 .
## Year2000 0.8087 0.3729 2.17 0.031 *
## Year2001 0.7162 0.3557 2.01 0.045 *
## Year2002 0.1962 0.2974 0.66 0.510
## Year2003 0.0794 0.3341 0.24 0.812
## Year2004 0.4982 0.3050 1.63 0.104
## Year2005 0.3383 0.3060 1.11 0.270
```

```

## Year2006          0.4869      0.3927      1.24      0.216
## Year2007          0.4004      0.2855      1.40      0.162
## Year2008          0.5104      0.2972      1.72      0.087 .
## Year2009          0.3669      0.3097      1.18      0.237
## Year2010          0.4494      0.2908      1.55      0.123
## Year2011          0.4827      0.2834      1.70      0.090 .
## Year2012          0.4188      0.2872      1.46      0.146
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.49
## Multiple R-squared:  0.0612, Adjusted R-squared:  -0.00748
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 245 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.206  0.876  0.955  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      3.77e-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 1.443 1      1.201
## Year              1.443 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
## -1.2963 -0.3145 0.0294 0.3186 1.6888
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6114 0.2680 2.28 0.023 *
## FirstAuthorFemale1 0.0776 0.0721 1.08 0.283
## Year1997 0.5014 0.3115 1.61 0.109
## Year1998 0.1754 0.2952 0.59 0.553
## Year1999 0.6004 0.3440 1.75 0.082 .
## Year2000 0.7975 0.3730 2.14 0.034 *
## Year2001 0.6849 0.3526 1.94 0.053 .
## Year2002 0.1669 0.2960 0.56 0.573
## Year2003 0.0485 0.3341 0.15 0.885
## Year2004 0.4693 0.3029 1.55 0.123
## Year2005 0.3078 0.3023 1.02 0.310
## Year2006 0.4677 0.3948 1.18 0.237
```

```

## Year2007          0.3701      0.2825      1.31      0.191
## Year2008          0.4806      0.2920      1.65      0.101
## Year2009          0.3426      0.3069      1.12      0.265
## Year2010          0.4191      0.2853      1.47      0.143
## Year2011          0.4602      0.2813      1.64      0.103
## Year2012          0.3920      0.2844      1.38      0.169
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.487
## Multiple R-squared:  0.0603, Adjusted R-squared:  -0.00442
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 239 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.204  0.865   0.956   0.902   0.985   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      3.77e-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 1.123  1      1.060
## Year              1.123 16      1.004

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.3294 -0.3209  0.0332  0.3095  1.6433

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.5560    0.2756   2.02   0.045 *
## LastAuthorFemale1 0.0554    0.0639   0.87   0.386
## Year1997        0.5495    0.3160   1.74   0.083 .
## Year1998        0.2367    0.2970   0.80   0.426
## Year1999        0.6791    0.3492   1.94   0.053 .
## Year2000        0.8751    0.3660   2.39   0.018 *
## Year2001        0.7734    0.3517   2.20   0.029 *
## Year2002        0.2296    0.3007   0.76   0.446
## Year2003        0.1315    0.3283   0.40   0.689
## Year2004        0.5322    0.3016   1.76   0.079 .
## Year2005        0.4087    0.2943   1.39   0.166
## Year2006        0.5182    0.3923   1.32   0.188
## Year2007        0.4436    0.2853   1.55   0.121
## Year2008        0.5675    0.2878   1.97   0.050 *
## Year2009        0.4155    0.3044   1.36   0.174
## Year2010        0.5042    0.2836   1.78   0.077 .
## Year2011        0.5295    0.2781   1.90   0.058 .
## Year2012        0.4664    0.2809   1.66   0.098 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.486
## Multiple R-squared:  0.0582, Adjusted R-squared:  -0.00666
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 241 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.230  0.882  0.953  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      3.77e-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: 265"
## [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
## 24 25 11 11 17 6 12 8 13 13 17 23 19 16 23
## 2011 2012
## 23 18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4 7 1 2 7 1 5 4 5 13 13 16 10 8 17
## 2011 2012
## 15 14
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1 7 1 2 4 1 5 3 4 12 11 12 8 5 12
## 2011 2012
## 13 13
## [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: 5.51784835276224"

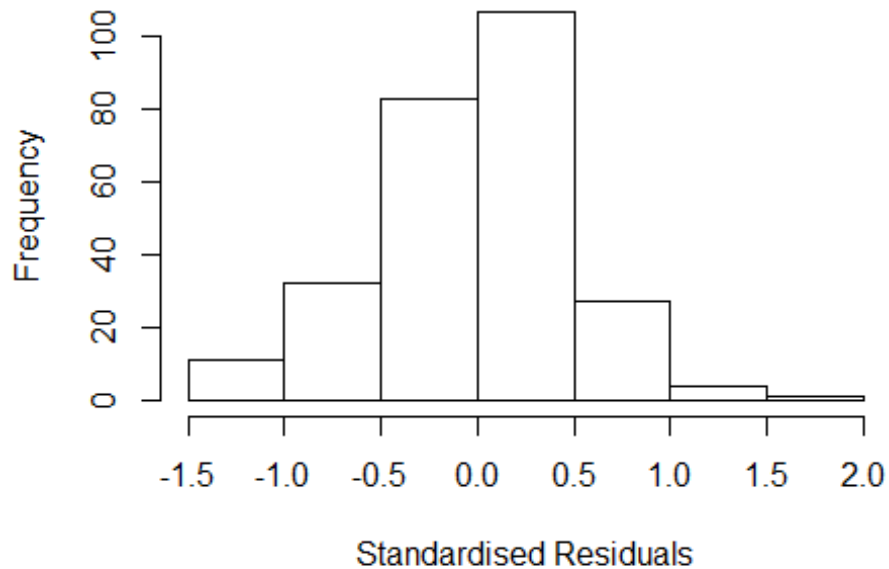
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## 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: 6.18786183886527"

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

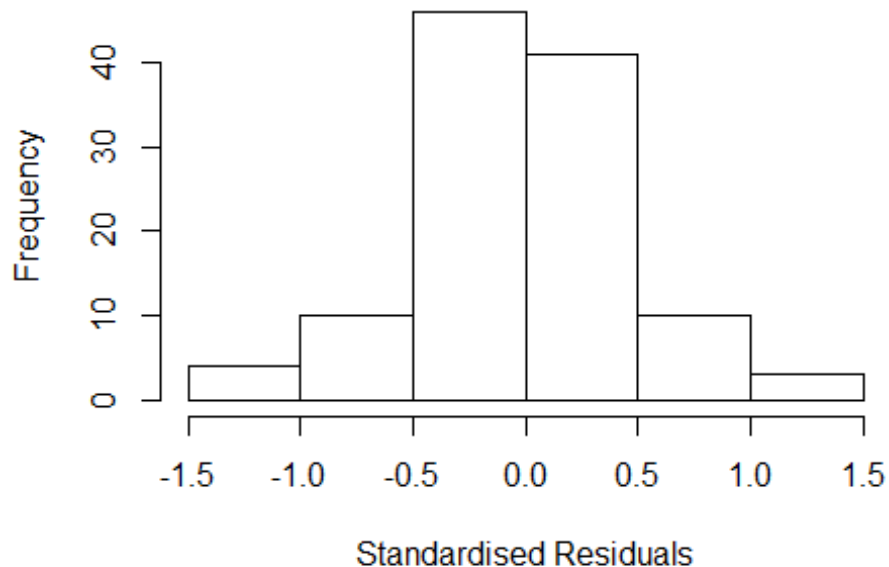
```


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,
Year as factors"
##               GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -2.060e+15  1          NaN
## LastAuthorFemale  1.086e+00  1          1.042
## UniqueAuthors    -2.866e+15  4          NaN
## Year              -2.271e+28 16          NaN
```

Residuals from first and last author and team size



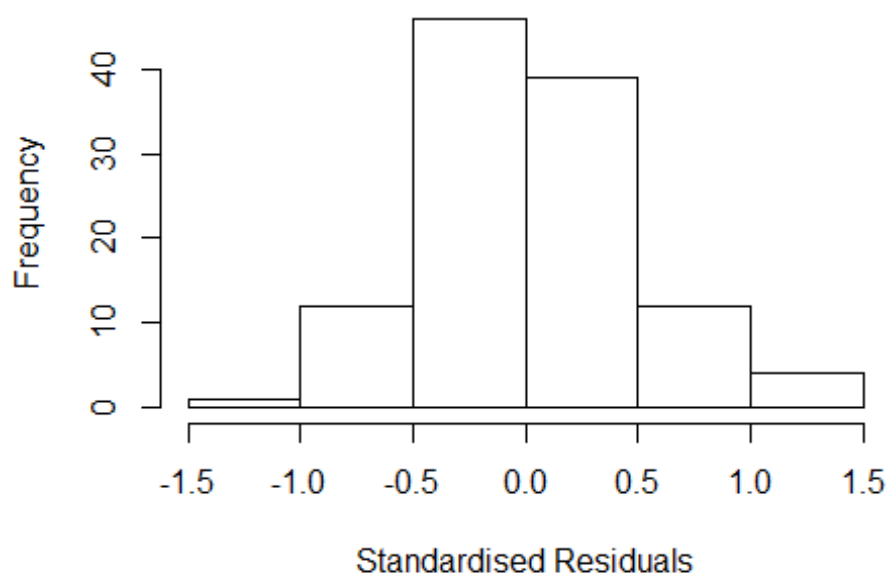
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.24910 -0.31512 -0.00115 0.28457 1.14163
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.25145 0.22061 5.67 1.6e-07 ***
## FirstAuthorFemale1 0.45647 0.12132 3.76 0.0003 ***
## LastAuthorFemale1 0.00210 0.27080 0.01 0.9938
## UniqueAuthors2 -0.46610 0.24514 -1.90 0.0604 .
## UniqueAuthors3 -0.44842 0.20736 -2.16 0.0332 *
## UniqueAuthors4 -0.10326 0.21447 -0.48 0.6313
## UniqueAuthors5 -0.09892 0.18864 -0.52 0.6013
## Year1997 -0.05681 0.54546 -0.10 0.9173
## Year1998 0.63447 0.12132 5.23 1.1e-06 ***
## Year1999 -0.15475 0.33775 -0.46 0.6479
```

```

## Year2000      -0.00227    0.29404   -0.01    0.9939
## Year2001      -0.04518    0.19353   -0.23    0.8159
## Year2002      -0.01994    0.18558   -0.11    0.9147
## Year2003      -0.02019    0.20539   -0.10    0.9219
## Year2004      -0.34188    0.24333   -1.41    0.1634
## Year2005       0.13044    0.20372    0.64    0.5236
## Year2006       0.02063    0.16607    0.12    0.9014
## Year2007      -0.27216    0.17357   -1.57    0.1203
## Year2008      -0.23485    0.15863   -1.48    0.1422
## Year2009      -0.25556    0.16035   -1.59    0.1144
## Year2010      -0.22315    0.16520   -1.35    0.1801
## Year2011      -0.04914    0.16944   -0.29    0.7724
## Year2012       0.10092    0.19512    0.52    0.6063
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.469
## Multiple R-squared:  0.281, Adjusted R-squared:  0.107
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 100 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.458  0.890  0.953  0.904  0.984  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00      5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      8.77e-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 -4.504e+15  1      NaN
## LastAuthorFemale  1.489e+00  1      1.22
## Year              -4.621e+15 16      NaN

```

Residuals from first and last author



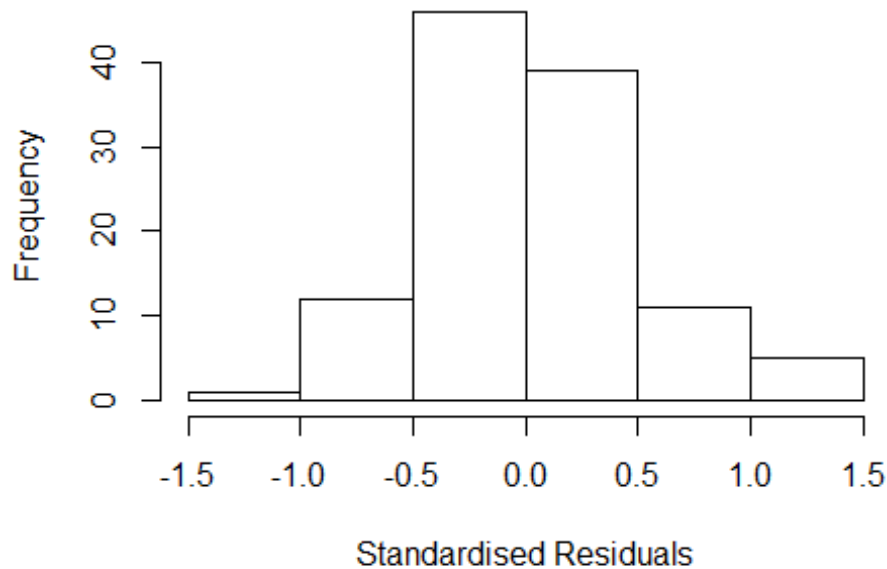
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.22e+00 -2.84e-01 8.53e-16 3.05e-01 1.28e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1478 0.1168 9.83 3.8e-16 ***
## FirstAuthorFemale1 0.4612 0.1168 3.95 0.00015 ***
## LastAuthorFemale1 0.0392 0.2821 0.14 0.88989
## Year1997 -0.1908 0.6036 -0.32 0.75255
## Year1998 0.6392 0.1168 5.47 3.6e-07 ***
## Year1999 -0.3295 0.1674 -1.97 0.05191 .
## Year2000 -0.0796 0.2587 -0.31 0.75888
## Year2001 -0.0448 0.1168 -0.38 0.70189
## Year2002 -0.2128 0.1981 -1.07 0.28531
## Year2003 -0.1444 0.2071 -0.70 0.48720
## Year2004 -0.5164 0.1427 -3.62 0.00048 ***
## Year2005 0.0618 0.1876 0.33 0.74240
```

```

## Year2006          -0.0360      0.1553   -0.23  0.81696
## Year2007          -0.4028      0.1560   -2.58  0.01138 *
## Year2008          -0.2537      0.1729   -1.47  0.14545
## Year2009          -0.2551      0.1288   -1.98  0.05056 .
## Year2010          -0.3787      0.1754   -2.16  0.03334 *
## Year2011          -0.1671      0.1267   -1.32  0.19034
## Year2012           0.0694      0.2102    0.33  0.74188
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.495
## Multiple R-squared:  0.203, Adjusted R-squared:  0.0515
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 100 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.485  0.871  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      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 125682 1      354.516
## Year              125682 16      1.443

```

Residuals from first author



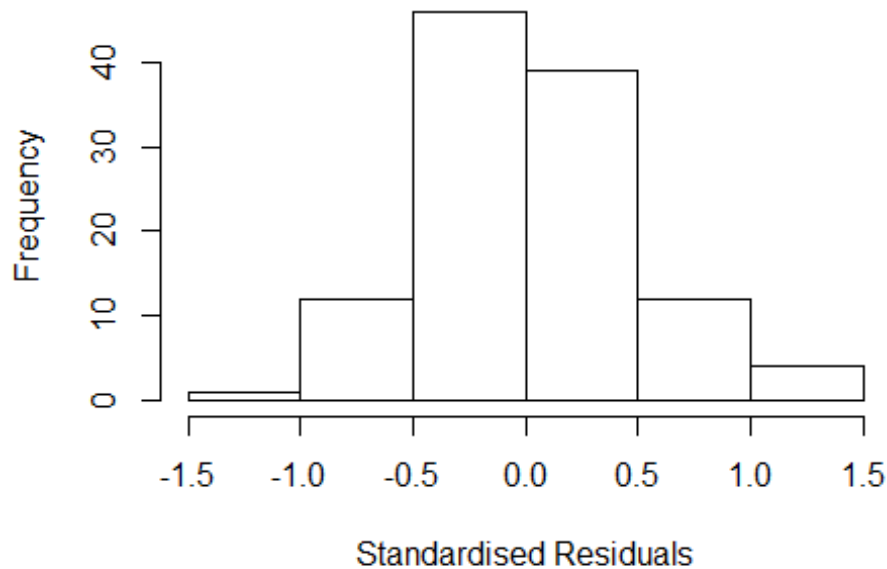
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.22e+00 -2.83e-01 -1.58e-15 3.04e-01 1.27e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1491 0.1154 9.96 < 2e-16 ***
## FirstAuthorFemale1 0.4599 0.1154 3.98 0.00013 ***
## Year1997 -0.1950 0.6051 -0.32 0.74803
## Year1998 0.6379 0.1154 5.53 2.8e-07 ***
## Year1999 -0.3295 0.1672 -1.97 0.05161 .
## Year2000 -0.0810 0.2571 -0.32 0.75340
## Year2001 -0.0461 0.1154 -0.40 0.69039
## Year2002 -0.2151 0.1981 -1.09 0.28018
## Year2003 -0.1450 0.2063 -0.70 0.48387
## Year2004 -0.5173 0.1417 -3.65 0.00043 ***
## Year2005 0.0604 0.1869 0.32 0.74729
## Year2006 -0.0370 0.1542 -0.24 0.81068
```

```

## Year2007          -0.4016      0.1557   -2.58  0.01139 *
## Year2008          -0.2554      0.1722   -1.48  0.14131
## Year2009          -0.2558      0.1283   -1.99  0.04894 *
## Year2010          -0.3744      0.1728   -2.17  0.03274 *
## Year2011          -0.1590      0.1246   -1.28  0.20491
## Year2012           0.0718      0.2076    0.35  0.72999
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.502
## Multiple R-squared:  0.201, Adjusted R-squared:  0.0592
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 100 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.499  0.873  0.956  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      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 87.45 1          9.352
## Year            87.45 16          1.150

```

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
## -1.26e+00 -3.08e-01 3.61e-16 3.01e-01 1.59e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.61e+00 3.03e-08 5.31e+07 < 2e-16 ***
## LastAuthorFemale1 -4.60e-02 4.40e-01 -1.00e-01 0.91703
## Year1997 -7.17e-01 7.41e-01 -9.70e-01 0.33541
## Year1998 1.78e-01 4.47e-08 3.98e+06 < 2e-16 ***
## Year1999 -3.30e-01 1.66e-01 -1.98e+00 0.05064 .
## Year2000 -4.40e-01 1.35e-01 -3.26e+00 0.00154 **
## Year2001 -5.06e-01 5.44e-08 -9.31e+06 < 2e-16 ***
## Year2002 -4.54e-01 2.25e-01 -2.02e+00 0.04611 *
## Year2003 -4.40e-01 1.67e-01 -2.63e+00 0.00993 **
## Year2004 -8.62e-01 6.30e-02 -1.37e+01 < 2e-16 ***
## Year2005 -3.14e-01 2.00e-01 -1.57e+00 0.11876
## Year2006 -4.54e-01 9.03e-02 -5.03e+00 2.3e-06 ***
```

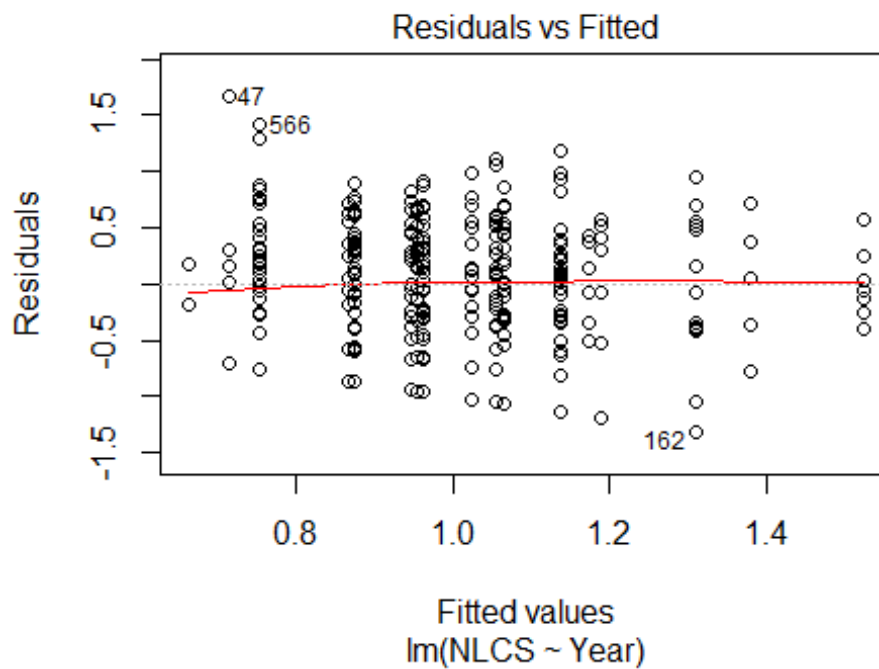


```

## Year2007          -5.49e-01    1.94e-01 -2.84e+00  0.00555 **
## Year2008          -6.50e-01    1.52e-01 -4.29e+00  4.3e-05 ***
## Year2009          -5.31e-01    1.29e-01 -4.10e+00  8.7e-05 ***
## Year2010          -6.73e-01    1.82e-01 -3.70e+00  0.00036 ***
## Year2011          -4.24e-01    1.73e-01 -2.44e+00  0.01640 *
## Year2012          -3.45e-01    2.13e-01 -1.62e+00  0.10850
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.532
## Multiple R-squared:  0.0906, Adjusted R-squared:  -0.0704
## Convergence in 30 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 99 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.349  0.875  0.956  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      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 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
##   14   22   18   29   26   26   36   24   38   39   44   60   74   64   55
## 2011 2012
##   35   52
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    5    7    6    7    7   16   11   23   19   21   29   38   44   34
## 2011 2012

```

```
## 24 37
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2 5 4 6 7 5 13 7 15 18 16 26 31 40 33
## 2011 2012
## 23 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 = 9.7, df = 16, p-value = 0.9
```



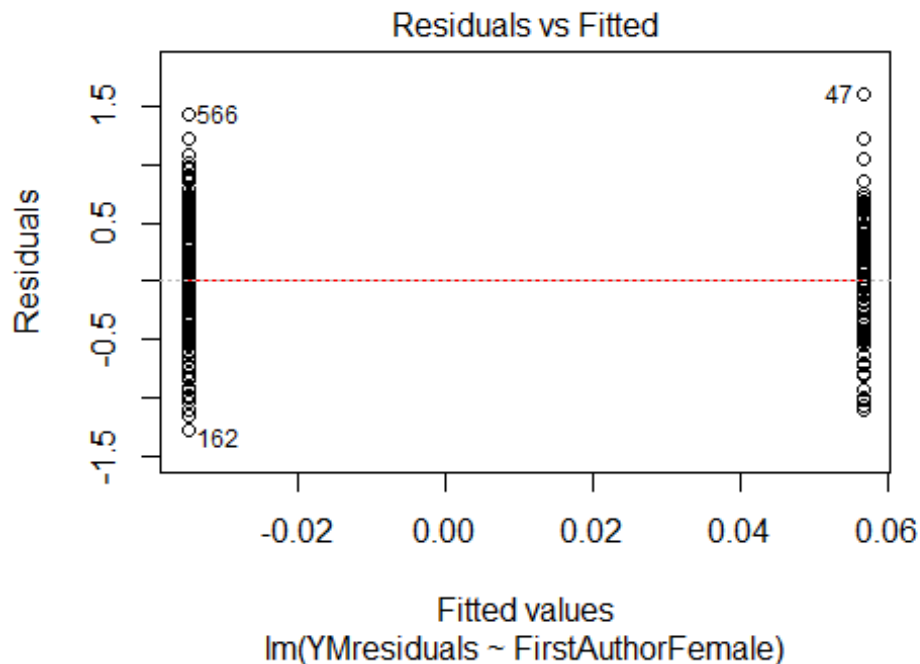
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.16, df = 1, p-value = 0.7

## [1] "Female first author team size 2018 geometric mean: 4.42313868060092"
## [1] "Male first author team size 2018 geometric mean: 4.65557323798701"

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

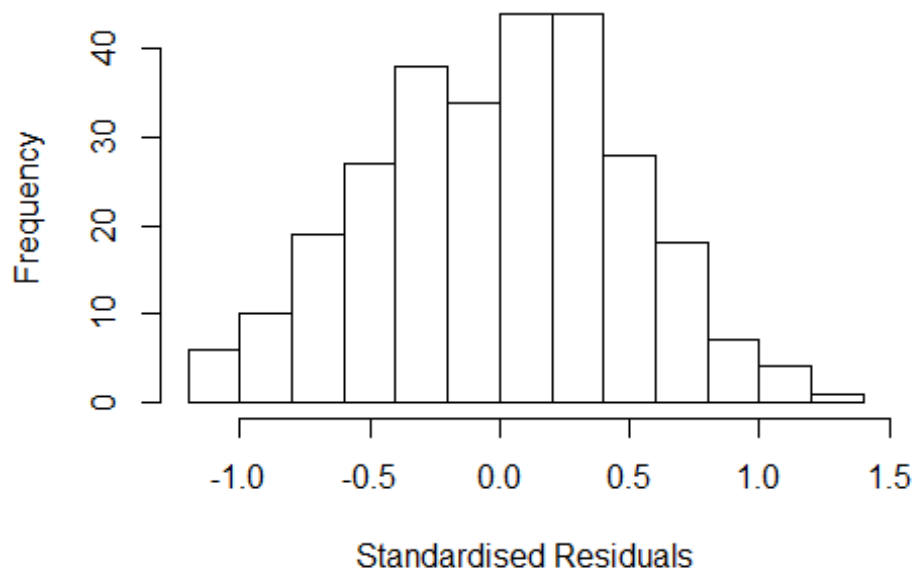
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 72, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.95360930140854"
## [1] "Male last author team size 2018 geometric mean: 5.00983970013357"

## Warning in wilcox.test.default(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 = 0.2
## 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 1.331 1          1.154
## LastAuthorFemale  1.448 1          1.203
## UniqueAuthors    3.010 4          1.148
## Year              3.580 16         1.041
```

Residuals from first and last author and team size



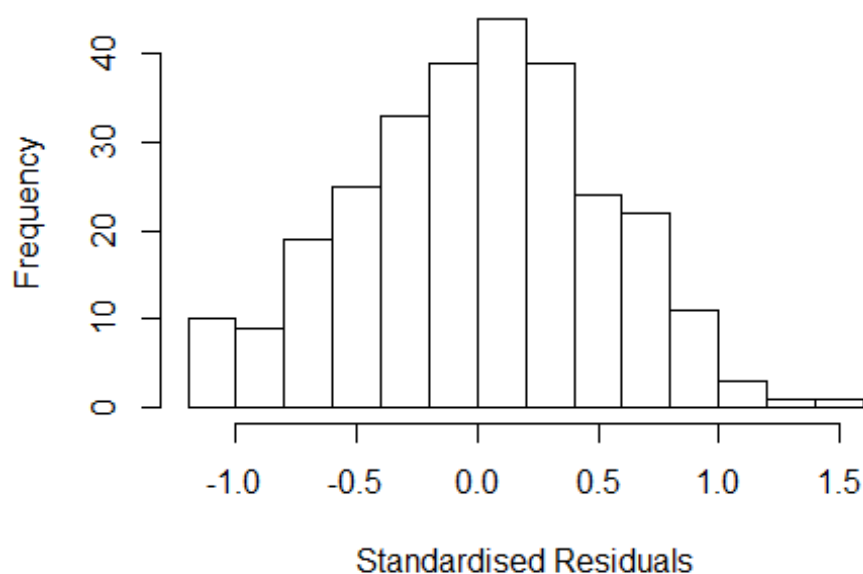
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1992 -0.3535  0.0344  0.3445  1.2901
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.4399    0.2139   2.06  0.0407 *
## FirstAuthorFemale1 0.0710    0.0668   1.06  0.2886
## LastAuthorFemale1 0.0690    0.0893   0.77  0.4403
## UniqueAuthors2    0.2341    0.1357   1.73  0.0856 .
## UniqueAuthors3    0.2588    0.1395   1.86  0.0646 .
## UniqueAuthors4    0.4379    0.1342   3.26  0.0013 **
## UniqueAuthors5    0.5866    0.1201   4.88 1.8e-06 ***
## Year1997          0.5160    0.2555   2.02  0.0445 *
## Year1998         -0.1130    0.3662  -0.31  0.7579
## Year1999          0.3033    0.2152   1.41  0.1599
```

```

## Year2000          0.6322      0.2735      2.31      0.0216 *
## Year2001          0.6036      0.2535      2.38      0.0180 *
## Year2002          0.2966      0.2727      1.09      0.2778
## Year2003          0.1173      0.2947      0.40      0.6911
## Year2004         -0.0488      0.2444     -0.20      0.8419
## Year2005          0.1454      0.2549      0.57      0.5689
## Year2006          0.1618      0.2390      0.68      0.4991
## Year2007          0.2504      0.2323      1.08      0.2820
## Year2008          0.0555      0.2345      0.24      0.8130
## Year2009         -0.1556      0.2197     -0.71      0.4793
## Year2010         -0.0127      0.2285     -0.06      0.9558
## Year2011          0.1713      0.2325      0.74      0.4619
## Year2012          0.1223      0.2252      0.54      0.5877
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.525
## Multiple R-squared:  0.216, Adjusted R-squared:  0.149
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 261 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.525  0.888  0.955  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      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 1.328 1      1.153
## LastAuthorFemale  1.441 1      1.201
## Year              1.842 16      1.019

```

Residuals from first and last author



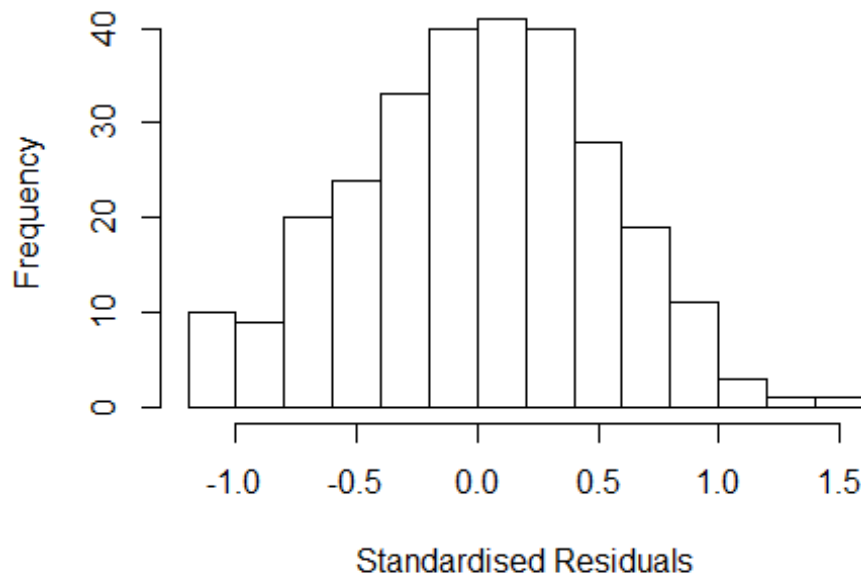
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1824 -0.3637 0.0364 0.3367 1.5006
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5715 0.1624 3.52 0.00051 ***
## FirstAuthorFemale1 0.1079 0.0688 1.57 0.11802
## LastAuthorFemale1 -0.0339 0.0950 -0.36 0.72154
## Year1997 0.7792 0.2871 2.71 0.00709 **
## Year1998 0.0482 0.2313 0.21 0.83517
## Year1999 0.5581 0.2109 2.65 0.00863 **
## Year2000 0.9367 0.2061 4.54 8.4e-06 ***
## Year2001 0.8581 0.2244 3.82 0.00016 ***
## Year2002 0.5876 0.2443 2.40 0.01687 *
## Year2003 0.4095 0.2735 1.50 0.13560
## Year2004 0.2972 0.1976 1.50 0.13380
## Year2005 0.4682 0.2145 2.18 0.02995 *
```

```

## Year2006          0.4566      0.1948      2.34  0.01985 *
## Year2007          0.5030      0.1838      2.74  0.00663 **
## Year2008          0.3177      0.1845      1.72  0.08631 .
## Year2009          0.0888      0.1804      0.49  0.62290
## Year2010          0.3555      0.1793      1.98  0.04848 *
## Year2011          0.4566      0.1784      2.56  0.01105 *
## Year2012          0.4434      0.1709      2.59  0.01000 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.543
## Multiple R-squared:  0.118, Adjusted R-squared:  0.0569
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 265 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.425  0.882  0.955  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      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 1.362 1      1.167
## Year              1.362 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
## -1.1732 -0.3535  0.0413  0.3386  1.5097
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.5598    0.1452   3.86 0.00015 ***
## FirstAuthorFemale1 0.1027    0.0699   1.47 0.14307
## Year1997        0.7877    0.2745   2.87 0.00444 **
## Year1998        0.0303    0.2203   0.14 0.89084
## Year1999        0.5608    0.2054   2.73 0.00676 **
## Year2000        0.9454    0.1909   4.95 1.3e-06 ***
## Year2001        0.8645    0.2168   3.99 8.7e-05 ***
## Year2002        0.5996    0.2323   2.58 0.01040 *
## Year2003        0.4188    0.2606   1.61 0.10928
## Year2004        0.2994    0.1887   1.59 0.11388
## Year2005        0.4739    0.2036   2.33 0.02068 *
## Year2006        0.4651    0.1837   2.53 0.01191 *
```



```

## Year2007          0.5107      0.1724      2.96  0.00332 **
## Year2008          0.3271      0.1704      1.92  0.05600 .
## Year2009          0.0915      0.1703      0.54  0.59141
## Year2010          0.3609      0.1679      2.15  0.03254 *
## Year2011          0.4613      0.1680      2.75  0.00645 **
## Year2012          0.4558      0.1544      2.95  0.00345 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.542
## Multiple R-squared:  0.117, Adjusted R-squared:  0.06
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 265 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.418  0.882  0.955  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      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 1.422  1      1.193
## Year              1.422 16      1.011

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.1914 -0.3277  0.0328  0.3464  1.4621

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.66681    0.13899   4.80 2.7e-06 ***
## LastAuthorFemale1 -0.00863    0.09588  -0.09 0.92838
## Year1997        0.72146    0.30103   2.40 0.01725 *
## Year1998        0.01523    0.24662   0.06 0.95081
## Year1999        0.51200    0.19897   2.57 0.01062 *
## Year2000        0.85269    0.18217   4.68 4.6e-06 ***
## Year2001        0.80553    0.22064   3.65 0.00032 ***
## Year2002        0.52459    0.23644   2.22 0.02736 *
## Year2003        0.34751    0.25033   1.39 0.16625
## Year2004        0.23964    0.18736   1.28 0.20200
## Year2005        0.39300    0.19727   1.99 0.04738 *
## Year2006        0.37276    0.17790   2.10 0.03711 *
## Year2007        0.44258    0.17247   2.57 0.01084 *
## Year2008        0.25599    0.17425   1.47 0.14301
## Year2009        0.03205    0.16702   0.19 0.84799
## Year2010        0.29375    0.16548   1.78 0.07705 .
## Year2011        0.40598    0.17177   2.36 0.01884 *
## Year2012        0.38519    0.16225   2.37 0.01831 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.547
## Multiple R-squared:  0.11, Adjusted R-squared:  0.0525
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 259 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.456 0.868 0.957 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      3.57e-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 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
##    3    5    3    7    7    8    7   12    9    9   10   14   15   32   27
## 2011 2012
##   30   26
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    2    3    2    5    5    7    8    6    5    9   12   13   20
## 2011 2012
##   22   15
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    2    2    2    5    5    7    6    6    5    6   10   12   15
## 2011 2012
##   20   14
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.05932717962681"
## [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

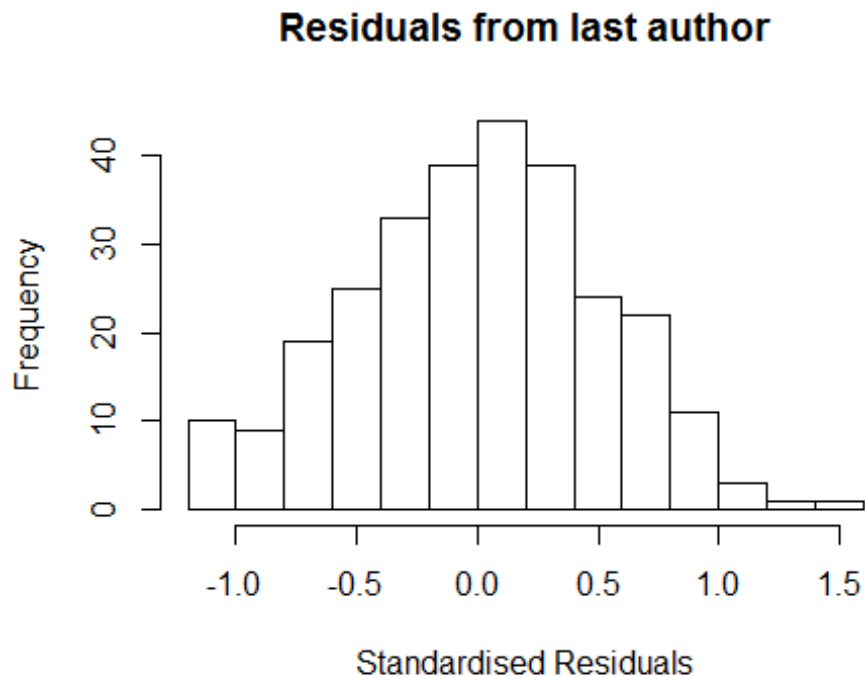
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 21, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.81273717198775"
## [1] "Male last author 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 = 18, 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
```



```
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.067e+01  1          3.267
## LastAuthorFemale  9.994e+00  1          3.161
## UniqueAuthors    6.683e+15  4         95.086
## Year             2.442e+16 16          3.252

## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9338 -0.2544  0.0426  0.2421  1.1263
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.5960      0.0000      Inf < 2e-16 ***
```

```

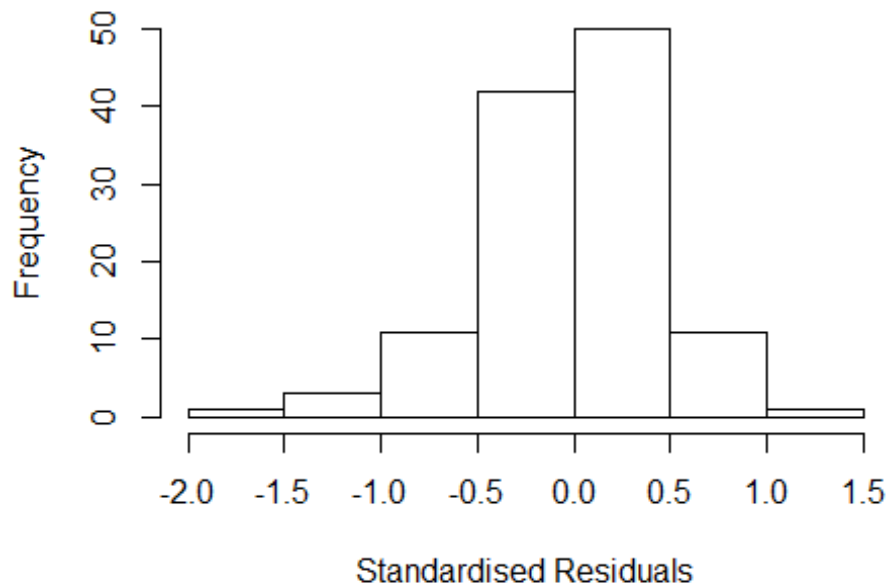
## FirstAuthorFemale1 -0.2431      0.1045    -2.33  0.02209 *
## LastAuthorFemale1 -0.0217      0.1255    -0.17  0.86311
## UniqueAuthors2     0.6872      0.2315     2.97  0.00378 **
## UniqueAuthors3     0.8170      0.2396     3.41  0.00095 ***
## UniqueAuthors4     0.9964      0.2817     3.54  0.00063 ***
## UniqueAuthors5     0.8932      0.2547     3.51  0.00069 ***
## Year1997           -1.7454      0.2817    -6.20  1.4e-08 ***
## Year1998           -0.8968      0.2423    -3.70  0.00036 ***
## Year1999           -1.3902      0.2199    -6.32  8.1e-09 ***
## Year2000           -0.4986      0.2809    -1.78  0.07903 .
## Year2001           -1.2518      0.2999    -4.17  6.6e-05 ***
## Year2002           -1.0462      0.1497    -6.99  3.7e-10 ***
## Year2003           -1.0173      0.2668    -3.81  0.00024 ***
## Year2004           -0.0409      0.3991    -0.10  0.91856
## Year2005           -0.5555      0.6240    -0.89  0.37557
## Year2006           -0.1202      0.3226    -0.37  0.71035
## Year2007           -0.6632      0.3352    -1.98  0.05075 .
## Year2008           -0.5647      0.1702    -3.32  0.00128 **
## Year2009           -0.9666      0.2501    -3.86  0.00020 ***
## Year2010           -0.8990      0.1940    -4.63  1.1e-05 ***
## Year2011           -1.1611      0.2096    -5.54  2.6e-07 ***
## Year2012           -1.2098      0.2242    -5.40  4.9e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.411
## Multiple R-squared:  0.496, Adjusted R-squared:  0.38
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## observation 37 is an outlier with |weight| = 0 ( < 0.00084);
## 10 weights are ~= 1. The remaining 108 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.434  0.881  0.960  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           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"

## 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  5.801  1          2.408
## LastAuthorFemale  3.742  1          1.934
## Year              11.482 16          1.079

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7976 -0.3375  0.0181  0.3037  1.2598
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.60e+00  0.00e+00      Inf < 2e-16 ***
```

```

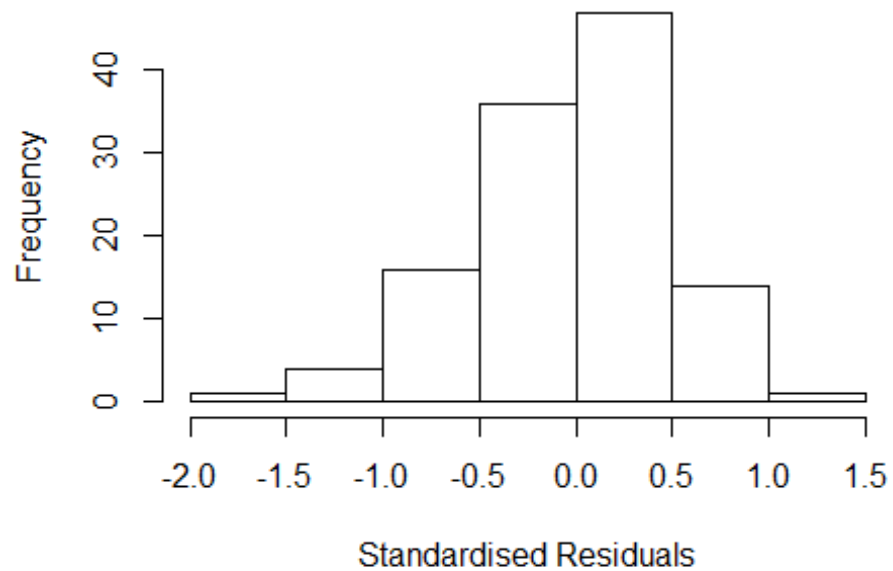
## FirstAuthorFemale1 -2.00e-01 1.18e-01 -1.70e+00 0.0922 .
## LastAuthorFemale1 -5.60e-02 1.11e-01 -5.10e-01 0.6138
## Year1997 -7.49e-01 3.47e-09 -2.16e+08 < 2e-16 ***
## Year1998 -1.32e-01 2.13e-01 -6.20e-01 0.5372
## Year1999 -6.64e-01 1.58e-01 -4.21e+00 5.6e-05 ***
## Year2000 1.67e-01 1.43e-01 1.17e+00 0.2464
## Year2001 -5.27e-01 1.90e-01 -2.77e+00 0.0067 **
## Year2002 -5.47e-01 2.58e-01 -2.12e+00 0.0362 *
## Year2003 -2.11e-01 2.18e-01 -9.70e-01 0.3343
## Year2004 7.01e-01 2.97e-01 2.36e+00 0.0202 *
## Year2005 2.02e-01 3.43e-01 5.90e-01 0.5584
## Year2006 1.71e-01 4.23e-01 4.00e-01 0.6866
## Year2007 1.13e-01 1.93e-01 5.90e-01 0.5578
## Year2008 9.99e-03 2.20e-01 5.00e-02 0.9638
## Year2009 -1.69e-01 1.73e-01 -9.80e-01 0.3288
## Year2010 -1.45e-01 1.23e-01 -1.18e+00 0.2418
## Year2011 -4.22e-01 1.91e-01 -2.20e+00 0.0297 *
## Year2012 -3.97e-01 1.30e-01 -3.06e+00 0.0029 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.486
## Multiple R-squared: 0.316, Adjusted R-squared: 0.193
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~ 1. The remaining 110 ones are summarized as
##   Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.141 0.874 0.953 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          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"

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

```

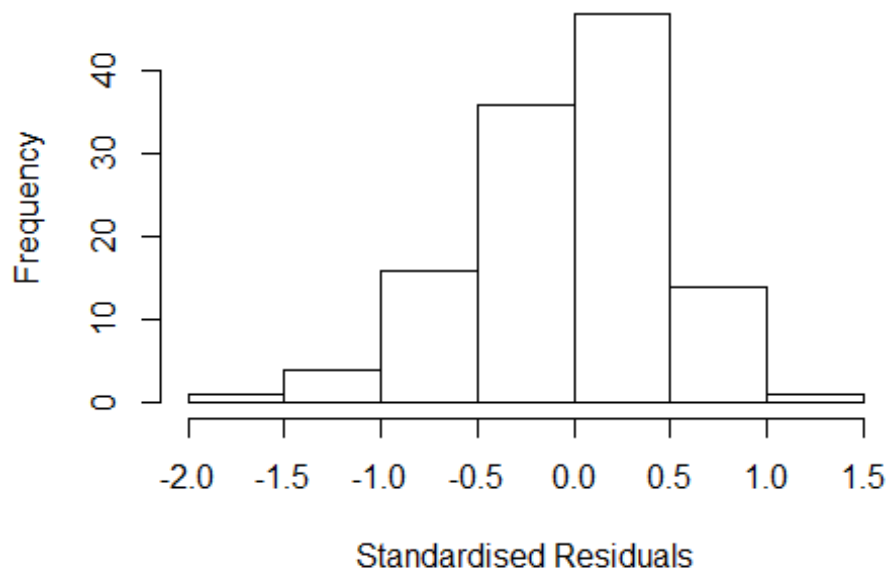
```
## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

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
## -1.754 -0.338 0.035 0.322 1.251
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.59600 0.00000 Inf < 2e-16 ***
## FirstAuthorFemale1 -0.21762 0.11337 -1.92 0.0577 .
## Year1997 -0.74900 0.00000 -Inf < 2e-16 ***
## Year1998 -0.14288 0.22949 -0.62 0.5349
## Year1999 -0.67438 0.14123 -4.77 6.1e-06 ***
## Year2000 0.17581 0.14805 1.19 0.2378
## Year2001 -0.52642 0.19100 -2.76 0.0069 **
## Year2002 -0.57434 0.24485 -2.35 0.0209 *
## Year2003 -0.23595 0.18956 -1.24 0.2161
## Year2004 0.67545 0.27836 2.43 0.0170 *
## Year2005 0.15813 0.34667 0.46 0.6493
## Year2006 0.18039 0.40677 0.44 0.6584
```

```

## Year2007          0.10475    0.19265    0.54    0.5878
## Year2008          -0.00715    0.21380   -0.03    0.9734
## Year2009          -0.21004    0.15594   -1.35    0.1810
## Year2010          -0.17011    0.11285   -1.51    0.1348
## Year2011          -0.45220    0.17922   -2.52    0.0132 *
## Year2012          -0.41921    0.12391   -3.38    0.0010 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.524
## Multiple R-squared:  0.306, Adjusted R-squared:  0.19
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 111 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.239  0.887  0.962  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      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 3.389 1      1.841
## Year              3.389 16      1.039

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.6256 -0.2913  0.0388  0.3313  1.4000

```

```

##
## Coefficients:
##              Estimate Std. Error   t value Pr(>|t|)
## (Intercept)   1.60e+00   2.75e-08   5.80e+07 < 2e-16 ***
## LastAuthorFemale1 -1.19e-01   1.08e-01  -1.09e+00   0.2766
## Year1997       -7.49e-01   4.66e-08  -1.61e+07 < 2e-16 ***
## Year1998       -3.01e-01   1.62e-01  -1.85e+00   0.0666 .
## Year1999       -8.33e-01   1.39e-01  -5.97e+00   3.5e-08 ***
## Year2000        6.70e-02   5.68e-02   1.18e+00   0.2413
## Year2001       -6.83e-01   1.16e-01  -5.88e+00   5.4e-08 ***
## Year2002       -6.25e-01   2.46e-01  -2.54e+00   0.0127 *
## Year2003       -2.50e-01   2.74e-01  -9.10e-01   0.3642
## Year2004        6.09e-01   2.98e-01   2.05e+00   0.0434 *
## Year2005        2.96e-02   3.36e-01   9.00e-02   0.9301
## Year2006        3.10e-02   3.81e-01   8.00e-02   0.9355
## Year2007        1.11e-02   2.07e-01   5.00e-02   0.9573
## Year2008       -8.80e-02   2.08e-01  -4.20e-01   0.6735
## Year2009       -2.44e-01   1.62e-01  -1.50e+00   0.1359
## Year2010       -2.22e-01   1.16e-01  -1.91e+00   0.0590 .
## Year2011       -5.57e-01   1.71e-01  -3.26e+00   0.0015 **
## Year2012       -4.68e-01   1.39e-01  -3.36e+00   0.0011 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.488
## Multiple R-squared:  0.289, Adjusted R-squared:  0.169
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 108 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.244  0.882  0.952  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      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 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
## 11 1 6 1 7 4 8 8 5 5 3 5 10 10 14
## 2011 2012
## 15 18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3 1 2 0 5 1 1 5 4 3 2 2 4 6 6
## 2011 2012
## 10 14
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3 1 2 0 5 1 1 3 3 3 2 1 1 5 5
## 2011 2012
## 5 12
## [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: 6"

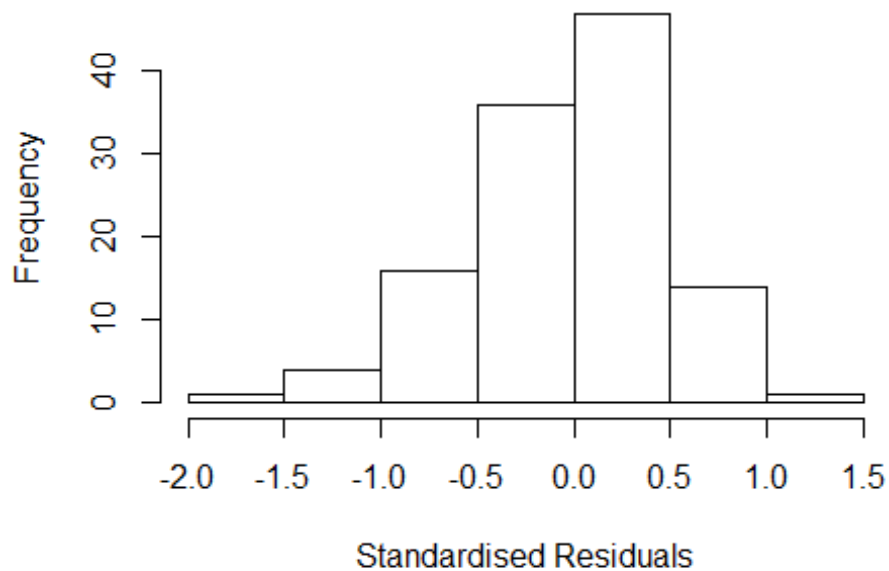
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0, 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: 4.24264068711928"

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

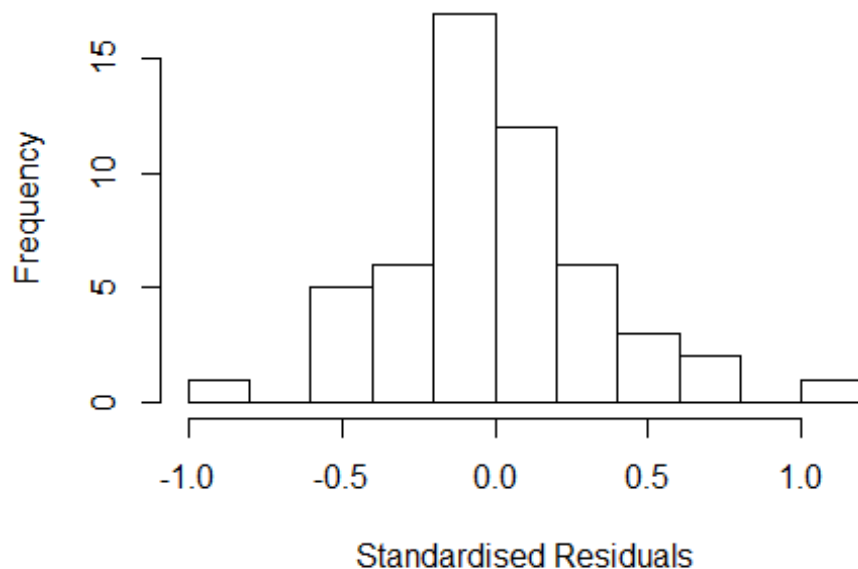
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0.5, p-value = 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 4.083e+14 1      2.021e+07
## LastAuthorFemale -8.462e+00 1           NaN
## UniqueAuthors    4.170e+16 4      1.195e+02
## Year              6.539e+30 15      1.065e+01
```

Residuals from first and last author and team size



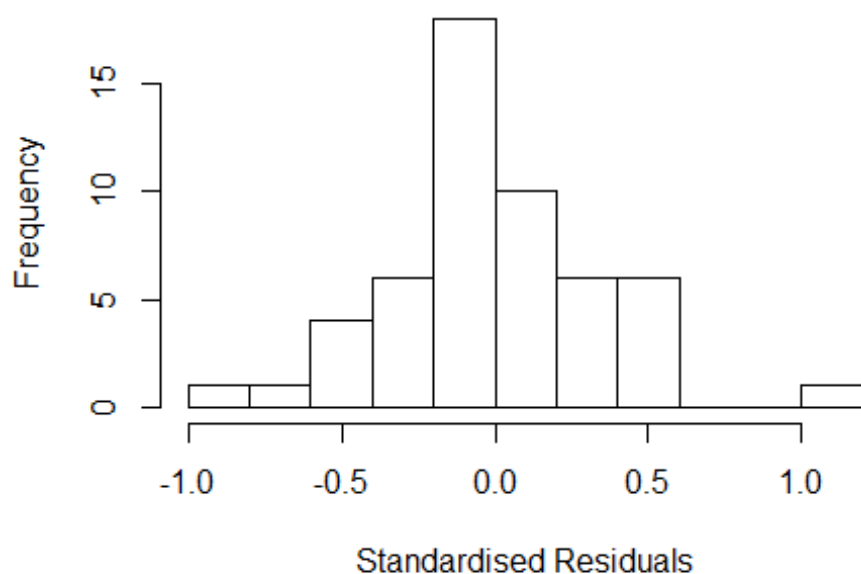
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -9.06e-01 -1.75e-01 3.33e-16 1.87e-01 1.02e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.3554 0.2946 1.21 0.2368
## FirstAuthorFemale1 0.2459 0.1896 1.30 0.2042
## LastAuthorFemale1 0.1158 0.2483 0.47 0.6442
## UniqueAuthors2 0.3302 0.2425 1.36 0.1831
## UniqueAuthors3 0.3412 0.2320 1.47 0.1514
## UniqueAuthors4 0.4223 0.1930 2.19 0.0363 *
## UniqueAuthors5 0.3032 0.2815 1.08 0.2897
## Year1997 0.4825 0.2208 2.19 0.0365 *
## Year1998 -0.0558 0.2887 -0.19 0.8480
## Year2000 0.6440 0.2293 2.81 0.0085 **
```

```

## Year2001      0.5284      0.2210      2.39      0.0231 *
## Year2002      0.2564      0.2715      0.94      0.3522
## Year2003      0.5215      0.2567      2.03      0.0509 .
## Year2004      0.1536      0.2184      0.70      0.4872
## Year2005      0.0766      0.2993      0.26      0.7998
## Year2006      0.3527      0.3174      1.11      0.2750
## Year2007     -0.2146      0.2715     -0.79      0.4353
## Year2008      0.1485      0.2208      0.67      0.5062
## Year2009      0.1211      0.3448      0.35      0.7278
## Year2010      0.1268      0.2598      0.49      0.6289
## Year2011     -0.0778      0.2641     -0.29      0.7703
## Year2012     -0.1409      0.2855     -0.49      0.6252
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.366
## Multiple R-squared:  0.479, Adjusted R-squared:  0.126
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 44 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.420  0.884  0.966  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      1.89e-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 1.713e+13 1 4.139e+06
## LastAuthorFemale 1.086e+01 1 3.295e+00
## Year 2.827e+14 15 3.032e+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
## -8.73e-01 -1.63e-01 -4.44e-16 1.98e-01 1.02e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6210 0.2957 2.10 0.043 *
## FirstAuthorFemale1 0.2449 0.1564 1.57 0.126
## LastAuthorFemale1 0.0488 0.1471 0.33 0.742
## Year1997 0.5211 0.3691 1.41 0.167
## Year1998 -0.1353 0.4029 -0.34 0.739
## Year2000 0.7371 0.3821 1.93 0.062 .
## Year2001 0.6850 0.2957 2.32 0.027 *
## Year2002 0.4141 0.3691 1.12 0.270
## Year2003 0.6248 0.3405 1.84 0.075 .
## Year2004 0.2598 0.3283 0.79 0.434
## Year2005 0.1374 0.3901 0.35 0.727
## Year2006 0.2732 0.4095 0.67 0.509
```

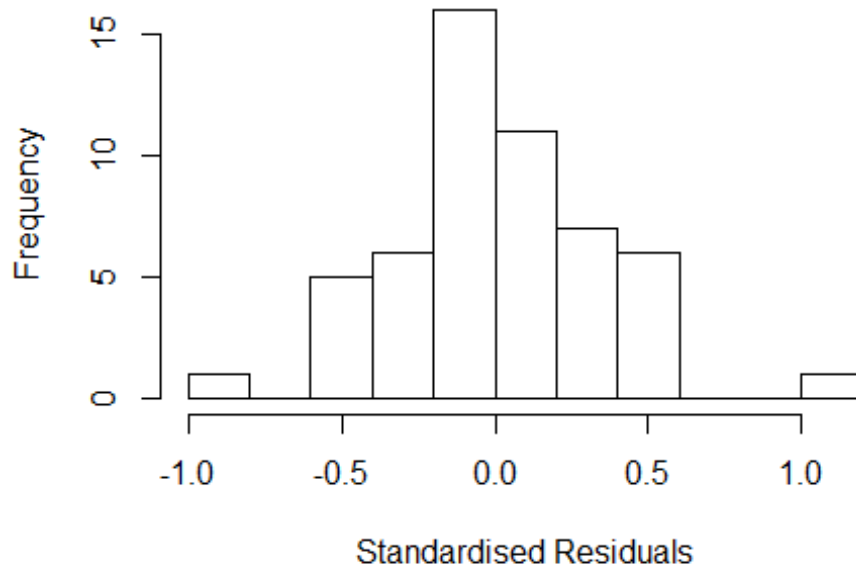


```

## Year2007          -0.0569      0.3691   -0.15    0.878
## Year2008           0.1871      0.3691    0.51    0.615
## Year2009           0.2800      0.4620    0.61    0.548
## Year2010           0.1732      0.3304    0.52    0.604
## Year2011          -0.0617      0.3779   -0.16    0.871
## Year2012          -0.0421      0.3981   -0.11    0.916
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.364
## Multiple R-squared:  0.449, Adjusted R-squared:  0.181
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.408  0.868  0.955  0.904  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      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 2.097e+16 1      1.448e+08
## Year              2.097e+16 15      3.500e+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
## -8.45e-01 -1.58e-01 4.44e-16 2.04e-01 1.05e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6266 0.2732 2.29 0.028 *
## FirstAuthorFemale1 0.2472 0.1634 1.51 0.139
## Year1997 0.5133 0.3359 1.53 0.135
## Year1998 -0.1187 0.3382 -0.35 0.728
## Year2000 0.7495 0.3269 2.29 0.028 *
## Year2001 0.6794 0.2732 2.49 0.018 *
## Year2002 0.4063 0.3359 1.21 0.234
## Year2003 0.6178 0.3093 2.00 0.053 .
## Year2004 0.2702 0.2907 0.93 0.359
## Year2005 0.1296 0.3586 0.36 0.720
## Year2006 0.2898 0.3491 0.83 0.412
## Year2007 -0.0647 0.3359 -0.19 0.848
```

```

## Year2008          0.1793      0.3359      0.53      0.597
## Year2009          0.3148      0.3759      0.84      0.408
## Year2010          0.1672      0.3119      0.54      0.595
## Year2011         -0.0556      0.3427     -0.16      0.872
## Year2012         -0.0289      0.3366     -0.09      0.932
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.369
## Multiple R-squared:  0.446, Adjusted R-squared:  0.199
## 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.
##  0.396  0.882  0.960  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      1.89e-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 22.26  1          4.718
## Year              22.26 15          1.109

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q    Median      3Q      Max
## -7.83e-01 -1.81e-01  2.78e-16  1.58e-01  1.11e+00
##

```

```

## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.6678    0.2027    3.29  0.00222 **
## LastAuthorFemale1 0.0595    0.1492    0.40  0.69234
## Year1997        0.7192    0.2027    3.55  0.00110 **
## Year1998        0.0574    0.2167    0.26  0.79262
## Year2000        0.8354    0.2181    3.83  0.00049 ***
## Year2001        0.6382    0.2027    3.15  0.00330 **
## Year2002        0.6122    0.2027    3.02  0.00463 **
## Year2003        0.7409    0.2153    3.44  0.00148 **
## Year2004        0.2096    0.2363    0.89  0.38097
## Year2005        0.3355    0.2386    1.41  0.16825
## Year2006        0.4659    0.2279    2.04  0.04830 *
## Year2007        0.1412    0.2027    0.70  0.49073
## Year2008        0.3852    0.2027    1.90  0.06549 .
## Year2009        0.3920    0.4172    0.94  0.35370
## Year2010        0.1685    0.2303    0.73  0.46897
## Year2011        0.0128    0.3164    0.04  0.96789
## Year2012        0.0561    0.2773    0.20  0.84077
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.364
## Multiple R-squared:  0.42, Adjusted R-squared:  0.162
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 40 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.329  0.866  0.957  0.899  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      1.89e-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] "#####"

```

```

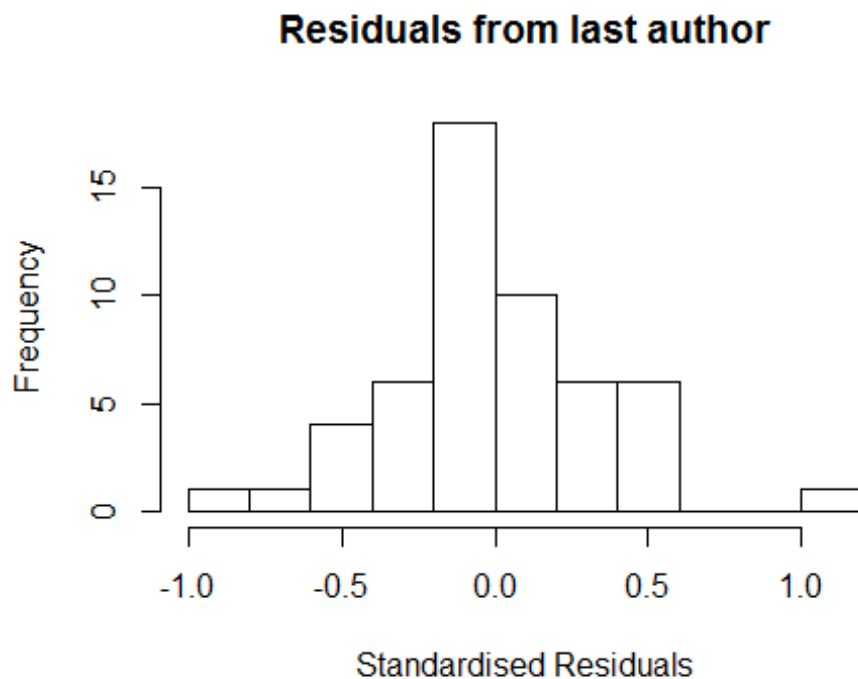
## [1] "Analysis of AJSC 2744"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997
##    2    1
##
## 1996 1997
##    2    1
##
## 1996 1997
##    2    1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: NaN"
## [1] "Male first author team size 2018 geometric mean: 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 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
##    3    2    6    5    2    8    4    3    3    4    8    11    13    16    19
## 2011 2012
##   19   14
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    0    2    3    1    2    2    1    3    3    3    9    10    11    14
## 2011 2012
##   16   13
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    0    2    2    1    2    2    1    1    3    2    7    7    8    13
## 2011 2012
##   13   11
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 5.53645891303657"
## [1] "Male first author team size 2018 geometric mean: 4.57885697021333"

```

```
## Warning in wilcox.test.default(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 = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.82590126048788"
## [1] "Male last author team size 2018 geometric mean: 4.74288121955862"

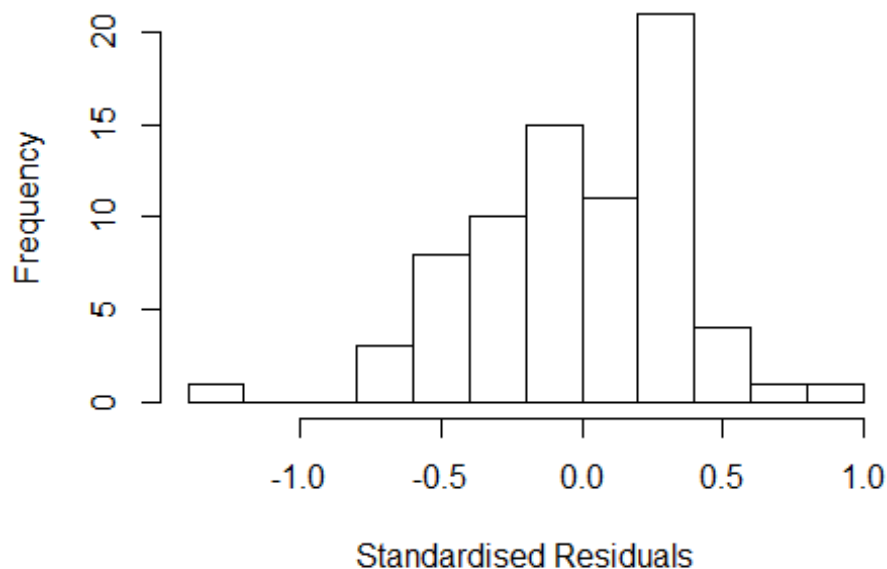
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 11, 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 1.057e+03  1      3.251e+01
```

| | | | |
|---------------------|-----------|----|-----------|
| ## LastAuthorFemale | 2.443e+13 | 1 | 4.943e+06 |
| ## UniqueAuthors | 3.605e+15 | 4 | 8.803e+01 |
| ## Year | 6.189e+18 | 14 | 4.690e+00 |

Residuals from first and last author and team size



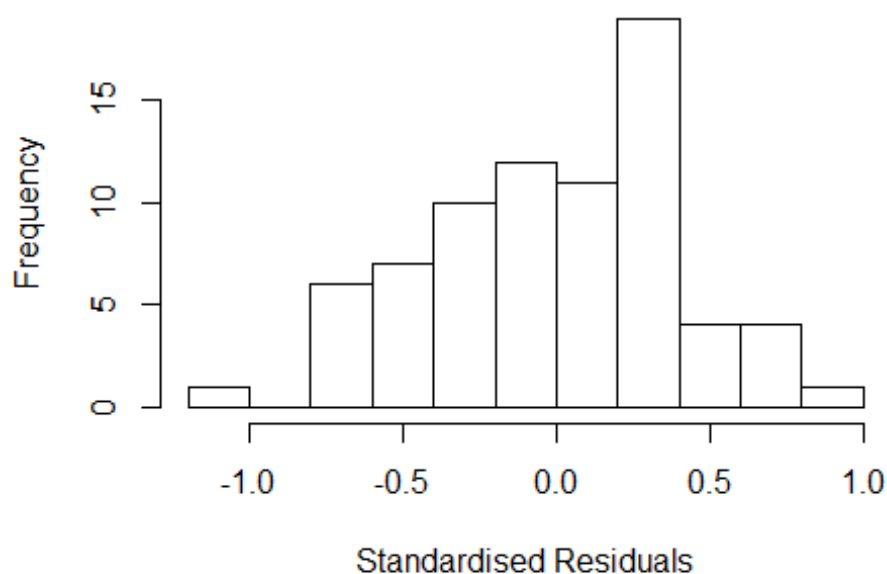
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.21438 -0.23557 0.00537 0.25915 0.82420
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0757 0.2007 5.36 1.8e-06 ***
## FirstAuthorFemale1 0.2668 0.1210 2.20 0.03175 *
## LastAuthorFemale1 -0.0175 0.1118 -0.16 0.87616
## UniqueAuthors2 0.1083 0.1808 0.60 0.55185
## UniqueAuthors3 0.1618 0.2379 0.68 0.49942
## UniqueAuthors4 -0.0253 0.2079 -0.12 0.90361
```

```

## UniqueAuthors5      0.4160      0.2226      1.87  0.06706 .
## Year1999             0.2490      0.1690      1.47  0.14645
## Year2000            -0.2624      0.1479     -1.77  0.08159 .
## Year2001             0.2878      0.2565      1.12  0.26678
## Year2002            -0.0603      0.1523     -0.40  0.69382
## Year2003            -0.2114      0.1479     -1.43  0.15854
## Year2004            -0.2317      0.2081     -1.11  0.27054
## Year2005            -0.5050      0.2017     -2.50  0.01534 *
## Year2006            -0.1147      0.1399     -0.82  0.41556
## Year2007            -0.3079      0.2054     -1.50  0.13977
## Year2008            -0.5898      0.1559     -3.78  0.00039 ***
## Year2009            -0.3066      0.1859     -1.65  0.10486
## Year2010            -0.5266      0.2272     -2.32  0.02426 *
## Year2011            -0.6811      0.1559     -4.37  5.7e-05 ***
## Year2012            -0.4791      0.1392     -3.44  0.00113 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.429
## Multiple R-squared:  0.423, Adjusted R-squared:  0.209
## 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.
##  0.403  0.911  0.962  0.928  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      1.33e-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 1.121e+03  1      3.348e+01
## LastAuthorFemale  1.212e+15  1      3.481e+07
## Year              9.903e+17 14      4.392e+00

```


Residuals from first and last author



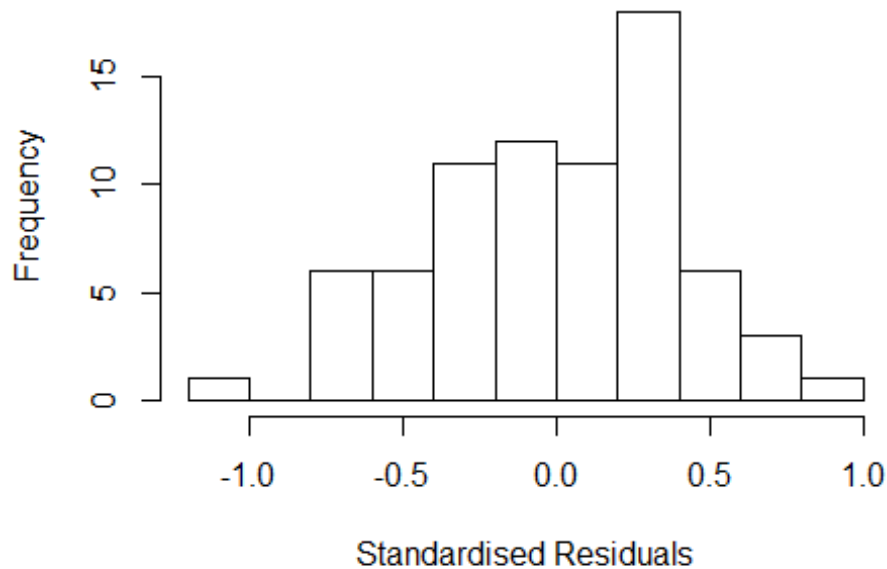
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0486 -0.3208 0.0159 0.2759 0.8976
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2710 0.1666 7.63 2.6e-10 ***
## FirstAuthorFemale1 0.3695 0.1164 3.17 0.0024 **
## LastAuthorFemale1 -0.0442 0.1077 -0.41 0.6826
## Year1999 0.2490 0.1666 1.49 0.1405
## Year2000 -0.1445 0.2032 -0.71 0.4800
## Year2001 0.0419 0.2934 0.14 0.8870
## Year2002 -0.2015 0.1813 -1.11 0.2710
## Year2003 -0.0935 0.2032 -0.46 0.6472
## Year2004 -0.5282 0.2447 -2.16 0.0350 *
## Year2005 -0.6968 0.2122 -3.28 0.0017 **
## Year2006 -0.1014 0.3786 -0.27 0.7898
## Year2007 -0.3556 0.2512 -1.42 0.1623
```

```

## Year2008          -0.5981      0.2088   -2.87   0.0058 **
## Year2009          -0.2675      0.2148   -1.25   0.2181
## Year2010          -0.5476      0.2698   -2.03   0.0469 *
## Year2011          -0.6049      0.2072   -2.92   0.0050 **
## Year2012          -0.5173      0.2253   -2.30   0.0253 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.481
## Multiple R-squared:  0.318, Adjusted R-squared:  0.13
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 65 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.614  0.932  0.959  0.933  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      1.33e-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 1246 1          35.30
## Year              1246 14          1.29

```

Residuals from first author



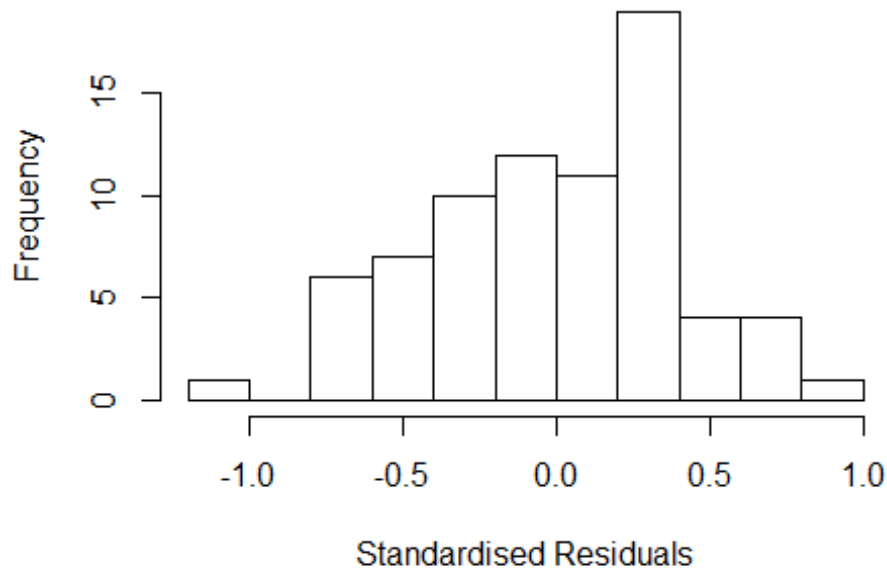
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0762 -0.3206 0.0315 0.2838 0.9127
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2710 0.1666 7.63 2.3e-10 ***
## FirstAuthorFemale1 0.3679 0.1217 3.02 0.0037 **
## Year1999 0.2490 0.1667 1.49 0.1405
## Year2000 -0.1429 0.2063 -0.69 0.4913
## Year2001 0.0206 0.3018 0.07 0.9459
## Year2002 -0.2015 0.1814 -1.11 0.2711
## Year2003 -0.0919 0.2063 -0.45 0.6577
## Year2004 -0.5709 0.2063 -2.77 0.0075 **
## Year2005 -0.7115 0.2078 -3.42 0.0011 **
## Year2006 -0.1235 0.3540 -0.35 0.7284
## Year2007 -0.3699 0.2439 -1.52 0.1348
## Year2008 -0.6163 0.2099 -2.94 0.0047 **
```

```

## Year2009          -0.2892      0.2007    -1.44    0.1549
## Year2010          -0.5627      0.2669    -2.11    0.0393 *
## Year2011          -0.6205      0.2042    -3.04    0.0035 **
## Year2012          -0.5265      0.2206    -2.39    0.0202 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.48
## Multiple R-squared:  0.318, Adjusted R-squared:  0.144
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.594  0.926  0.960  0.934  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      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 9682 1          98.398
## Year            9682 14          1.388

```

Residuals from last author



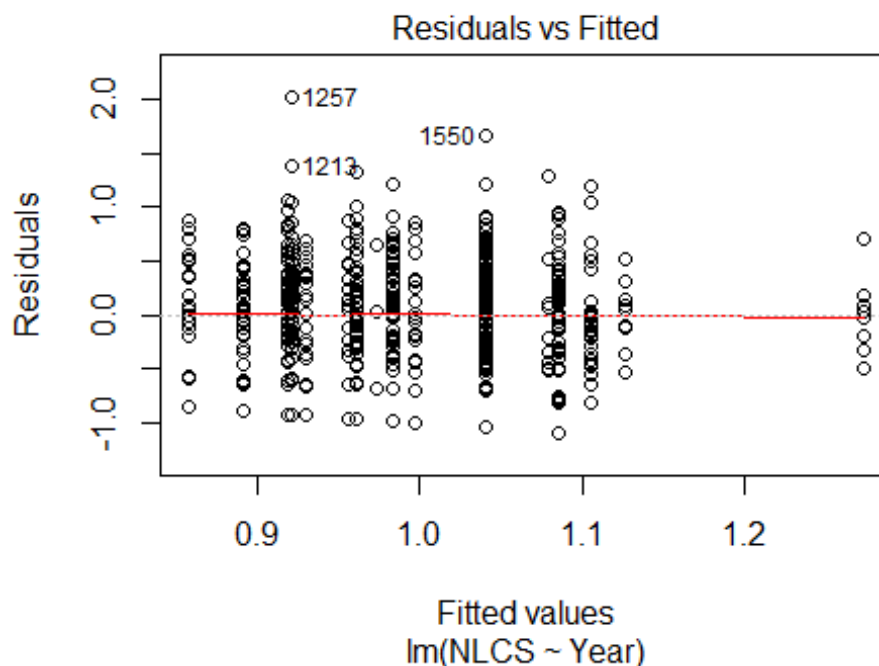
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.8255 -0.2836  0.0545  0.2443  0.8621
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2710     0.1664   7.64 2.3e-10 ***
## LastAuthorFemale1 -0.0266     0.1156  -0.23  0.819
## Year1999          0.2490     0.1665   1.50  0.140
## Year2000          0.2250     0.1664   1.35  0.182
## Year2001          0.2178     0.1957   1.11  0.270
## Year2002         -0.2015     0.1812  -1.11  0.271
## Year2003          0.2760     0.1664   1.66  0.103
## Year2004         -0.1764     0.2026  -0.87  0.388
## Year2005         -0.5753     0.2182  -2.64  0.011 *
## Year2006         -0.1102     0.3701  -0.30  0.767
## Year2007         -0.2974     0.2431  -1.22  0.226
## Year2008         -0.5493     0.2163  -2.54  0.014 *
```

```

## Year2009          -0.0859      0.2334   -0.37    0.714
## Year2010          -0.4455      0.2580   -1.73    0.089 .
## Year2011          -0.5303      0.2193   -2.42    0.019 *
## Year2012          -0.4573      0.2434   -1.88    0.065 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.486
## Multiple R-squared:  0.236, Adjusted R-squared:  0.0423
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.733  0.897  0.960  0.928  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      1.33e-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 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
##   54   59   45   36   44   76   59   65   76  102  107  123   94   99  115
## 2011 2012
##  116  117
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    9   18    8    3    9   20   22   23   31   45   57   62   52   55   73
## 2011 2012
##   69   69
##

```

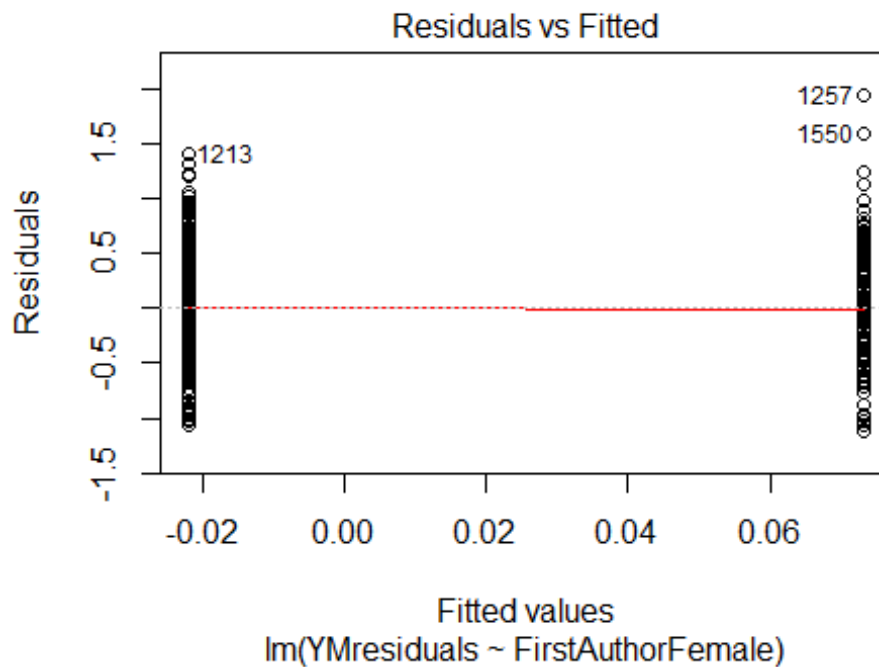
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7   17    7    2    5   12   14   12   25   37   42   51   42   45   65
## 2011 2012
##   61   56
## [1] "Heteroscedasticity checks, confirming that there are problems with
##      these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 13, df = 16, p-value = 0.7
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 5.2, df = 1, p-value = 0.02
##
## [1] "Female first author team size 2018 geometric mean: 4.93029265468846"
## [1] "Male first author team size 2018 geometric mean: 4.05240435733512"
##
## Warning in wilcox.test.default(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 = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.49559654719498"
## [1] "Male last author team size 2018 geometric mean: 4.13301303249998"

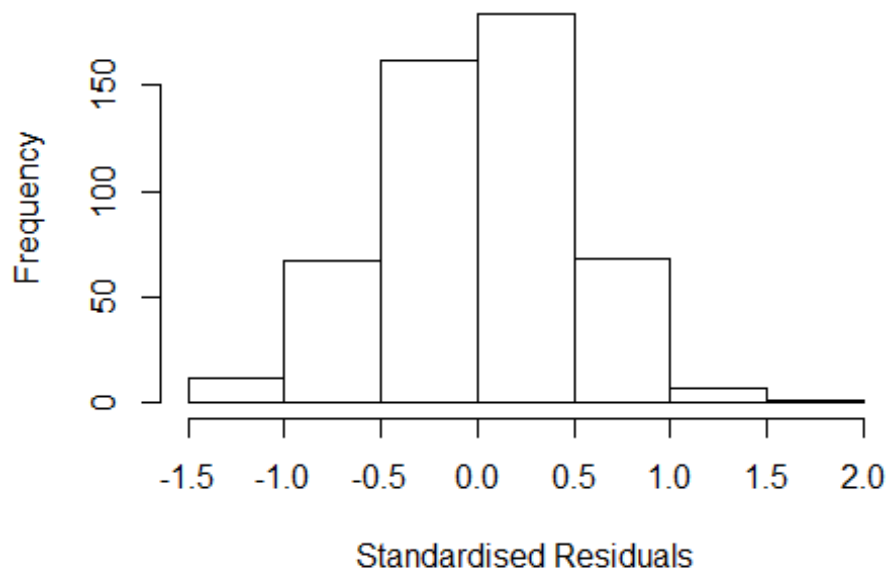
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 250, p-value = 0.06
## 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 | 1.138 | 1 | 1.067 |
| LastAuthorFemale | 1.221 | 1 | 1.105 |
| UniqueAuthors | 1.758 | 4 | 1.073 |
| Year | 2.029 | 16 | 1.022 |

Residuals from first and last author and team size



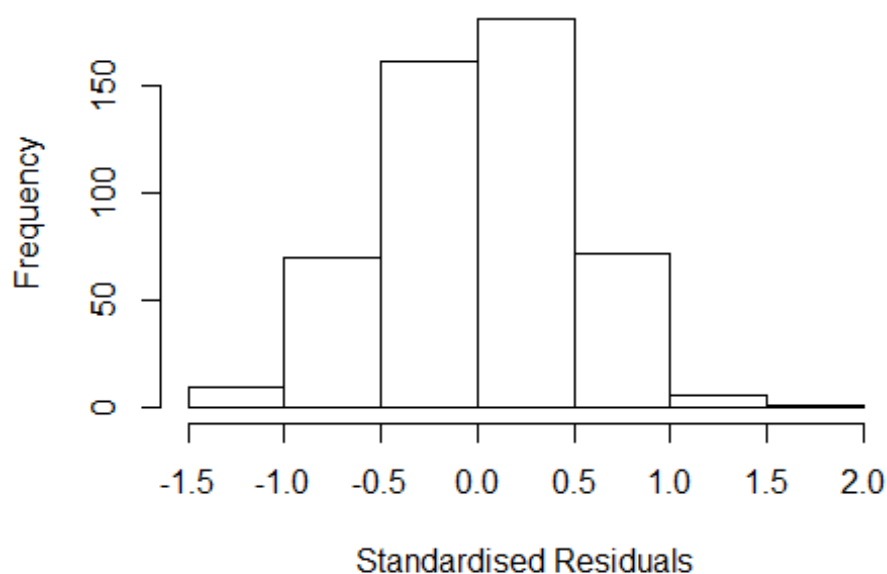
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1994 -0.3234 0.0279 0.3028 1.7926
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.70831 0.30458 2.33 0.0205 *
## FirstAuthorFemale1 0.12451 0.05711 2.18 0.0297 *
## LastAuthorFemale1 -0.13193 0.08112 -1.63 0.1045
## UniqueAuthors2 0.09260 0.19320 0.48 0.6320
## UniqueAuthors3 0.23672 0.18369 1.29 0.1981
## UniqueAuthors4 0.34184 0.18186 1.88 0.0608 .
## UniqueAuthors5 0.47839 0.17927 2.67 0.0079 **
## Year1997 0.00388 0.27785 0.01 0.9889
## Year1998 0.20883 0.26413 0.79 0.4295
## Year1999 -0.42186 0.48085 -0.88 0.3808
```

```

## Year2000      -0.04897    0.26218   -0.19    0.8519
## Year2001      0.09326    0.28165    0.33    0.7407
## Year2002     -0.13029    0.29079   -0.45    0.6543
## Year2003     -0.12847    0.29674   -0.43    0.6653
## Year2004      0.07052    0.26123    0.27    0.7873
## Year2005     -0.22728    0.25855   -0.88    0.3798
## Year2006     -0.12465    0.25901   -0.48    0.6306
## Year2007     -0.01530    0.25611   -0.06    0.9524
## Year2008     -0.17656    0.26457   -0.67    0.5049
## Year2009      0.01273    0.25708    0.05    0.9605
## Year2010     -0.16479    0.25682   -0.64    0.5214
## Year2011     -0.04701    0.25495   -0.18    0.8538
## Year2012     -0.04129    0.25596   -0.16    0.8719
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.49
## Multiple R-squared:  0.115, Adjusted R-squared:  0.0741
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 39 weights are ~= 1. The remaining 461 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.152  0.870  0.955  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      2.00e-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 1.098 1      1.048
## LastAuthorFemale  1.154 1      1.074
## Year              1.243 16      1.007

```

Residuals from first and last author



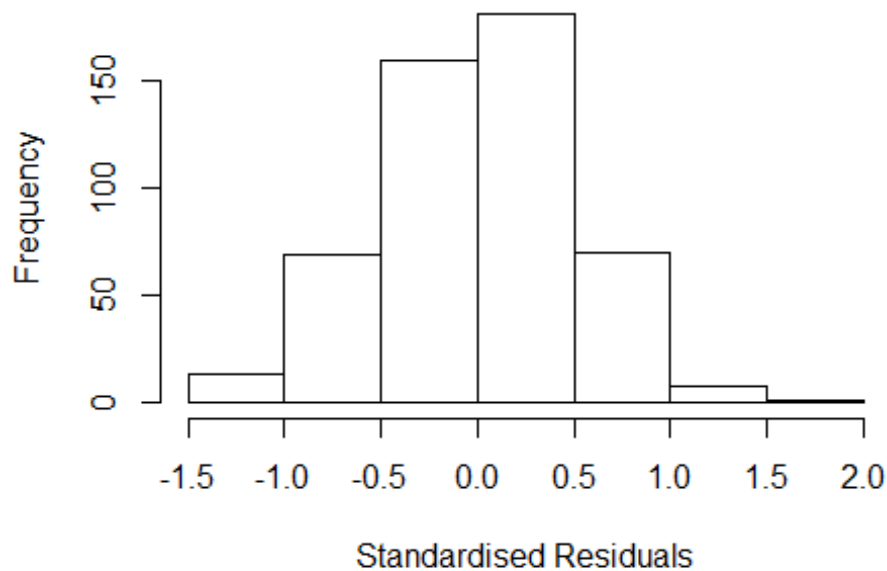
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1359 -0.3414 0.0241 0.3418 1.9659
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.02848 0.22036 4.67 4e-06 ***
## FirstAuthorFemale1 0.10036 0.06028 1.66 0.097 .
## LastAuthorFemale1 -0.16175 0.09071 -1.78 0.075 .
## Year1997 -0.07478 0.26034 -0.29 0.774
## Year1998 0.26463 0.24395 1.08 0.279
## Year1999 -0.38448 0.35730 -1.08 0.282
## Year2000 0.00838 0.23218 0.04 0.971
## Year2001 0.10570 0.25116 0.42 0.674
## Year2002 -0.16199 0.27074 -0.60 0.550
## Year2003 -0.07235 0.28881 -0.25 0.802
## Year2004 0.06594 0.24157 0.27 0.785
## Year2005 -0.18382 0.23363 -0.79 0.432
```

```

## Year2006      -0.05673    0.23656   -0.24    0.811
## Year2007      0.03865    0.23314    0.17    0.868
## Year2008     -0.13732    0.24116   -0.57    0.569
## Year2009     -0.00617    0.23416   -0.03    0.979
## Year2010     -0.15578    0.23193   -0.67    0.502
## Year2011      0.01372    0.22838    0.06    0.952
## Year2012      0.00702    0.23139    0.03    0.976
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.505
## Multiple R-squared:  0.0468, Adjusted R-squared:  0.0111
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 44 weights are ~= 1. The remaining 456 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.096  0.867   0.951   0.909   0.986   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      2.00e-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 1.085 1      1.042
## Year              1.085 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
## -1.120 -0.345 0.019 0.337 1.985
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.02909 0.22056 4.67 4e-06 ***
## FirstAuthorFemale1 0.09300 0.06047 1.54 0.12
## Year1997 -0.07434 0.26056 -0.29 0.78
## Year1998 0.26705 0.24447 1.09 0.28
## Year1999 -0.38509 0.35573 -1.08 0.28
## Year2000 0.00925 0.23260 0.04 0.97
## Year2001 0.10789 0.25135 0.43 0.67
## Year2002 -0.16563 0.26772 -0.62 0.54
## Year2003 -0.08938 0.27974 -0.32 0.75
## Year2004 0.05800 0.24130 0.24 0.81
## Year2005 -0.18839 0.23367 -0.81 0.42
## Year2006 -0.06635 0.23710 -0.28 0.78
```

```

## Year2007          0.03675    0.23320    0.16    0.87
## Year2008          -0.16224    0.24075   -0.67    0.50
## Year2009          -0.02328    0.23532   -0.10    0.92
## Year2010          -0.16816    0.23200   -0.72    0.47
## Year2011          -0.01710    0.22824   -0.07    0.94
## Year2012          -0.00242    0.23189   -0.01    0.99
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.506
## Multiple R-squared:  0.0399, Adjusted R-squared:  0.00599
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 46 weights are ~= 1. The remaining 454 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0899 0.8660 0.9510 0.9090 0.9860 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      2.00e-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 1.13 1          1.063
## Year          1.13 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
## -1.086 -0.338  0.025  0.331  2.036

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.02831    0.22039   4.67 4e-06 ***
## LastAuthorFemale1 -0.15025    0.08920  -1.68  0.093 .
## Year1997        -0.05888    0.25952  -0.23  0.821
## Year1998         0.30540    0.24774   1.23  0.218
## Year1999        -0.38431    0.35576  -1.08  0.281
## Year2000         0.02859    0.23506   0.12  0.903
## Year2001         0.14352    0.24876   0.58  0.564
## Year2002        -0.13440    0.26835  -0.50  0.617
## Year2003        -0.06105    0.28817  -0.21  0.832
## Year2004         0.07793    0.24104   0.32  0.747
## Year2005        -0.16420    0.23341  -0.70  0.482
## Year2006        -0.03512    0.23641  -0.15  0.882
## Year2007         0.05811    0.23315   0.25  0.803
## Year2008        -0.11537    0.24113  -0.48  0.633
## Year2009         0.00777    0.23373   0.03  0.973
## Year2010        -0.12561    0.23158  -0.54  0.588
## Year2011         0.03254    0.22813   0.14  0.887
## Year2012         0.02989    0.23186   0.13  0.897
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.505
## Multiple R-squared:  0.0405, Adjusted R-squared:  0.00671
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 42 weights are ~= 1. The remaining 458 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0668 0.8680 0.9540 0.9090 0.9870 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      2.00e-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: 500"
## [1] ""

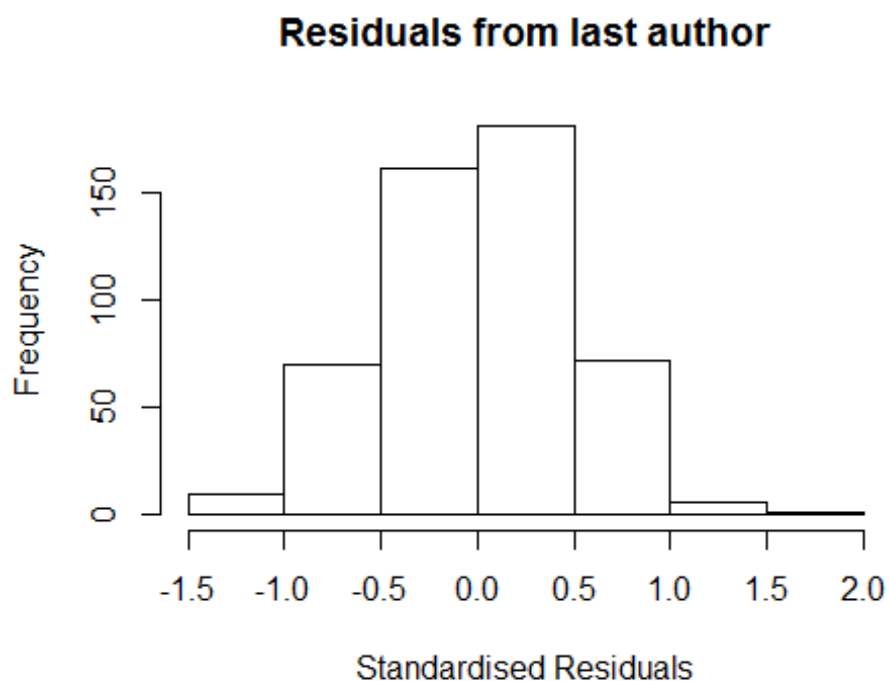
```

```

## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2747"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2009 2010 2011
##    2    7    4    3    6    4    8    4    9    2    6    6    7    10    7
## 2012
##    7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2009 2010 2011
##    0    1    2    0    2    2    7    3    4    2    3    3    5    6    6
## 2012
##    4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2009 2010 2011
##    0    1    2    0    0    0    4    1    2    2    3    2    2    6    5
## 2012
##    4
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 7.48331477354788"
## [1] "Male first author team size 2018 geometric mean: 6.86828545531999"

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

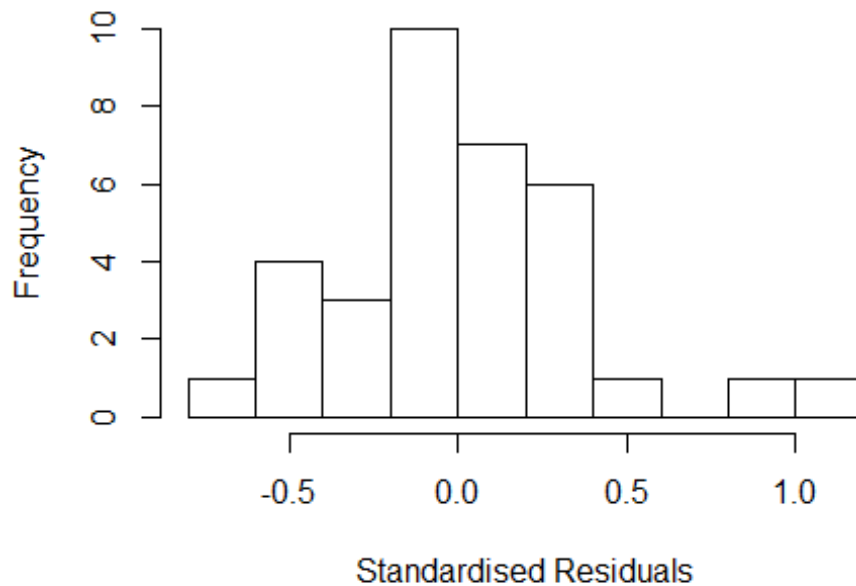
```

```
##
## 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: NaN"
## [1] "Male last author team size 2018 geometric mean: 7.10798671525346"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
```

| | GVIF | Df | GVIF ^{1/(2*Df)} |
|-------------------|-----------|----|--------------------------|
| FirstAuthorFemale | 1.646e+03 | 1 | 40.569 |
| LastAuthorFemale | 6.145e+00 | 1 | 2.479 |
| UniqueAuthors | 7.202e+18 | 3 | 1389.677 |
| Year | 7.061e+20 | 11 | 8.865 |

Residuals from first and last author and team size



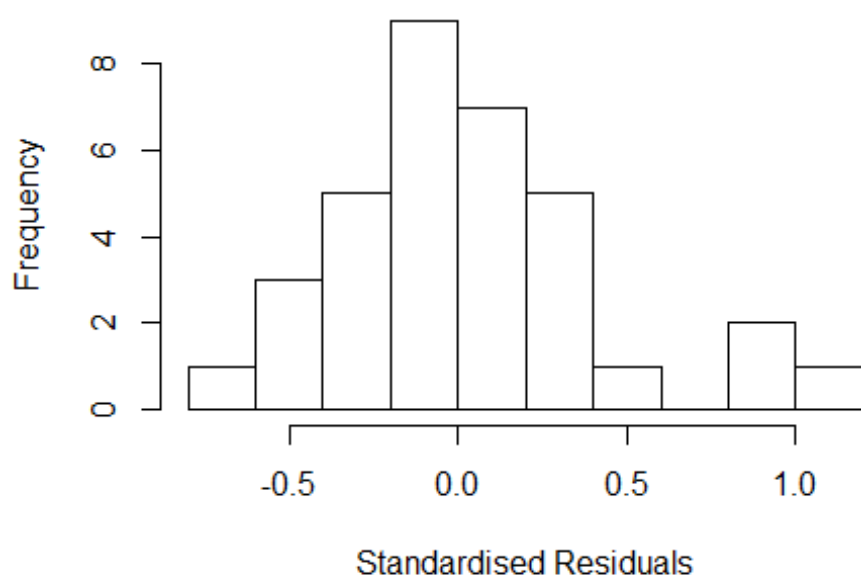
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -6.23e-01 -1.67e-01 -1.11e-16 2.12e-01 1.05e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.0216 0.4074 0.05 0.9583
## FirstAuthorFemale1 0.6153 0.2904 2.12 0.0491 *
## LastAuthorFemale1 0.3302 0.2748 1.20 0.2460
## UniqueAuthors3 -0.1281 0.1922 -0.67 0.5140
## UniqueAuthors4 0.4045 0.3022 1.34 0.1983
## UniqueAuthors5 0.4821 0.2869 1.68 0.1112
## Year1998 0.9713 0.2963 3.28 0.0044 **
## Year2002 0.4750 0.3726 1.27 0.2195
## Year2003 -0.1279 0.2869 -0.45 0.6612
## Year2004 0.0667 0.2914 0.23 0.8217
```

```

## Year2005          0.2835      0.1263      2.25      0.0383 *
## Year2006          0.2012      0.2828      0.71      0.4865
## Year2007          0.7866      0.3505      2.24      0.0384 *
## Year2009          0.0961      0.3880      0.25      0.8074
## Year2010          0.0793      0.5448      0.15      0.8859
## Year2011          0.0566      0.2505      0.23      0.8238
## Year2012         -0.2823      0.3046     -0.93      0.3670
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.496
## Multiple R-squared:  0.546, Adjusted R-squared:  0.119
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 26 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.636  0.911  0.976  0.937  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 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 120.0 1      10.95
## LastAuthorFemale  457.6 1      21.39
## Year              40846.3 11      1.62

```

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
## -6.25e-01 -2.22e-01 -6.11e-16  2.05e-01  1.03e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    5.38e-01   3.35e-01   1.61e+00   0.124
## FirstAuthorFemale1 5.81e-01   3.35e-01   1.73e+00   0.098 .
## LastAuthorFemale1 2.01e-01   2.53e-01   7.90e-01   0.436
## Year1998         9.37e-01   3.40e-01   2.75e+00   0.012 *
## Year2002        -8.28e-03   3.15e-01  -3.00e-02   0.979
## Year2003        -6.10e-01   2.58e-08  -2.36e+07 <2e-16 ***
## Year2004        -1.49e-01   1.27e-01  -1.17e+00   0.255
## Year2005         2.83e-01   1.26e-01   2.24e+00   0.036 *
## Year2006         2.03e-01   2.84e-01   7.20e-01   0.482
## Year2007         7.13e-01   3.37e-01   2.12e+00   0.047 *
## Year2009         1.43e-01   3.45e-01   4.10e-01   0.683
## Year2010         2.54e-02   5.03e-01   5.00e-02   0.960
```

```

## Year2011          -1.94e-01   2.09e-01 -9.30e-01   0.363
## Year2012          -3.01e-01   2.73e-01 -1.10e+00   0.283
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.479
## Multiple R-squared:  0.501, Adjusted R-squared:  0.176
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 28 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.624  0.914  0.965   0.930   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           2.94e-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"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 101.3  1          10.066
## Year              101.3 11          1.234
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6.27e-01 -2.06e-01  1.11e-16  1.92e-01  1.06e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   5.07e-01   3.08e-01   1.65e+00   0.1146

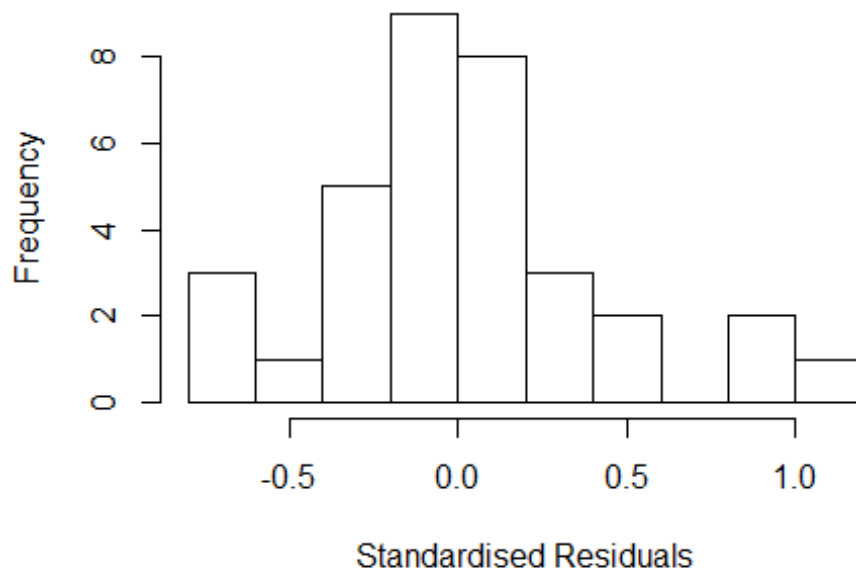
```

```

## FirstAuthorFemale1  6.12e-01  3.08e-01  1.98e+00  0.0605 .
## Year1998            9.68e-01  3.14e-01  3.08e+00  0.0056 **
## Year2002            9.27e-03  3.03e-01  3.00e-02  0.9759
## Year2003           -6.10e-01  1.98e-08 -3.08e+07  <2e-16 ***
## Year2004           -4.80e-02  7.78e-02 -6.20e-01  0.5441
## Year2005            2.83e-01  1.27e-01  2.24e+00  0.0361 *
## Year2006            2.05e-01  2.85e-01  7.20e-01  0.4803
## Year2007            7.44e-01  3.11e-01  2.40e+00  0.0260 *
## Year2009            2.59e-01  2.27e-01  1.14e+00  0.2663
## Year2010            9.39e-02  4.53e-01  2.10e-01  0.8380
## Year2011           -1.77e-01  1.92e-01 -9.30e-01  0.3651
## Year2012           -2.98e-01  2.72e-01 -1.09e+00  0.2867
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.468
## Multiple R-squared:  0.5,   Adjusted R-squared:  0.214
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 30 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.589  0.913  0.980  0.931  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           2.94e-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
## -0.7247 -0.1663  0.0263  0.2087  1.5577
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.1190     0.0000   Inf < 2e-16 ***
## LastAuthorFemale1 0.3675     0.1650    2.23  0.0370 *
## Year1998        0.3560     0.0589    6.04  5.4e-06 ***
## Year2002       -0.3943     0.2860   -1.38  0.1826
## Year2003       -0.6100     0.0000  -Inf < 2e-16 ***
## Year2004       -0.2318     0.0981   -2.36  0.0279 *
## Year2005        0.2835     0.1257    2.26  0.0349 *
## Year2006        0.1960     0.2708    0.72  0.4771
```

```

## Year2007          0.1325      0.0372      3.56      0.0018 **
## Year2009          -0.2303      0.1077     -2.14      0.0444 *
## Year2010          -0.5997      0.2864     -2.09      0.0486 *
## Year2011          -0.5309      0.2513     -2.11      0.0468 *
## Year2012          -0.5032      0.2476     -2.03      0.0550 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.552
## Multiple R-squared:  0.394, Adjusted R-squared:  0.047
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 30 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.406  0.902  0.986  0.938  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      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 2748"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7    7   14    5   11   11   15    9   13    8   14   17   19   16   11
## 2011 2012
##   22   23
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    2    1    3    4    7    4    9    7   10   13   11   13    7
## 2011 2012
##   15   17

```

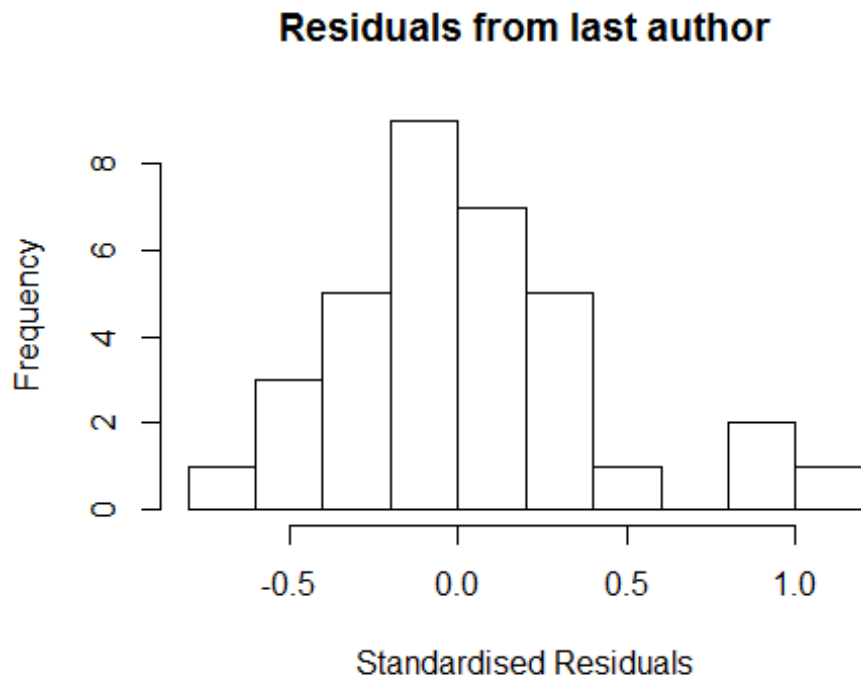


```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    2    1    2    4    7    4    8    7   10   11   10   10    6
## 2011 2012
##   14   14
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.87298334620742"
## [1] "Male first author team size 2018 geometric mean: 3.94870097166964"

## Warning in wilcox.test.default(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: 2.82842712474619"
## [1] "Male last author team size 2018 geometric mean: 4.47769492694043"

## Warning in wilcox.test.default(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 = 0.3
## 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 1.172e+00  1      1.0827
## LastAuthorFemale  9.543e-01  1      0.9769
## UniqueAuthors    1.127e+16  4     101.5002
## Year              2.183e+16 16      3.2404

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##      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 -2.54e-01 -1.86e-15  2.71e-01  1.07e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.13e+00   3.52e-01   3.22e+00  0.0018 **
## FirstAuthorFemale1 -1.27e-01   9.74e-02  -1.30e+00  0.1957
## LastAuthorFemale1  3.19e-01   2.22e-01   1.43e+00  0.1550
## UniqueAuthors2    3.46e-01   3.52e-01   9.90e-01  0.3271
## UniqueAuthors3    4.73e-01   2.55e-01   1.86e+00  0.0668 .
## UniqueAuthors4    3.82e-01   2.55e-01   1.50e+00  0.1374
## UniqueAuthors5    4.68e-01   2.23e-01   2.10e+00  0.0390 *
## Year1997          5.90e-01   6.36e-08   9.28e+06 <2e-16 ***
## Year1998          7.03e-02   3.55e-01   2.00e-01  0.8436
## Year1999         -7.73e-01   2.78e-01  -2.78e+00  0.0066 **
## Year2000         -6.07e-01   5.93e-01  -1.02e+00  0.3085
## Year2001         -2.60e-01   3.16e-01  -8.20e-01  0.4116
## Year2002         -1.14e-01   3.09e-01  -3.70e-01  0.7120
## Year2003         -5.28e-01   3.73e-01  -1.42e+00  0.1597
## Year2004         -4.28e-01   2.50e-01  -1.71e+00  0.0899 .
## Year2005         -6.48e-01   2.43e-01  -2.66e+00  0.0091 **
## Year2006         -3.02e-01   3.18e-01  -9.50e-01  0.3452
## Year2007         -4.24e-01   3.13e-01  -1.36e+00  0.1783
```

```

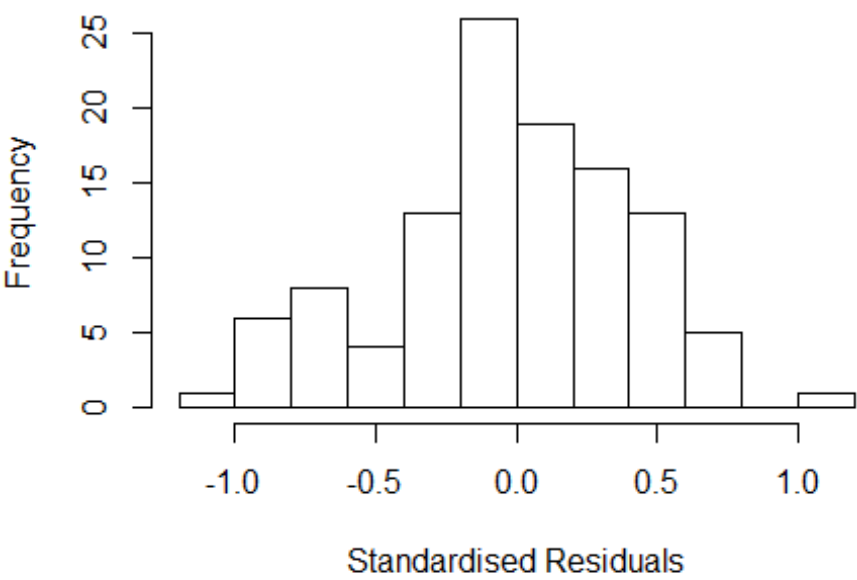
## Year2008          -2.15e-01   3.38e-01 -6.40e-01   0.5267
## Year2009          -3.87e-01   3.28e-01 -1.18e+00   0.2418
## Year2010          -6.92e-01   4.07e-01 -1.70e+00   0.0924 .
## Year2011          -6.43e-01   2.94e-01 -2.19e+00   0.0312 *
## Year2012          -7.50e-01   3.42e-01 -2.19e+00   0.0311 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.444
## Multiple R-squared:  0.25,   Adjusted R-squared:  0.0649
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 96 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.538  0.855  0.940  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      8.93e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##  nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##  trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"

## 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

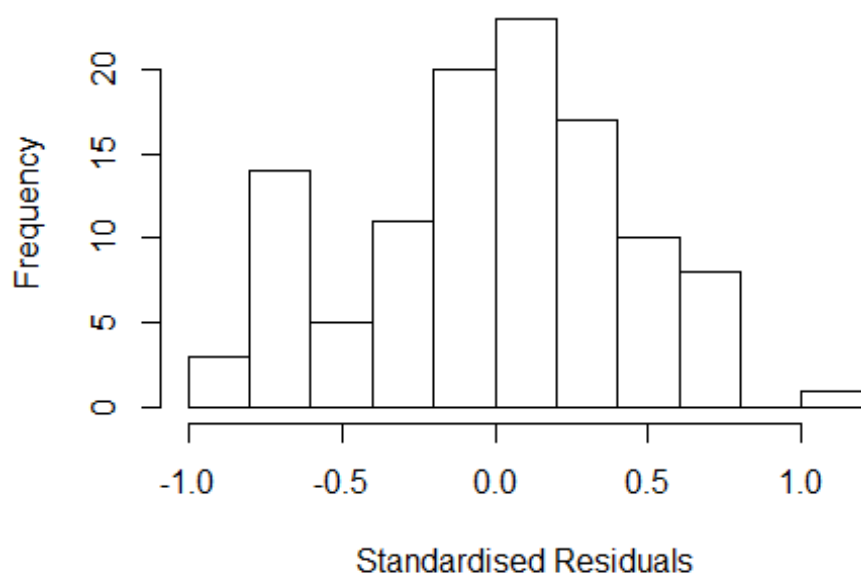
```

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 |

Residuals from first and last author



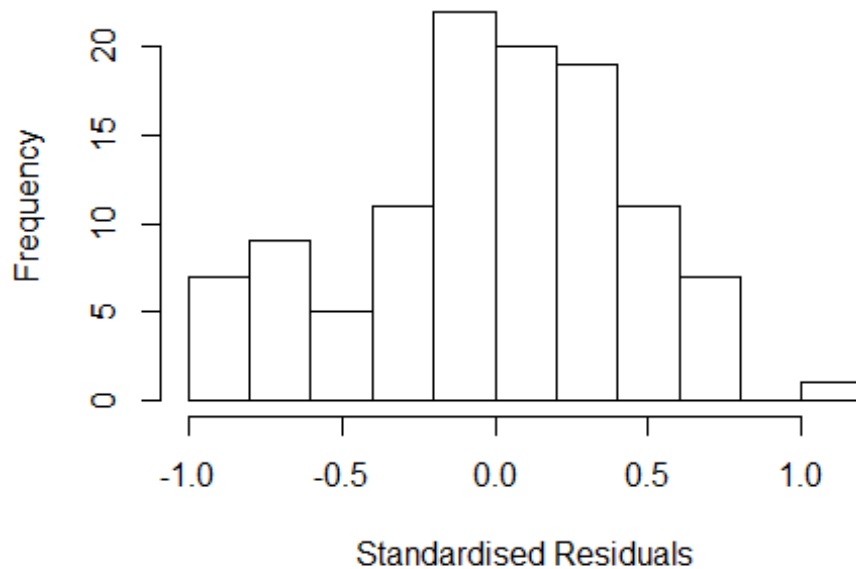
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9879 -0.2838 0.0186 0.2663 1.1270
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.47900 0.00000 Inf < 2e-16 ***
## FirstAuthorFemale1 -0.12105 0.09284 -1.30 0.19550
## LastAuthorFemale1 0.31301 0.19887 1.57 0.11889
## Year1997 0.59000 0.00000 Inf < 2e-16 ***
## Year1998 0.18853 0.22003 0.86 0.39375
## Year1999 -0.65200 0.00000 -Inf < 2e-16 ***
## Year2000 -0.71750 1.19390 -0.60 0.54932
## Year2001 -0.17943 0.13919 -1.29 0.20054
## Year2002 -0.00842 0.13247 -0.06 0.94948
## Year2003 -0.41051 0.25147 -1.63 0.10597
## Year2004 -0.32962 0.08675 -3.80 0.00026 ***
## Year2005 -0.57596 0.23683 -2.43 0.01693 *
```

```

## Year2006      -0.20049    0.16004   -1.25   0.21342
## Year2007      -0.32065    0.12582   -2.55   0.01246 *
## Year2008      -0.31815    0.14152   -2.25   0.02693 *
## Year2009      -0.35213    0.19429   -1.81   0.07315 .
## Year2010      -0.59501    0.28103   -2.12   0.03691 *
## Year2011      -0.53948    0.08808   -6.13   2.1e-08 ***
## Year2012      -0.68304    0.22606   -3.02   0.00325 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.468
## Multiple R-squared:  0.209, Adjusted R-squared:  0.0564
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 96 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.541  0.854  0.956  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      8.93e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500           50           2           1           1000      200
##      trace.lev      mts      compute.rd
##      0             1000      0
##      psi          subsampling          cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.054 1      1.433
## Year              2.054 16      1.023

```

Residuals from first author



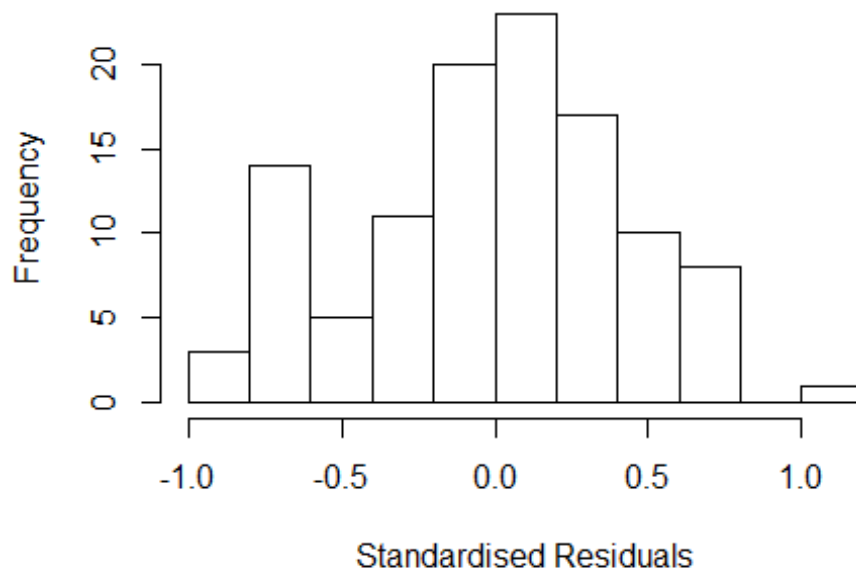
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.908 -0.272 0.016 0.275 1.015
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.48e+00 2.88e-08 5.13e+07 < 2e-16 ***
## FirstAuthorFemale1 -8.09e-02 9.92e-02 -8.20e-01 0.41687
## Year1997 5.90e-01 5.52e-08 1.07e+07 < 2e-16 ***
## Year1998 1.68e-01 2.38e-01 7.10e-01 0.48108
## Year1999 -6.52e-01 1.69e-08 -3.87e+07 < 2e-16 ***
## Year2000 -7.17e-01 1.19e+00 -6.00e-01 0.54694
## Year2001 -1.90e-01 1.36e-01 -1.40e+00 0.16575
## Year2002 -1.51e-02 1.32e-01 -1.10e-01 0.90908
## Year2003 -4.11e-01 2.52e-01 -1.63e+00 0.10598
## Year2004 -3.44e-01 8.55e-02 -4.02e+00 0.00012 ***
## Year2005 -5.86e-01 2.38e-01 -2.46e+00 0.01581 *
## Year2006 -2.15e-01 1.59e-01 -1.36e+00 0.17861
```

```

## Year2007          -3.09e-01    1.17e-01 -2.63e+00    0.00992 **
## Year2008          -2.98e-01    1.52e-01 -1.96e+00    0.05258 .
## Year2009          -3.10e-01    1.94e-01 -1.60e+00    0.11357
## Year2010          -6.15e-01    2.82e-01 -2.18e+00    0.03159 *
## Year2011          -5.35e-01    8.84e-02 -6.05e+00    2.9e-08 ***
## Year2012          -5.71e-01    2.33e-01 -2.45e+00    0.01607 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.469
## Multiple R-squared:  0.185, Adjusted R-squared:  0.0375
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 98 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.618  0.861  0.958  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      8.93e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.411 1          1.188
## Year            1.411 16          1.011

```


Residuals from last author



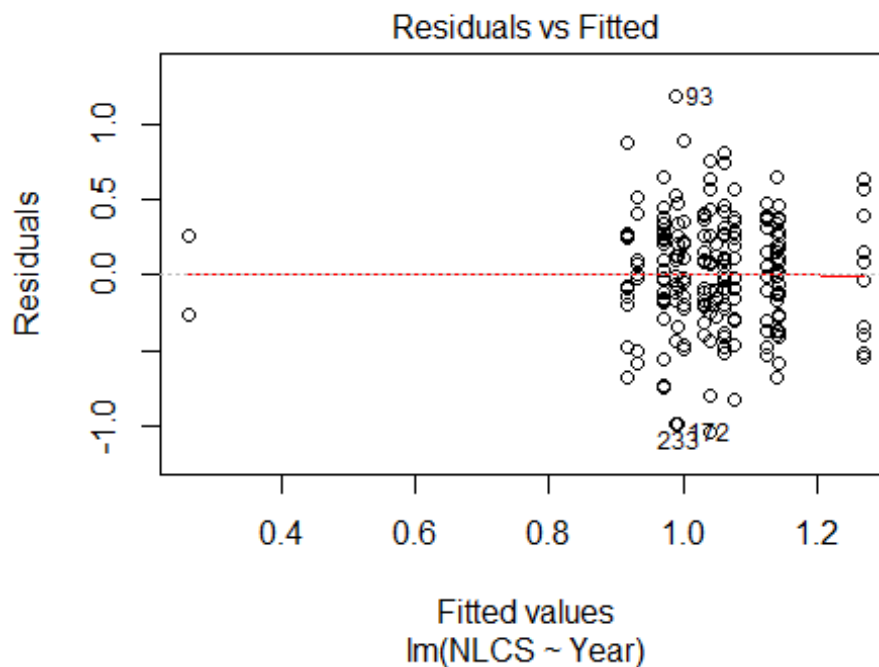
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.051 -0.282 0.040 0.279 1.137
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.48e+00 4.91e-08 3.01e+07 < 2e-16 ***
## LastAuthorFemale1 2.65e-01 2.09e-01 1.27e+00 0.2066
## Year1997 5.90e-01 5.14e-08 1.15e+07 < 2e-16 ***
## Year1998 1.28e-01 2.71e-01 4.70e-01 0.6383
## Year1999 -6.52e-01 5.43e-08 -1.20e+07 < 2e-16 ***
## Year2000 -7.17e-01 1.24e+00 -5.80e-01 0.5648
## Year2001 -2.11e-01 1.27e-01 -1.67e+00 0.0986 .
## Year2002 -2.77e-02 1.29e-01 -2.10e-01 0.8305
## Year2003 -4.09e-01 2.52e-01 -1.62e+00 0.1076
## Year2004 -3.72e-01 7.96e-02 -4.67e+00 9.9e-06 ***
## Year2005 -6.05e-01 2.44e-01 -2.48e+00 0.0150 *
## Year2006 -2.45e-01 1.46e-01 -1.68e+00 0.0964 .
```

```

## Year2007          -3.70e-01    1.12e-01 -3.31e+00    0.0013 **
## Year2008          -3.49e-01    1.47e-01 -2.37e+00    0.0197 *
## Year2009          -3.82e-01    1.92e-01 -1.99e+00    0.0491 *
## Year2010          -6.51e-01    2.81e-01 -2.32e+00    0.0227 *
## Year2011          -5.96e-01    7.96e-02 -7.49e+00    3.7e-11 ***
## Year2012          -6.93e-01    2.39e-01 -2.89e+00    0.0047 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.462
## Multiple R-squared:  0.199, Adjusted R-squared:  0.0536
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 97 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.524  0.836  0.957  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      8.93e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 112"
## [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
##   10   21   22   21   24   28   14   15   21   22   20   34   23   17   27
## 2011 2012
##   21   28
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2   10   10   11    8    6   11    8   13   13   14   22   13   14   16
## 2011 2012

```

```
## 16 20
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2 10 10 10 5 5 11 7 10 12 12 18 11 14 14
## 2011 2012
## 15 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 = 18, df = 16, p-value = 0.3
```



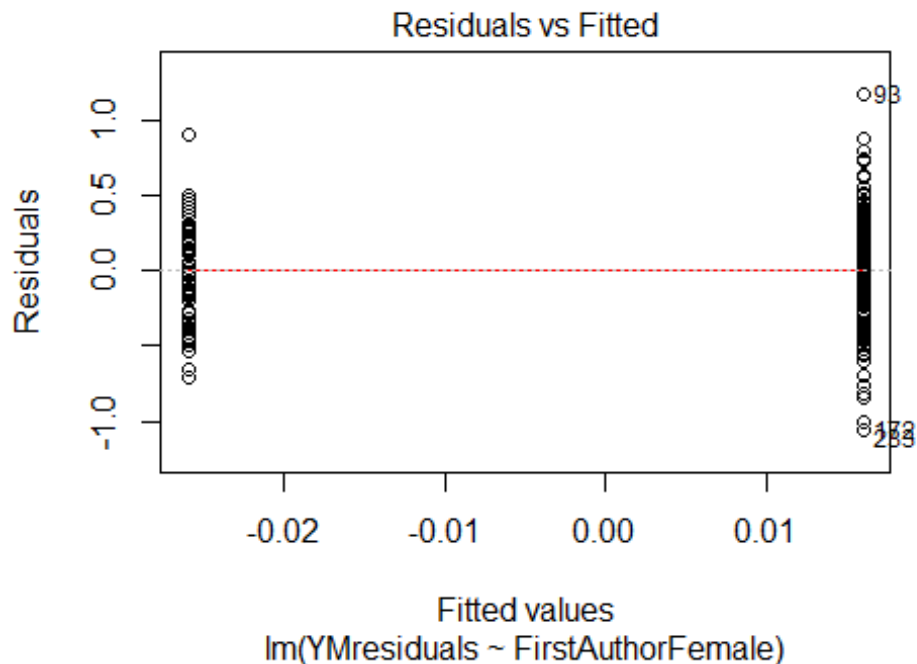
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 4.8, df = 1, p-value = 0.03

## [1] "Female first author team size 2018 geometric mean: 3.8221305957095"
## [1] "Male first author team size 2018 geometric mean: 3.5065589508876"

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

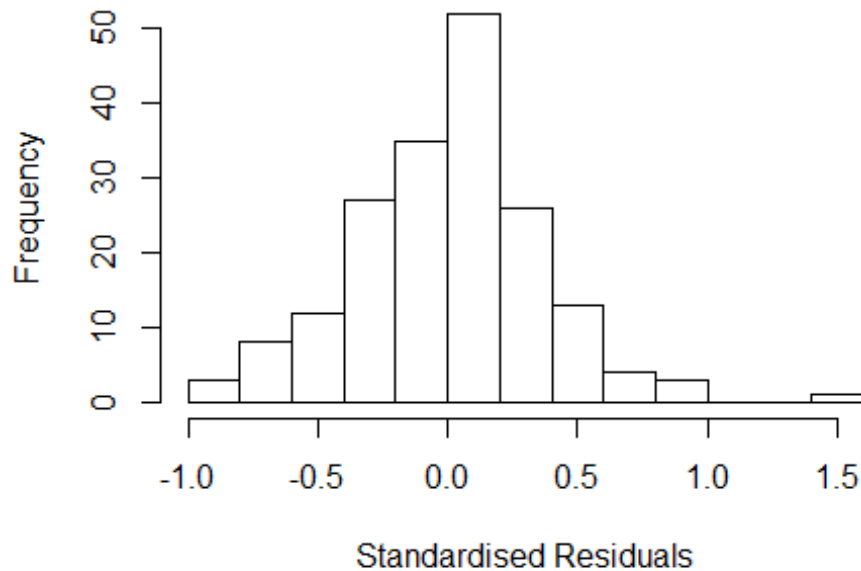
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 80, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.20734182673286"
## [1] "Male last author team size 2018 geometric mean: 4.02371933548209"

## Warning in wilcox.test.default(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 = 0.6
## 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 1.468 1      1.211
## LastAuthorFemale 1.552 1      1.246
## UniqueAuthors    8.130 4      1.299
## Year              8.348 16     1.069
```

Residuals from first and last author and team size



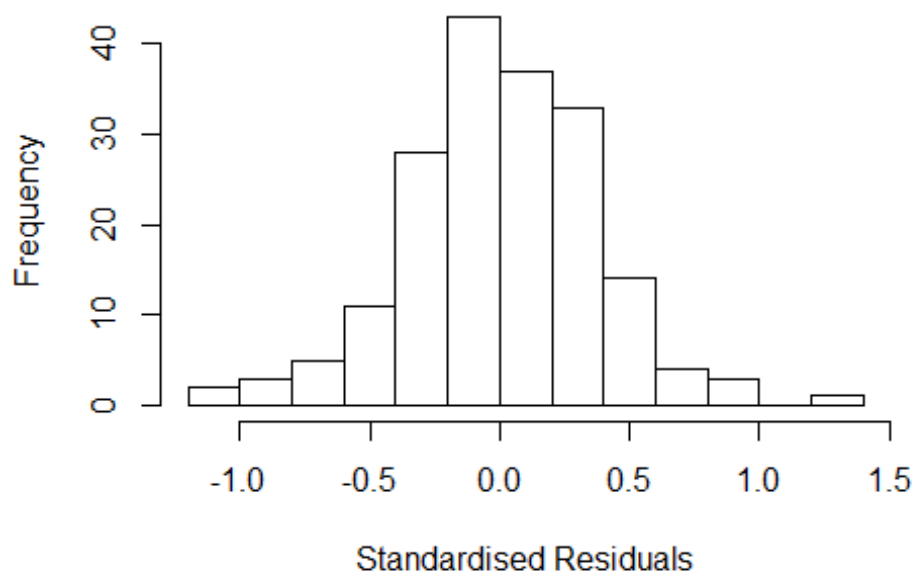
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9719 -0.2341 0.0279 0.2014 1.4244
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.1257 0.2799 0.45 0.65394
## FirstAuthorFemale1 -0.1183 0.0569 -2.08 0.03905 *
## LastAuthorFemale1 0.1318 0.0654 2.01 0.04561 *
## UniqueAuthors2 0.1082 0.2111 0.51 0.60879
## UniqueAuthors3 0.2787 0.2306 1.21 0.22857
## UniqueAuthors4 0.2108 0.2079 1.01 0.31209
## UniqueAuthors5 0.2394 0.2141 1.12 0.26505
## Year1997 1.0457 0.2615 4.00 9.7e-05 ***
## Year1998 0.6542 0.2316 2.82 0.00533 **
## Year1999 0.6269 0.2328 2.69 0.00783 **
```

```

## Year2000          0.6355      0.2576      2.47  0.01466 *
## Year2001          0.6880      0.2169      3.17  0.00181 **
## Year2002          0.7679      0.2294      3.35  0.00101 **
## Year2003          0.8461      0.2252      3.76  0.00024 ***
## Year2004          0.7047      0.2094      3.37  0.00096 ***
## Year2005          0.8080      0.2550      3.17  0.00183 **
## Year2006          0.8660      0.2080      4.16  5.1e-05 ***
## Year2007          0.7347      0.1976      3.72  0.00028 ***
## Year2008          0.7170      0.2295      3.12  0.00211 **
## Year2009          0.8000      0.2136      3.75  0.00025 ***
## Year2010          0.6848      0.2035      3.36  0.00096 ***
## Year2011          0.8455      0.1997      4.23  3.8e-05 ***
## Year2012          0.8644      0.1964      4.40  1.9e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.346
## Multiple R-squared:  0.19, Adjusted R-squared:  0.0788
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 169 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0517 0.8720 0.9550 0.8990 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          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 1.224 1 1.106
## LastAuthorFemale 1.259 1 1.122
## Year 1.534 16 1.013

```

Residuals from first and last author



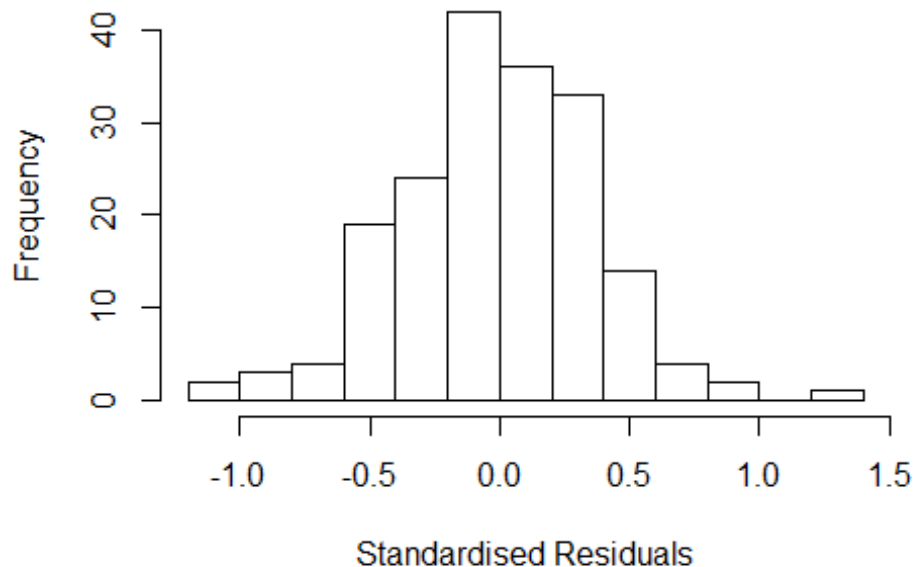
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.18804 -0.23633 0.00011 0.22601 1.25578
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.3149 0.2594 1.21 0.2265
## FirstAuthorFemale1 -0.1098 0.0549 -2.00 0.0471 *
## LastAuthorFemale1 0.1472 0.0662 2.22 0.0276 *
## Year1997 0.9568 0.3036 3.15 0.0019 **
## Year1998 0.5914 0.2883 2.05 0.0418 *
## Year1999 0.6063 0.2942 2.06 0.0409 *
## Year2000 0.6240 0.3080 2.03 0.0444 *
## Year2001 0.6787 0.2718 2.50 0.0135 *
## Year2002 0.7568 0.2968 2.55 0.0117 *
## Year2003 0.8731 0.2877 3.03 0.0028 **
## Year2004 0.6753 0.2772 2.44 0.0159 *
## Year2005 0.8381 0.3154 2.66 0.0087 **
```

```

## Year2006          0.8518      0.2762      3.08      0.0024 **
## Year2007          0.7264      0.2686      2.70      0.0076 **
## Year2008          0.7574      0.2834      2.67      0.0083 **
## Year2009          0.8059      0.2762      2.92      0.0040 **
## Year2010          0.6959      0.2684      2.59      0.0104 *
## Year2011          0.8347      0.2727      3.06      0.0026 **
## Year2012          0.8812      0.2679      3.29      0.0012 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.37
## Multiple R-squared:  0.146, Adjusted R-squared:  0.0526
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 162 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.225  0.891  0.949  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      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 1.222 1      1.105
## Year              1.222 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
## -1.1733 -0.2298 -0.0139 0.2617 1.2339
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.3054 0.2467 1.24 0.21758
## FirstAuthorFemale1 -0.0908 0.0560 -1.62 0.10690
## Year1997 0.9746 0.2876 3.39 0.00088 ***
## Year1998 0.6448 0.2787 2.31 0.02192 *
## Year1999 0.6377 0.2838 2.25 0.02597 *
## Year2000 0.6301 0.2956 2.13 0.03451 *
## Year2001 0.7389 0.2617 2.82 0.00533 **
## Year2002 0.7845 0.2833 2.77 0.00626 **
## Year2003 0.8679 0.2802 3.10 0.00229 **
## Year2004 0.6950 0.2687 2.59 0.01055 *
## Year2005 0.8540 0.3034 2.81 0.00548 **
## Year2006 0.8648 0.2680 3.23 0.00151 **
```

```

## Year2007          0.7519      0.2577      2.92  0.00402 **
## Year2008          0.7896      0.2695      2.93  0.00388 **
## Year2009          0.8516      0.2663      3.20  0.00166 **
## Year2010          0.7192      0.2550      2.82  0.00538 **
## Year2011          0.8932      0.2595      3.44  0.00073 ***
## Year2012          0.9180      0.2549      3.60  0.00042 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.383
## Multiple R-squared:  0.124, Adjusted R-squared:  0.0337
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 168 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.279  0.893  0.956  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      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 1.301 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
## -1.1513 -0.2143  0.0111  0.2491  1.2899

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.2600    0.2027    1.28 0.20150
## LastAuthorFemale1 0.1264    0.0696    1.82 0.07127 .
## Year1997        0.9929    0.2656    3.74 0.00025 ***
## Year1998        0.5870    0.2419    2.43 0.01633 *
## Year1999        0.6271    0.2448    2.56 0.01130 *
## Year2000        0.6628    0.2515    2.64 0.00921 **
## Year2001        0.6983    0.2193    3.18 0.00173 **
## Year2002        0.7626    0.2457    3.10 0.00225 **
## Year2003        0.8913    0.2364    3.77 0.00023 ***
## Year2004        0.6844    0.2320    2.95 0.00364 **
## Year2005        0.8532    0.2830    3.01 0.00298 **
## Year2006        0.8823    0.2307    3.82 0.00019 ***
## Year2007        0.7352    0.2194    3.35 0.00100 ***
## Year2008        0.7475    0.2326    3.21 0.00157 **
## Year2009        0.8322    0.2263    3.68 0.00032 ***
## Year2010        0.7053    0.2140    3.30 0.00120 **
## Year2011        0.8489    0.2203    3.85 0.00017 ***
## Year2012        0.8868    0.2144    4.14 5.6e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.368
## Multiple R-squared:  0.132, Adjusted R-squared:  0.0429
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 166 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.193 0.889  0.951  0.903  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      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 2801"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 2010 2011 2012
##    1    3    3
##
## 2010 2011 2012
##    1    3    3
##
## 2010 2011 2012
##    0    1    3
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.89897948556636"
## [1] "Male first author team size 2018 geometric mean: 5.91607978309962"
##
## 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: 7.48331477354788"
## [1] "Male last author team size 2018 geometric mean: 3.87298334620742"
##
## Wilcoxon rank sum test
##
## 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"
## [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 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
##      3      5      1      6      2      4      4      2      8      7     10     11      9      8      7
## 2011 2012
##      9      6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      3      1      6      2      3      3      1      5      4      9      9      8      7      6
## 2011 2012
##      6      5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      3      1      5      2      3      3      1      5      4      7      8      8      7      6
## 2011 2012
##      5      5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 5.8557053508852"
## [1] "Male first author team size 2018 geometric mean: 4.67948669187545"

## Warning in wilcox.test.default(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 = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6.58623218930013"
## [1] "Male last author team size 2018 geometric mean: 4.2958782681456"

## Warning in wilcox.test.default(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] "Regression 1: First author gender, last author gender, team size,
Year as factors"

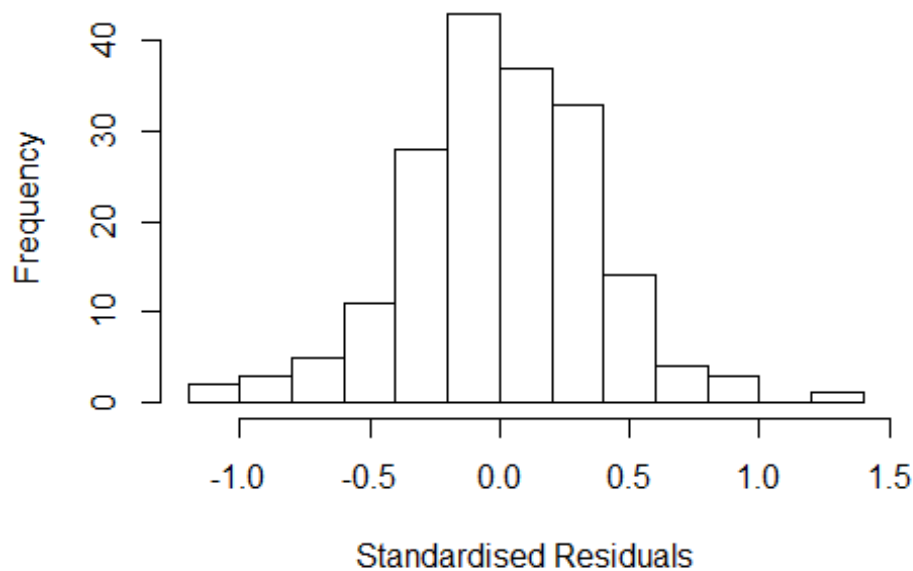
## Warning in lmrob.S(x, y, control = control, mf = mf): S refinements did
not
## converge (to refine.tol=1e-07) in 1000 (= k.max) steps

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

```

```
## Warning in lmrob.S(x, y, control = control, mf = mf): S refinements did
not
## converge (to refine.tol=1e-07) in 1000 (= k.max) steps
## 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 last author



```
## [1] "Regression 3: First author gender, Year as factors"
##               GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.849  1          1.962
## Year              3.849 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 ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.45e+00 -2.01e-01  5.27e-16  1.83e-01  7.75e-01
##
## Coefficients:
```

```

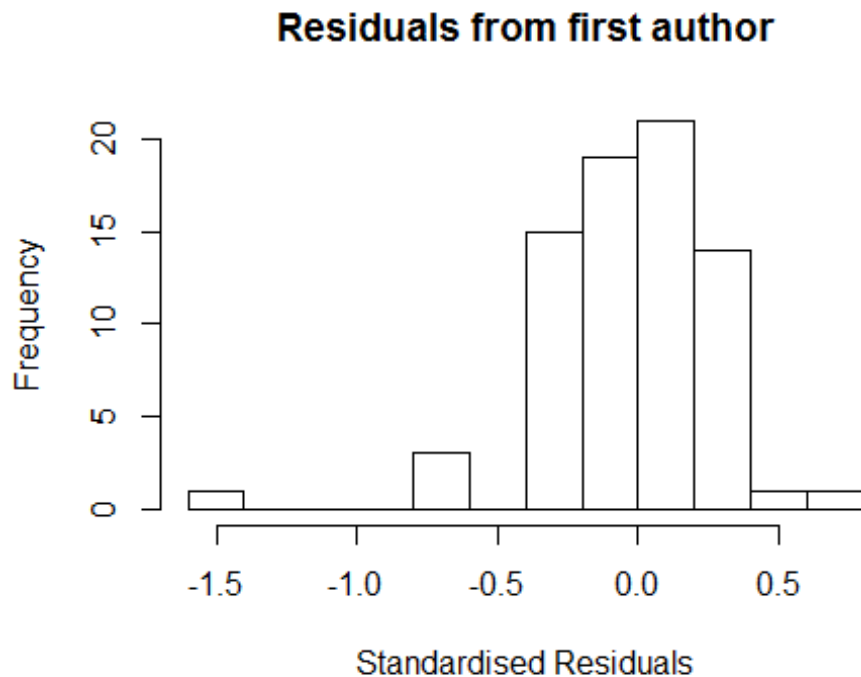
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.4204    0.3005   4.73 1.5e-05 ***
## FirstAuthorFemale1 -0.0608    0.0770  -0.79   0.43
## Year1997          0.1225    0.3296   0.37   0.71
## Year1998         -0.0924    0.3005  -0.31   0.76
## Year1999         -0.1847    0.3181  -0.58   0.56
## Year2000         -0.1739    0.3501  -0.50   0.62
## Year2001         -0.1673    0.3448  -0.49   0.63
## Year2002         -0.5275    0.3484  -1.51   0.14
## Year2003         -0.1564    0.3005  -0.52   0.60
## Year2004         -0.0765    0.3282  -0.23   0.82
## Year2005         -0.1052    0.3609  -0.29   0.77
## Year2006         -0.4350    0.3209  -1.36   0.18
## Year2007         -0.0716    0.3133  -0.23   0.82
## Year2008         -0.5699    0.3302  -1.73   0.09 .
## Year2009         -0.3311    0.3153  -1.05   0.30
## Year2010         -0.0114    0.3213  -0.04   0.97
## Year2011         -0.0993    0.3098  -0.32   0.75
## Year2012          0.0291    0.3205   0.09   0.93
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.324
## Multiple R-squared:  0.333, Adjusted R-squared:  0.134
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~ = 1. The remaining 69 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0076 0.9110 0.9620 0.9180 0.9860 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      1.33e-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"

## 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

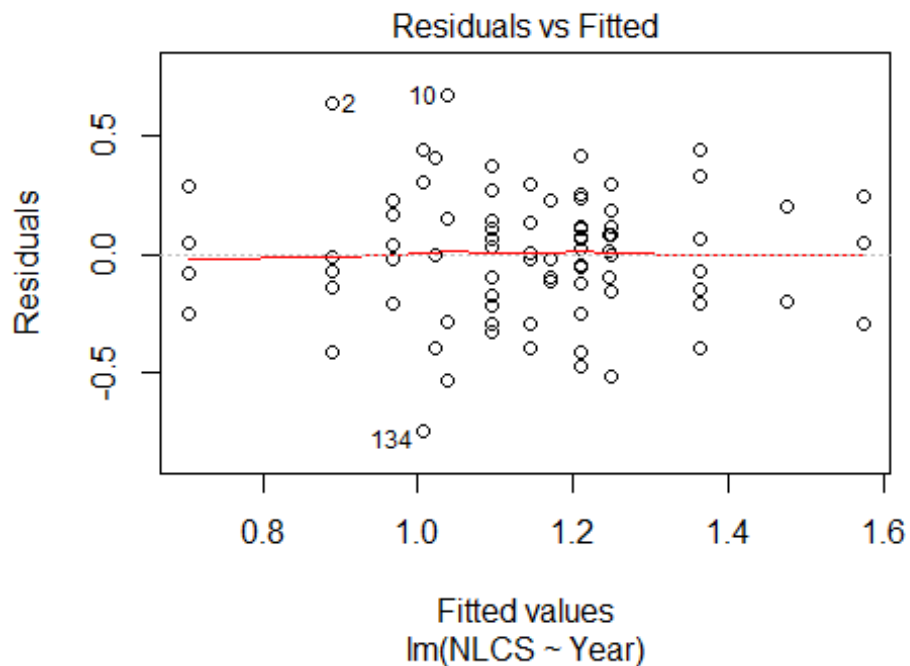
## Warning in lmrob.S(x, y, control = control, mf = mf): S refinements did
not
## converge (to refine.tol=1e-07) in 1000 (= k.max) steps
```



```
## [1] "Sample size for the above analysis: 75"
## [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
##    6    7    7    5    5    8    9    4    3    3    8    6   12    7    8
## 2011 2012
##    5   21
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    4    3    3    3    6    4    4    2    3    7    3    9    4    7
## 2011 2012
##    3   17
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    3    3    2    3    5    1    4    2    2    4    3    6    3    6
```



```
## 2011 2012
## 3 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 = 15, df = 16, p-value = 0.5
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2, df = 1, p-value = 0.2

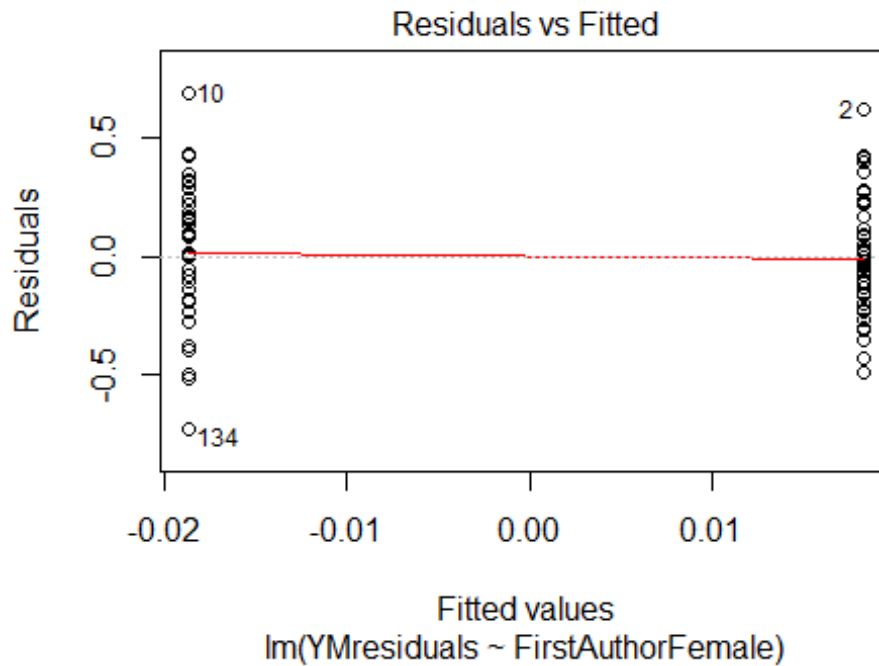
## [1] "Female first author team size 2018 geometric mean: 4.23605776337559"
## [1] "Male first author team size 2018 geometric mean: 2.6207413942089"

## 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 = 20, p-value = 0.1
```

```
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.16016764610381"
## [1] "Male last author team size 2018 geometric mean: 3.24534222319921"

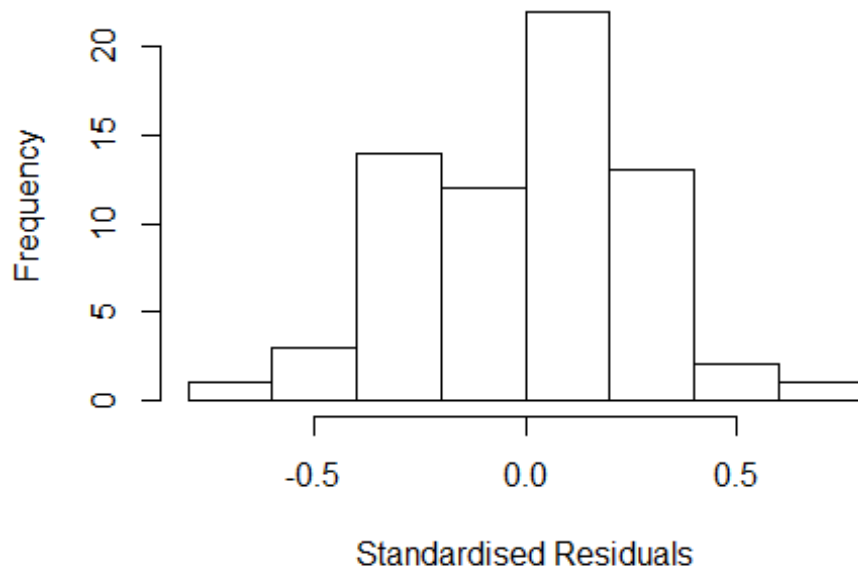
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 18, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
## Year as factors"
##
```

| | GVIF | Df | GVIF ^{1/(2*Df)} |
|-------------------|-----------|----|--------------------------|
| FirstAuthorFemale | 7.355 | 1 | 2.712 |
| LastAuthorFemale | 4.116 | 1 | 2.029 |
| UniqueAuthors | 11152.867 | 4 | 3.206 |
| Year | 31512.618 | 16 | 1.382 |

Residuals from first and last author and team size



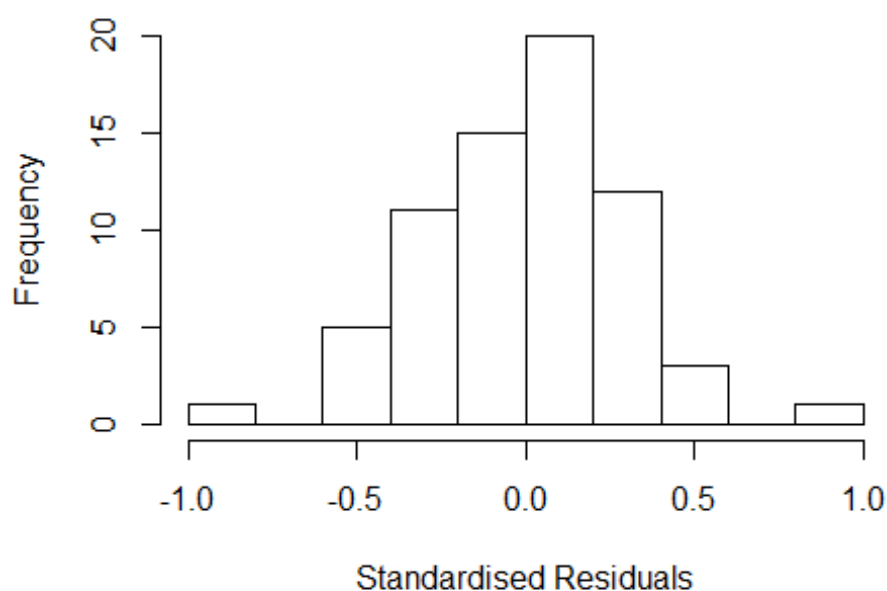
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.6485 -0.2018 0.0397 0.1948 0.7962
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.84445 0.32705 2.58 0.01315 *
## FirstAuthorFemale1 -0.02844 0.08550 -0.33 0.74093
## LastAuthorFemale1 -0.07378 0.11532 -0.64 0.52559
## UniqueAuthors2 0.06783 0.30153 0.22 0.82303
## UniqueAuthors3 -0.05818 0.37187 -0.16 0.87636
## UniqueAuthors4 0.24606 0.40876 0.60 0.55022
## UniqueAuthors5 -0.10253 0.30476 -0.34 0.73812
## Year1997 0.12556 0.45488 0.28 0.78379
## Year1998 0.47701 0.13176 3.62 0.00074 ***
## Year1999 -0.00895 0.24349 -0.04 0.97084
```

```

## Year2000      0.25102      0.21760      1.15      0.25476
## Year2001      0.16713      0.11052      1.51      0.13746
## Year2002     -0.39245      0.32705     -1.20      0.23643
## Year2003      0.39268      0.16886      2.33      0.02461 *
## Year2004      0.38449      0.36307      1.06      0.29525
## Year2005      0.41830      0.28128      1.49      0.14395
## Year2006      0.39884      0.17853      2.23      0.03050 *
## Year2007      0.90082      0.19074      4.72      2.3e-05 ***
## Year2008      0.32640      0.16000      2.04      0.04725 *
## Year2009      0.44164      0.15757      2.80      0.00745 **
## Year2010      0.68519      0.12178      5.63      1.1e-06 ***
## Year2011      0.16958      0.45914      0.37      0.71360
## Year2012      0.45637      0.13631      3.35      0.00165 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.319
## Multiple R-squared:  0.405, Adjusted R-squared:  0.114
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 63 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.512  0.915   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          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 7.051 1      2.655
## LastAuthorFemale  2.507 1      1.583
## Year             14.464 16      1.087

```

Residuals from first and last author



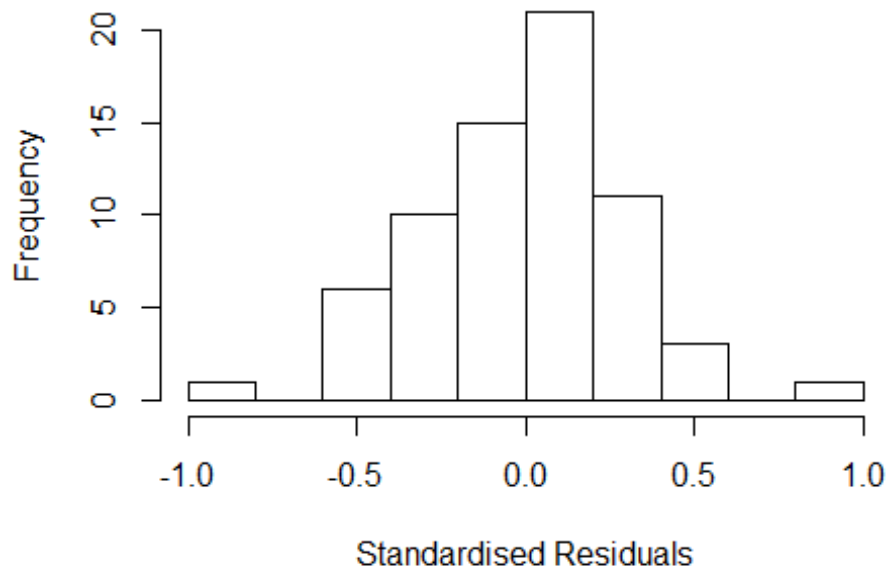
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.905 -0.199 0.015 0.199 0.849
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.733967 0.084057 8.73 1.5e-11 ***
## FirstAuthorFemale1 0.000251 0.082669 0.00 0.99759
## LastAuthorFemale1 -0.016509 0.096940 -0.17 0.86548
## Year1997 0.124734 0.439597 0.28 0.77780
## Year1998 0.517763 0.098169 5.27 3.0e-06 ***
## Year1999 0.229782 0.177922 1.29 0.20260
## Year2000 0.288305 0.237575 1.21 0.23074
## Year2001 0.194522 0.111604 1.74 0.08761 .
## Year2002 -0.281967 0.084057 -3.35 0.00154 **
## Year2003 0.424045 0.166650 2.54 0.01414 *
## Year2004 0.741033 0.173188 4.28 8.7e-05 ***
## Year2005 0.411908 0.259136 1.59 0.11837
```

```

## Year2006          0.421966    0.177435    2.38  0.02134 *
## Year2007          0.854436    0.181494    4.71  2.1e-05 ***
## Year2008          0.366690    0.129977    2.82  0.00689 **
## Year2009          0.471903    0.126638    3.73  0.00050 ***
## Year2010          0.693922    0.135990    5.10  5.4e-06 ***
## Year2011          0.434467    0.411547    1.06  0.29629
## Year2012          0.479349    0.117593    4.08  0.00017 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.335
## Multiple R-squared:  0.377, Adjusted R-squared:  0.149
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 61 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.445  0.929  0.963  0.934  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] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 6.419 1      2.534
## Year              6.419 16      1.060

```

Residuals from first author



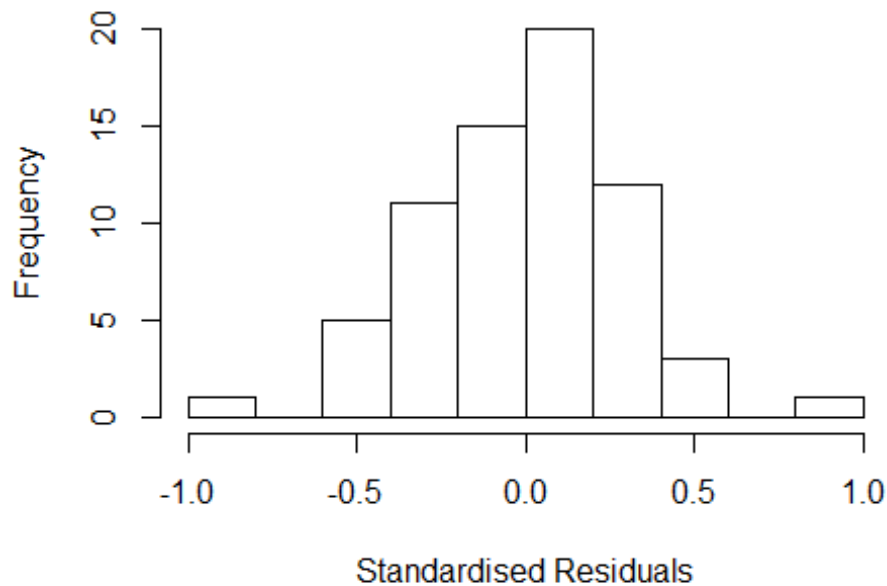
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.914 -0.199 0.016 0.198 0.857
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.73382 0.08406 8.73 1.3e-11 ***
## FirstAuthorFemale1 0.00112 0.08184 0.01 0.98917
## Year1997 0.11709 0.43937 0.27 0.79095
## Year1998 0.51188 0.09347 5.48 1.4e-06 ***
## Year1999 0.22906 0.17798 1.29 0.20401
## Year2000 0.28841 0.23865 1.21 0.23254
## Year2001 0.18780 0.10394 1.81 0.07680 .
## Year2002 -0.28182 0.08406 -3.35 0.00153 **
## Year2003 0.41952 0.15859 2.65 0.01088 *
## Year2004 0.74118 0.17347 4.27 8.7e-05 ***
## Year2005 0.41162 0.25971 1.58 0.11929
## Year2006 0.42197 0.17780 2.37 0.02152 *
```

```

## Year2007          0.84267    0.16607    5.07  5.7e-06 ***
## Year2008          0.36619    0.13001    2.82  0.00693 **
## Year2009          0.46618    0.12150    3.84  0.00035 ***
## Year2010          0.69029    0.13113    5.26  3.0e-06 ***
## Year2011          0.44287    0.40639    1.09  0.28104
## Year2012          0.47816    0.11845    4.04  0.00019 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.33
## Multiple R-squared:  0.38,   Adjusted R-squared:  0.169
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 60 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.424  0.921  0.962  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      1.47e-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 2.443 1          1.563
## Year            2.443 16          1.028

```


Residuals from last author



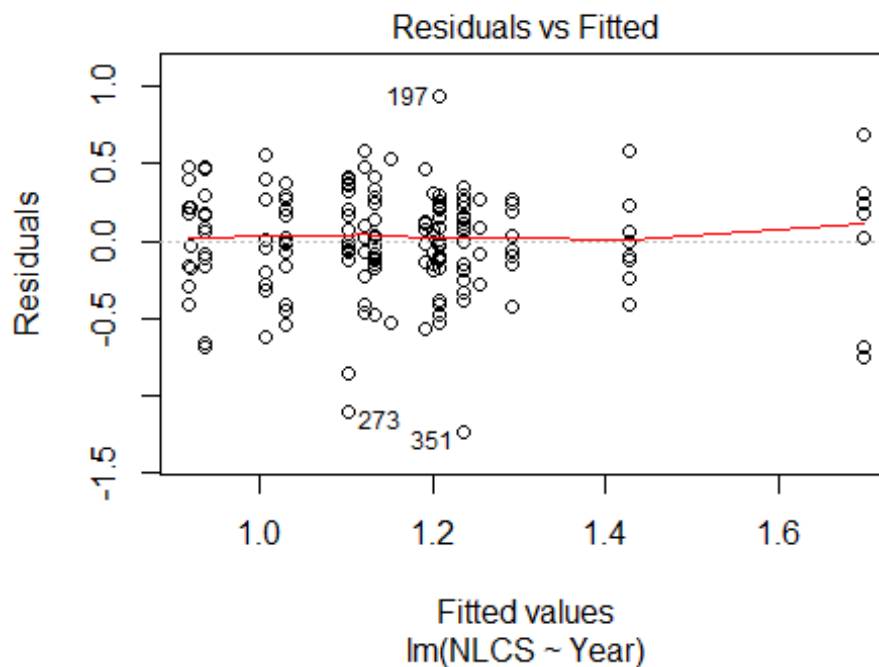
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.908 -0.199 0.015 0.199 0.851
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7341 0.0790 9.29 1.8e-12 ***
## LastAuthorFemale1 -0.0165 0.0957 -0.17 0.86405
## Year1997 0.1227 0.4265 0.29 0.77473
## Year1998 0.5178 0.0930 5.57 1.0e-06 ***
## Year1999 0.2299 0.1701 1.35 0.18254
## Year2000 0.2882 0.2361 1.22 0.22790
## Year2001 0.1945 0.1109 1.75 0.08565 .
## Year2002 -0.2821 0.0790 -3.57 0.00080 ***
## Year2003 0.4241 0.1660 2.55 0.01371 *
## Year2004 0.7409 0.1709 4.34 7.0e-05 ***
## Year2005 0.4119 0.2592 1.59 0.11829
## Year2006 0.4220 0.1772 2.38 0.02107 *
```

```

## Year2007          0.8546      0.1693      5.05  6.3e-06 ***
## Year2008          0.3667      0.1277      2.87  0.00597 **
## Year2009          0.4720      0.1203      3.92  0.00027 ***
## Year2010          0.6939      0.1380      5.03  6.7e-06 ***
## Year2011          0.4371      0.4055      1.08  0.28618
## Year2012          0.4795      0.1126      4.26  9.1e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.333
## Multiple R-squared:  0.378, Adjusted R-squared:  0.167
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 61 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.438  0.929  0.963  0.934  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 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
##    9   17    6   15   13   16    5    5   13   10   12   24   15   17   20
## 2011 2012
##   19   23
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7    8    4    8    6    8    3    2   10    8    8   20   12   16   16
## 2011 2012

```

```
## 16 19
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 6 7 2 5 5 5 2 0 6 4 7 18 12 15 15
## 2011 2012
## 16 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 = 16, df = 16, p-value = 0.5
```



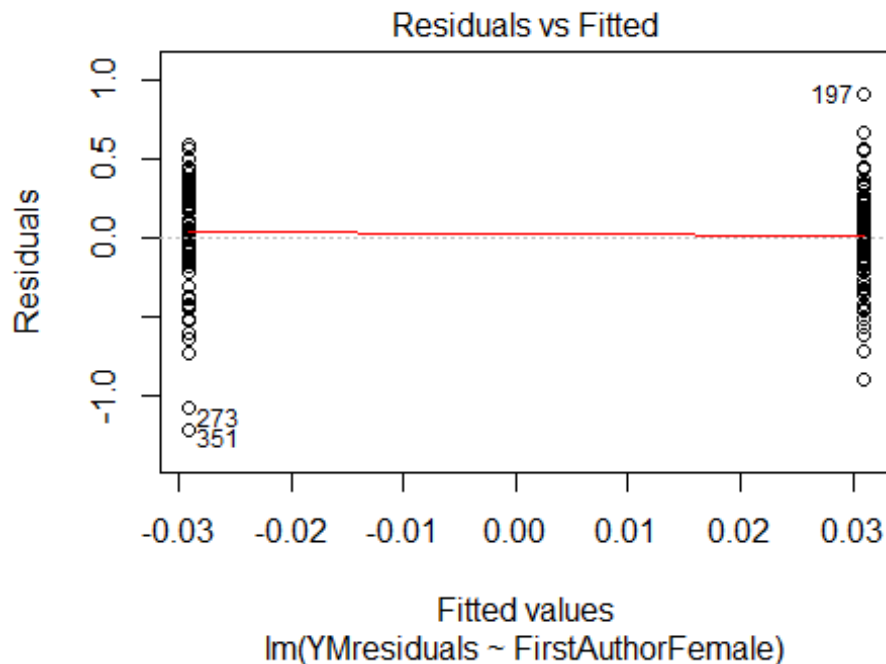
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.53, df = 1, p-value = 0.5
## [1] "Female first author team size 2018 geometric mean: 5.21000730958691"
## [1] "Male first author team size 2018 geometric mean: 5.83258067488062"
## Warning in wilcox.test.default(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 = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6.54389389941237"
## [1] "Male last author team size 2018 geometric mean: 4.73360578552818"

## Warning in wilcox.test.default(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] "Regression 1: First author gender, last author gender, team size,
Year as factors"

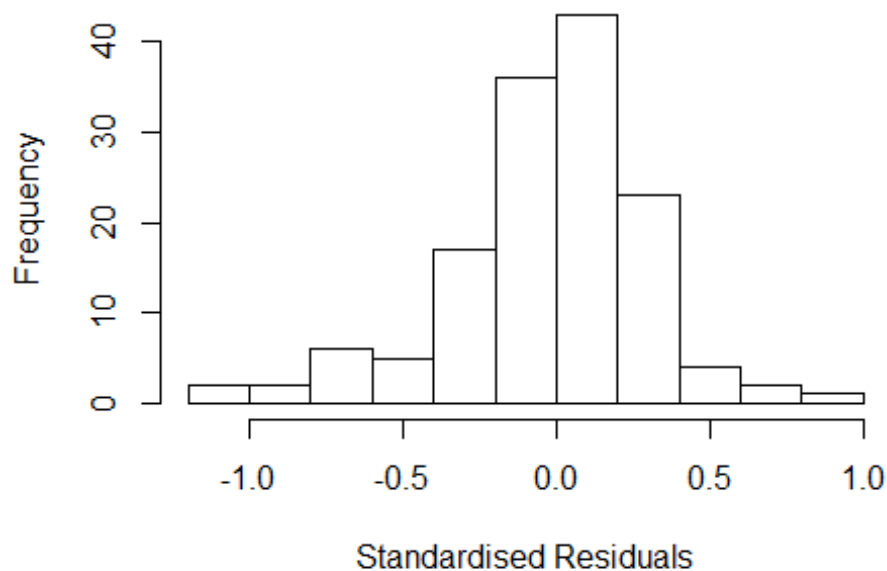
## Warning in lmrob.fit(x, y, control, init = init, mf = mf): M-step did NOT
## converge. Returning unconverged SM-estimate
```



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

| | GVIF | Df | GVIF ^{1/(2*Df)} |
|----------------------|--------|----|--------------------------|
| ## FirstAuthorFemale | 3.605 | 1 | 1.899 |
| ## LastAuthorFemale | 3.994 | 1 | 1.999 |
| ## Year | 14.076 | 15 | 1.092 |

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 |
|----|---------|---------|--------|--------|--------|
| ## | -1.1678 | -0.1791 | 0.0113 | 0.1882 | 0.9873 |

```
##
```

```
## Coefficients:
```

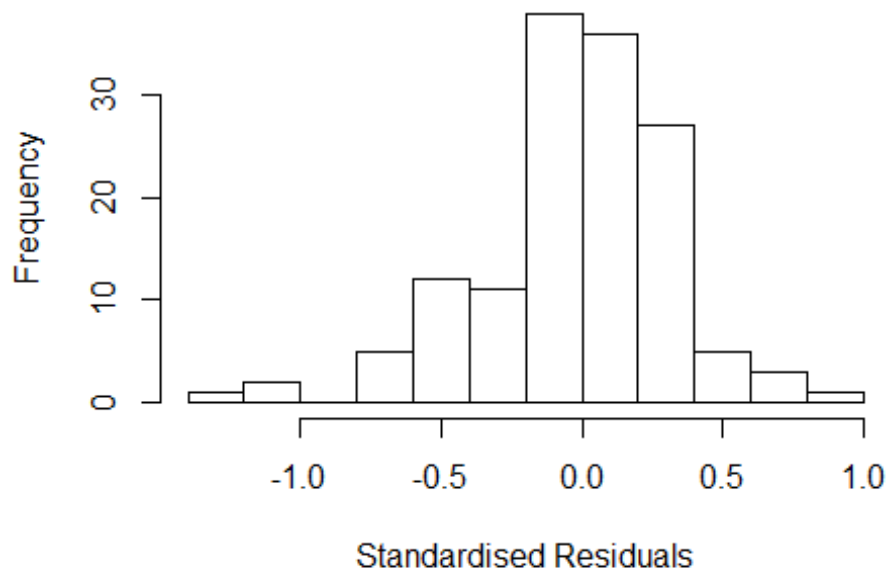
| | Estimate | Std. Error | t value | Pr(> t) |
|-----------------------|----------|------------|---------|-------------|
| ## (Intercept) | 1.7338 | 0.3790 | 4.57 | 1.1e-05 *** |
| ## FirstAuthorFemale1 | 0.0391 | 0.0639 | 0.61 | 0.541 |
| ## LastAuthorFemale1 | 0.1558 | 0.0845 | 1.84 | 0.068 . |
| ## Year1997 | -0.3801 | 0.3933 | -0.97 | 0.336 |
| ## Year1998 | -0.2958 | 0.3847 | -0.77 | 0.443 |

```

## Year1999          -0.4069      0.4015    -1.01      0.313
## Year2000          -0.5208      0.3902    -1.33      0.185
## Year2001          -1.0483      0.4022    -2.61      0.010 *
## Year2002          -0.9987      0.3963    -2.52      0.013 *
## Year2004          -0.7397      0.4470    -1.65      0.100
## Year2005          -0.6789      0.4649    -1.46      0.147
## Year2006          -0.5886      0.3965    -1.48      0.140
## Year2007          -0.6153      0.3970    -1.55      0.124
## Year2008          -0.8270      0.4062    -2.04      0.044 *
## Year2009          -0.6259      0.3896    -1.61      0.111
## Year2010          -0.5660      0.3937    -1.44      0.153
## Year2011          -0.7338      0.3894    -1.88      0.062 .
## Year2012          -0.5881      0.4064    -1.45      0.150
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.3
## Multiple R-squared:  0.303, Adjusted R-squared:  0.207
## Convergence in 27 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 125 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0959 0.8830 0.9600 0.8890 0.9860 0.9990
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          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 2.492 1          1.579
## Year              2.492 15          1.031

```

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
## -1.2199 -0.1810  0.0075  0.2052  0.9197
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.6918     0.3352   5.05 1.6e-06 ***
## FirstAuthorFemale1  0.0597     0.0676   0.88  0.3786
## Year1997         -0.3345     0.3542  -0.94  0.3468
## Year1998         -0.2538     0.3417  -0.74  0.4590
## Year1999         -0.3340     0.3553  -0.94  0.3490
## Year2000         -0.4870     0.3482  -1.40  0.1645
## Year2001        -1.0068     0.3667  -2.75  0.0069 **
## Year2002         -0.8788     0.3395  -2.59  0.0108 *
## Year2004         -0.6343     0.4104  -1.55  0.1247
## Year2005         -0.5329     0.3869  -1.38  0.1709
## Year2006         -0.5471     0.3578  -1.53  0.1288
## Year2007         -0.5262     0.3454  -1.52  0.1302
```

```

## Year2008          -0.7456      0.3522   -2.12    0.0363 *
## Year2009          -0.5747      0.3465   -1.66    0.0997 .
## Year2010          -0.5076      0.3474   -1.46    0.1465
## Year2011          -0.6775      0.3440   -1.97    0.0511 .
## Year2012          -0.4719      0.3462   -1.36    0.1753
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.3
## Multiple R-squared:  0.267, Adjusted R-squared:  0.172
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 124 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0614 0.8580 0.9540 0.8850 0.9840 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      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 4.072 1          2.018
## Year              4.072 15          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
## -1.18721 -0.16577  0.00664  0.18164  1.01147
##

```



```

## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.737      0.374   4.64 8.7e-06 ***
## LastAuthorFemale1  0.163      0.085   1.92  0.0569 .
## Year1997         -0.377      0.386  -0.98  0.3313
## Year1998         -0.299      0.380  -0.79  0.4332
## Year1999         -0.400      0.396  -1.01  0.3145
## Year2000         -0.508      0.383  -1.33  0.1868
## Year2001         -1.039      0.396  -2.63  0.0097 **
## Year2002         -1.006      0.391  -2.57  0.0113 *
## Year2004         -0.720      0.435  -1.65  0.1007
## Year2005         -0.671      0.460  -1.46  0.1474
## Year2006         -0.561      0.388  -1.44  0.1511
## Year2007         -0.603      0.389  -1.55  0.1231
## Year2008         -0.813      0.401  -2.03  0.0446 *
## Year2009         -0.611      0.383  -1.60  0.1129
## Year2010         -0.550      0.391  -1.41  0.1621
## Year2011         -0.719      0.384  -1.87  0.0632 .
## Year2012         -0.566      0.400  -1.42  0.1594
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.303
## Multiple R-squared:  0.296, Adjusted R-squared:  0.205
## Convergence in 32 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 125 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0914 0.8920 0.9570 0.8900 0.9840 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      7.09e-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 2805"
## [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
##    7    5    5    3    2    2    4    6    5    3    8    3    7    6    12
## 2012
##    8
##
## 1996 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    4    2    5    2    2    1    3    4    4    3    8    3    6    6    12
## 2012
##    7
##
## 1996 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    4    2    4    2    2    1    3    4    4    3    7    3    5    6    12
## 2012
##    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: 3.66284150148471"

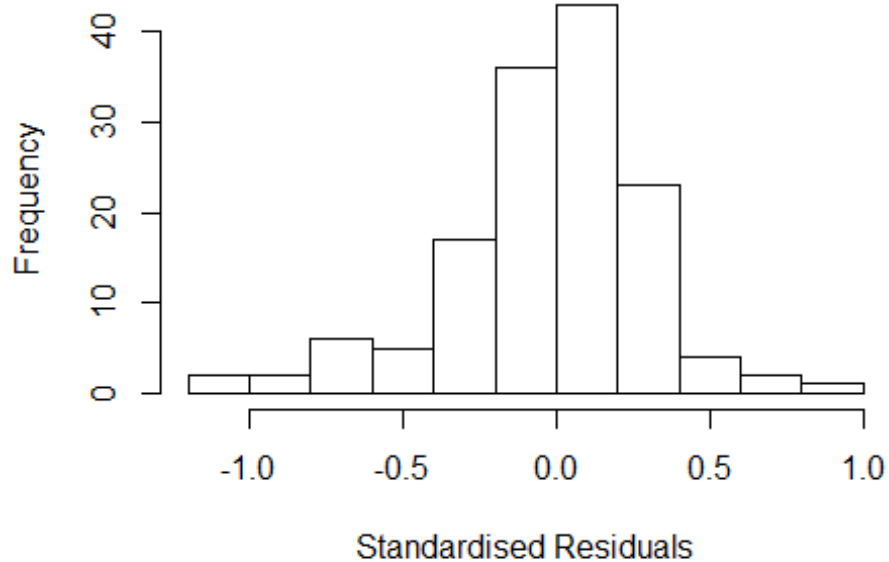
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6.06154651402987"
## [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

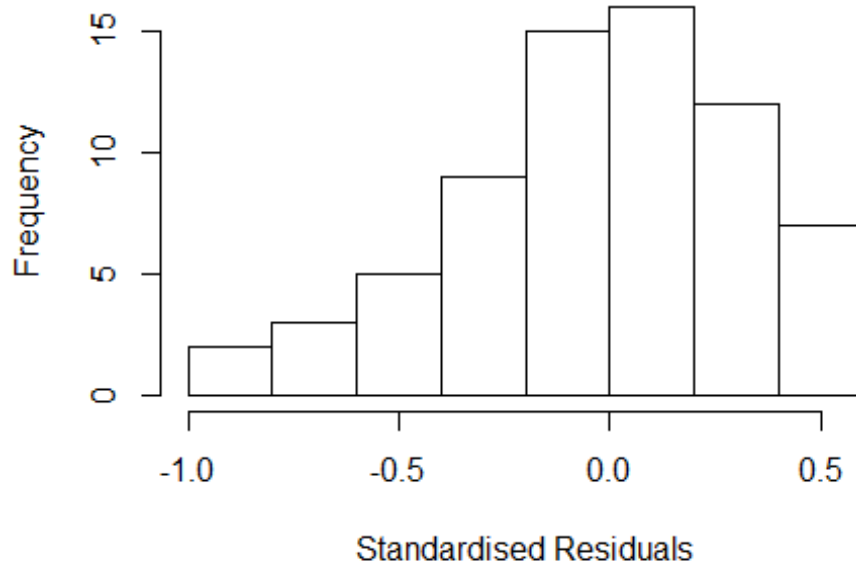
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 12, 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 11.73 1      3.425
## LastAuthorFemale  19.41 1      4.406
## UniqueAuthors    1142.03 4      2.411
## Year              4898.91 15     1.327
```

Residuals from first and last author and team size



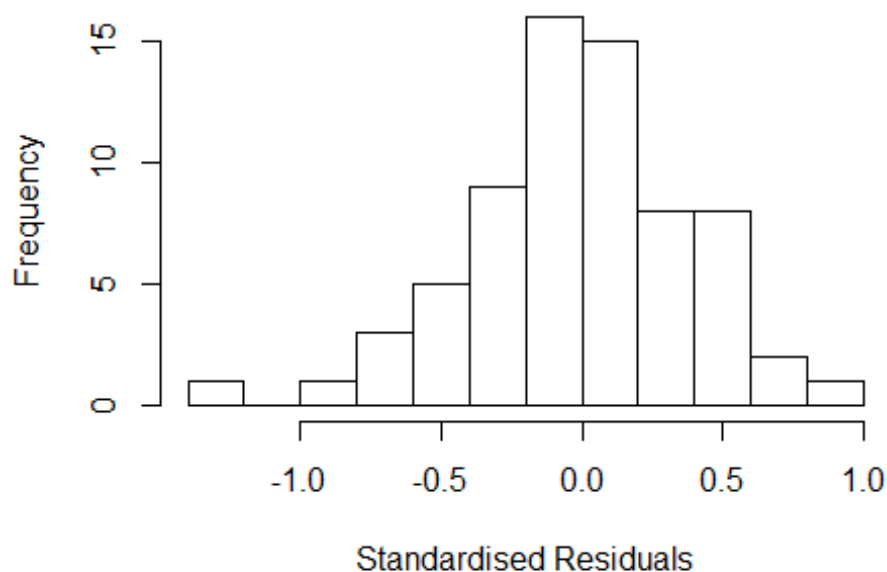
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.92333 -0.25259 0.00681 0.23496 0.55535
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8214 0.2253 3.65 0.00067 ***
## FirstAuthorFemale1 0.0831 0.1647 0.50 0.61622
## LastAuthorFemale1 -0.0494 0.2394 -0.21 0.83754
## UniqueAuthors2 0.2345 0.2503 0.94 0.35361
## UniqueAuthors3 0.2244 0.2274 0.99 0.32879
## UniqueAuthors4 0.7091 0.2515 2.82 0.00701 **
## UniqueAuthors5 0.4378 0.2072 2.11 0.03995 *
## Year1997 -0.3618 0.2173 -1.66 0.10261
## Year1999 0.1020 0.2982 0.34 0.73391
## Year2000 0.5542 0.2326 2.38 0.02129 *
```

```

## Year2001          -0.3041      0.6034   -0.50   0.61660
## Year2002          -0.1515      0.1590   -0.95   0.34551
## Year2003          -0.6240      0.2584   -2.41   0.01970 *
## Year2004           0.1813      0.1513    1.20   0.23701
## Year2005           0.0233      0.2321    0.10   0.92060
## Year2006           0.0318      0.1601    0.20   0.84341
## Year2007          -0.1367      0.2513   -0.54   0.58895
## Year2008          -0.0681      0.2437   -0.28   0.78103
## Year2009          -0.2965      0.3316   -0.89   0.37570
## Year2010          -0.1612      0.3376   -0.48   0.63516
## Year2011           0.0014      0.1689    0.01   0.99343
## Year2012           0.1508      0.1780    0.85   0.40120
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.385
## Multiple R-squared:  0.451, Adjusted R-squared:  0.206
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 59 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.545  0.901  0.949  0.920  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      1.45e-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 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 6.251 1 2.500
## LastAuthorFemale 5.597 1 2.366
## Year 11.929 15 1.086

```

Residuals from first and last author



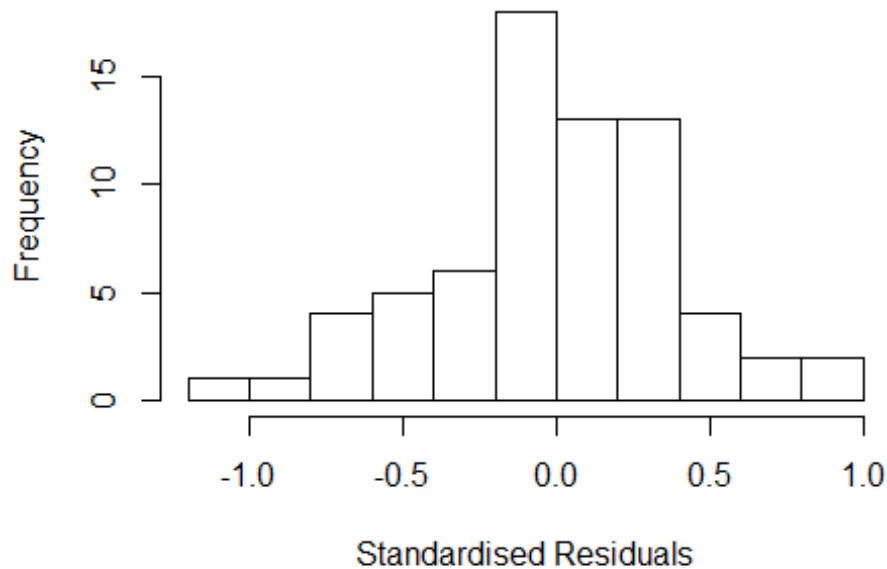
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.21e+00 -2.22e-01 -3.48e-16 2.28e-01 8.74e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3833 0.1474 9.38 1.1e-12 ***
## FirstAuthorFemale1 0.1564 0.1186 1.32 0.193
## LastAuthorFemale1 -0.1508 0.1566 -0.96 0.340
## Year1997 -0.7048 0.4069 -1.73 0.089 .
## Year1999 -0.1729 0.4606 -0.38 0.709
## Year2000 0.2217 0.2068 1.07 0.289
## Year2001 -0.7905 0.7653 -1.03 0.307
## Year2002 -0.0043 0.1474 -0.03 0.977
## Year2003 -0.9785 0.4545 -2.15 0.036 *
## Year2004 -0.0933 0.1524 -0.61 0.543
## Year2005 -0.2552 0.2807 -0.91 0.368
## Year2006 -0.0225 0.1603 -0.14 0.889
```

```

## Year2007          -0.4236      0.2518   -1.68    0.099 .
## Year2008          -0.4311      0.2350   -1.83    0.072 .
## Year2009          -0.6360      0.3435   -1.85    0.070 .
## Year2010          -0.3343      0.2039   -1.64    0.107
## Year2011          -0.2053      0.1642   -1.25    0.217
## Year2012          -0.1297      0.1861   -0.70    0.489
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.401
## Multiple R-squared:  0.337, Adjusted R-squared:  0.116
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 62 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.343  0.901  0.957  0.912  0.985  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      1.45e-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 4.183 1          2.045
## Year              4.183 15          1.049

```

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
## -1.09e+00 -1.96e-01 -1.31e-16  2.34e-01  8.65e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.37562    0.17258   7.97 1.4e-10 ***
## FirstAuthorFemale1 0.10663    0.10731   0.99  0.325
## Year1997      -0.69712    0.41395  -1.68  0.098 .
## Year1999      -0.28449    0.41920  -0.68  0.500
## Year2000       0.22938    0.22513   1.02  0.313
## Year2001      -0.75793    0.81505  -0.93  0.357
## Year2002       0.00338    0.17258   0.02  0.984
## Year2003      -0.96185    0.46316  -2.08  0.043 *
## Year2004      -0.11072    0.17783  -0.62  0.536
## Year2005      -0.27784    0.26981  -1.03  0.308
## Year2006       0.00156    0.19022   0.01  0.993
## Year2007      -0.45940    0.24763  -1.86  0.069 .
```

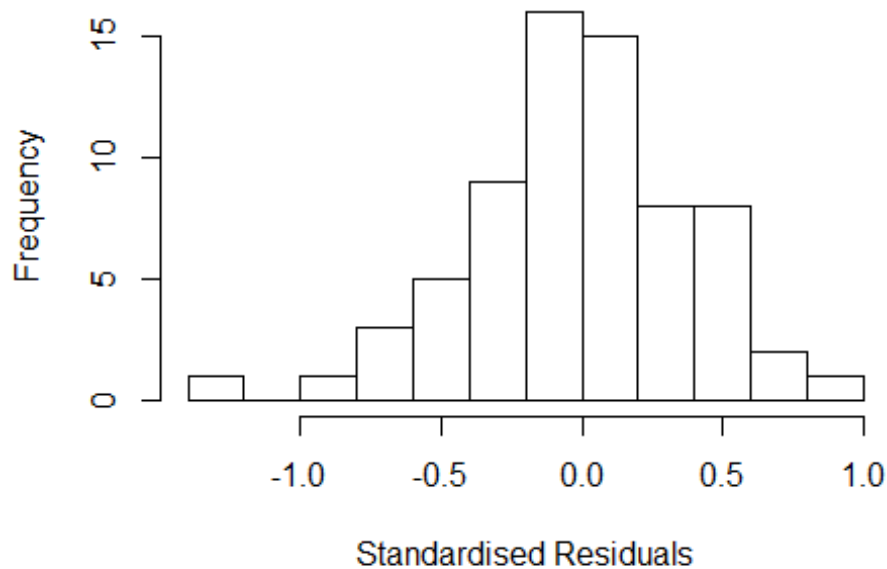


```

## Year2008          -0.40390    0.24757   -1.63    0.109
## Year2009          -0.63064    0.34730   -1.82    0.075 .
## Year2010          -0.33186    0.22802   -1.46    0.152
## Year2011          -0.19745    0.19046   -1.04    0.305
## Year2012          -0.13704    0.21195   -0.65    0.521
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.414
## Multiple R-squared:  0.319, Adjusted R-squared:  0.11
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 64 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.466  0.905  0.970  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      1.45e-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 4.909 1          2.216
## Year            4.909 15          1.054

```

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
## -1.20e+00 -2.29e-01 3.47e-16 2.54e-01 8.65e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4405 0.1472 9.79 2.2e-13 ***
## LastAuthorFemale1 -0.0674 0.1356 -0.50 0.621
## Year1997 -0.7620 0.4033 -1.89 0.064 .
## Year1999 -0.2449 0.4043 -0.61 0.547
## Year2000 0.1645 0.2063 0.80 0.429
## Year2001 -0.7695 1.0497 -0.73 0.467
## Year2002 -0.0615 0.1472 -0.42 0.678
## Year2003 -1.0263 0.4474 -2.29 0.026 *
## Year2004 -0.1322 0.1516 -0.87 0.387
## Year2005 -0.2747 0.2488 -1.10 0.275
## Year2006 -0.0289 0.1873 -0.15 0.878
## Year2007 -0.4663 0.2575 -1.81 0.076 .
```

```

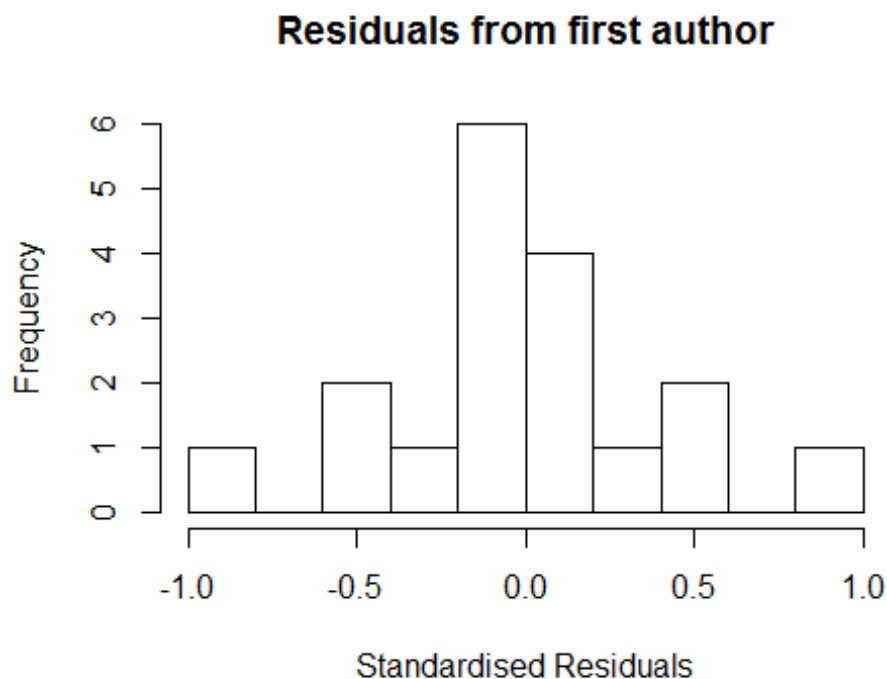
## Year2008          -0.4280      0.2315   -1.85    0.070 .
## Year2009          -0.6957      0.3363   -2.07    0.044 *
## Year2010          -0.3048      0.2100   -1.45    0.153
## Year2011          -0.1940      0.1713   -1.13    0.262
## Year2012          -0.1104      0.1835   -0.60    0.550
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.414
## Multiple R-squared:  0.315, Adjusted R-squared:  0.105
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 59 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.385  0.907  0.959  0.911  0.982  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.45e-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: 69"
## [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 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    1    1    2    1    1    3    2    6    1    1    4    1
##
## 1997 1998 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    0    1    0    2    1    0    3    2    4    1    1    4    1
##
## 1997 1998 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    0    0    0    0    2    0    0    3    2    4    1    1    4    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"
##
      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -223 1      NaN
## Year              -223 7      NaN

```



```

## [1] "Regression 4: Last author gender, Year as factors"
##
      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 3.956e+16 1      1.989e+08
## Year              3.956e+16 7      1.533e+01

## [1] "Sample size for the above analysis: 18"
## [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]"
##
## 2000 2001 2002 2004 2005 2007 2008 2009 2010 2011 2012

```

```

##      2      3      1      2      1      1      3      3      8      5     12
##
## 2000 2001 2002 2004 2005 2007 2008 2009 2010 2011 2012
##      1      0      0      2      1      1      2      3      6      3      8
##
## 2000 2001 2002 2004 2005 2007 2008 2009 2010 2011 2012
##      1      0      0      2      0      1      2      3      6      2      8
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 8"
## [1] "Male first author team size 2018 geometric mean: 5.59344471040698"

## Warning in wilcox.test.default(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: 7.48331477354788"
## [1] "Male last author team size 2018 geometric mean: 5"

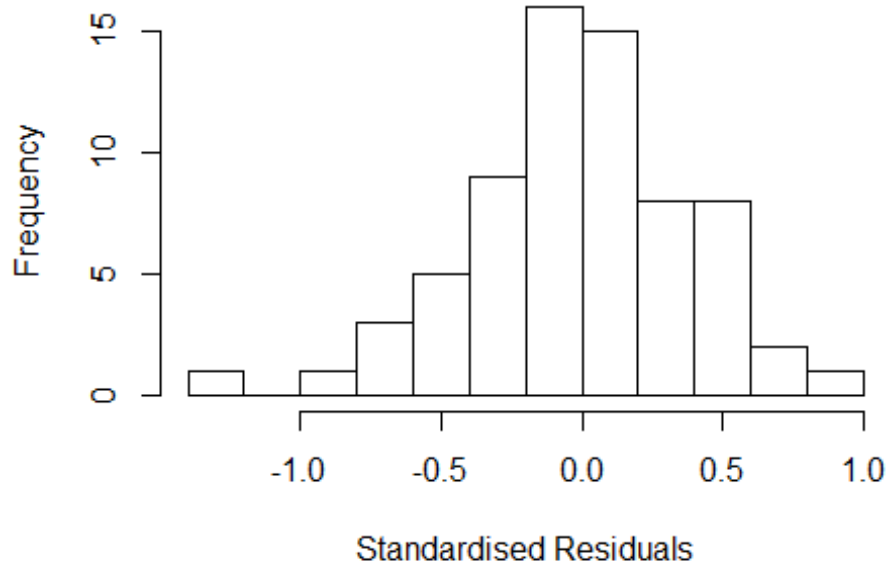
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4, 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"

## 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 -3.420e+27  1      NaN
## LastAuthorFemale  1.450e+01  1      3.807
## Year              -4.541e+28  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
## -0.963 -0.199 -0.015  0.258  0.584
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.401      0.000    Inf <2e-16 ***
## FirstAuthorFemale1  0.157      0.463    0.34  0.7391
## LastAuthorFemale1 -0.233      0.131   -1.78  0.0952 .
## Year2004          -0.904      0.491   -1.84  0.0854 .
## Year2007          -0.253      0.463   -0.55  0.5926
## Year2008          -1.097      0.314   -3.50  0.0032 **
## Year2009          -0.199      0.396   -0.50  0.6221
```

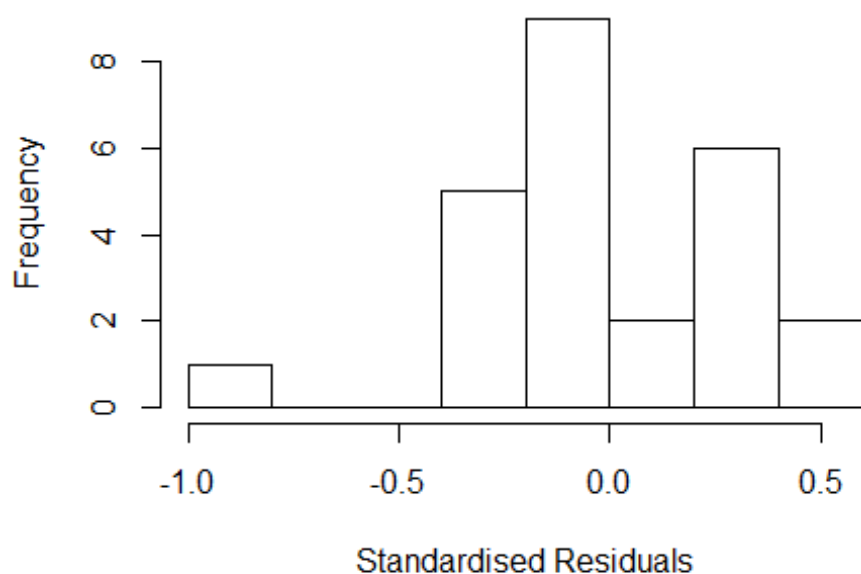
```

## Year2010          -0.019      0.290   -0.07   0.9485
## Year2011          -1.024      0.324   -3.16   0.0065 **
## Year2012          -0.438      0.673   -0.65   0.5248
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.39
## Multiple R-squared:  0.595, Adjusted R-squared:  0.352
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 20 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.522  0.916  0.937  0.922  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      4.00e-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

```

Residuals from first and last author



```
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.501e+15  1      3.875e+07
## Year              1.501e+15  7      1.213e+01

## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.87712 -0.28963 -0.00583  0.29908  0.65637
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.401      0.000    Inf <2e-16 ***
## FirstAuthorFemale1  0.202      0.409    0.49  0.6284
## Year2004         -1.066      0.422   -2.52  0.0225 *
## Year2007         -0.298      0.409   -0.73  0.4771
## Year2008         -1.236      0.374   -3.31  0.0045 **
## Year2009         -0.300      0.311   -0.97  0.3486
## Year2010         -0.136      0.209   -0.65  0.5231
## Year2011        -1.024      0.326   -3.14  0.0064 **
```



```

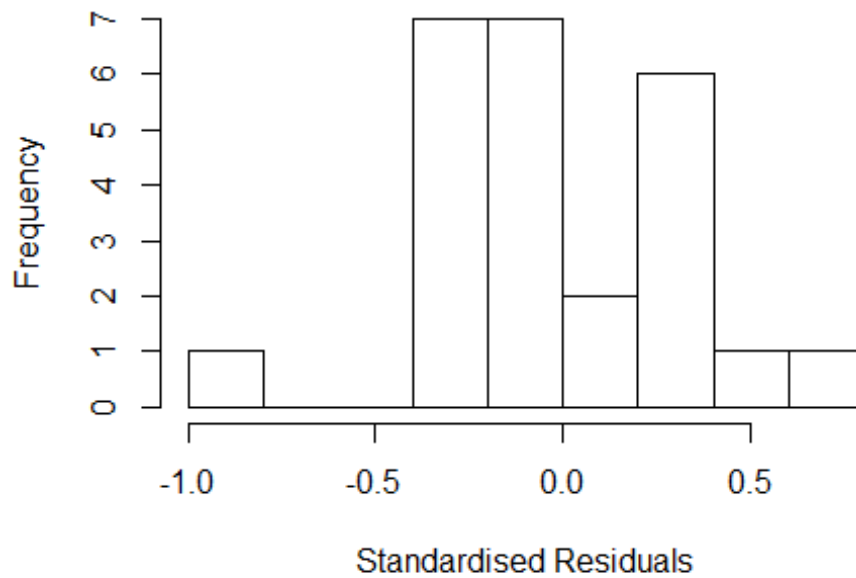
## Year2012          -0.524      0.592   -0.88   0.3896
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.385
## Multiple R-squared:  0.554, Adjusted R-squared:  0.331
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 3 weights are ~= 1. The remaining 22 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.582  0.916  0.946  0.922  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      4.00e-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 cov2cor(v): diag(.) had 0 or NA entries; non-finite result is
## doubtful

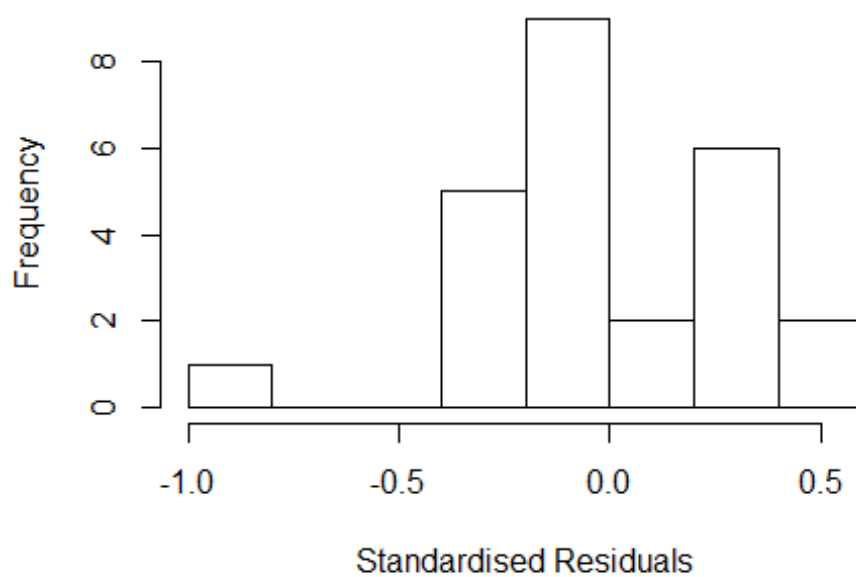
```

Residuals from first author



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

Residuals from last author



```

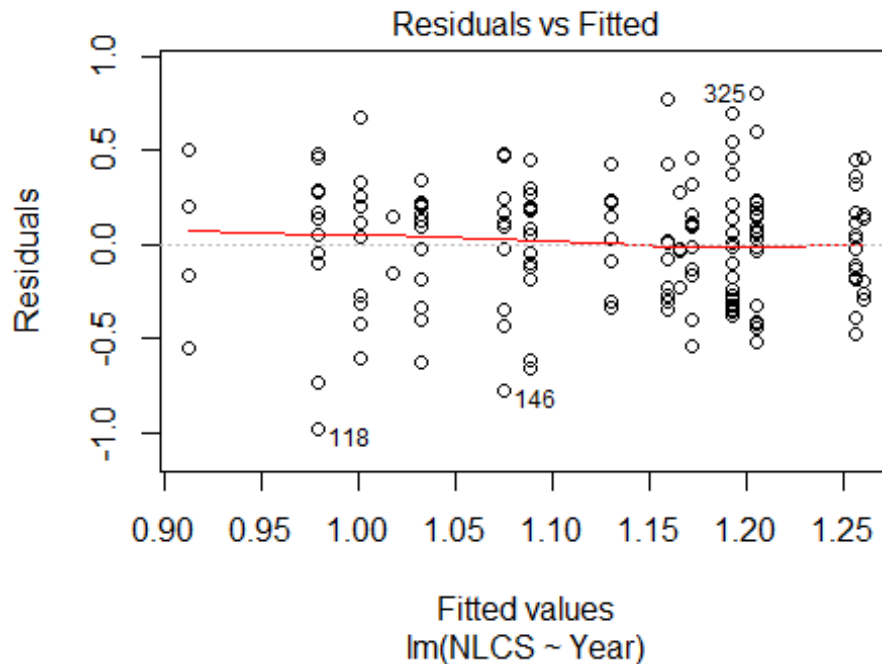
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.14e+00 -1.58e-01 -1.55e-15 2.34e-01 6.37e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4010 0.0000 Inf < 2e-16 ***
## LastAuthorFemale1 -0.2530 0.1068 -2.37 0.0308 *
## Year2004 -0.7375 0.0548 -13.45 3.9e-10 ***
## Year2007 -0.0960 0.0000 -Inf < 2e-16 ***
## Year2008 -1.0090 0.1141 -8.84 1.5e-07 ***
## Year2009 -0.0829 0.1148 -0.72 0.4805
## Year2010 0.0846 0.1559 0.54 0.5950
## Year2011 -1.0240 0.3304 -3.10 0.0069 **
## Year2012 -0.2565 0.2493 -1.03 0.3188
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.375
## Multiple R-squared: 0.614, Adjusted R-squared: 0.421
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 18 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.331 0.910 0.943 0.902 0.980 0.992
## Algorithmic parameters:
## tuning.chi bb tuning.psi refine.tol
## 1.55e+00 5.00e-01 4.69e+00 1.00e-07
## rel.tol solve.tol eps.outlier eps.x
## 1.00e-07 1.00e-07 4.00e-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: 25"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2808"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      3      8      2     12     13     10      7     12     13     16     14     19     20     19     25
## 2011 2012
##     31     27
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      0      2      0      9      6      4      4      9     12     11     10     14     13     11     14
## 2011 2012
##     18     19
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      0      2      0      6      6      3      1      8      9     11      9     10      8      9     12
## 2011 2012
##     17     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 = 7.6, df = 14, p-value = 0.9

```



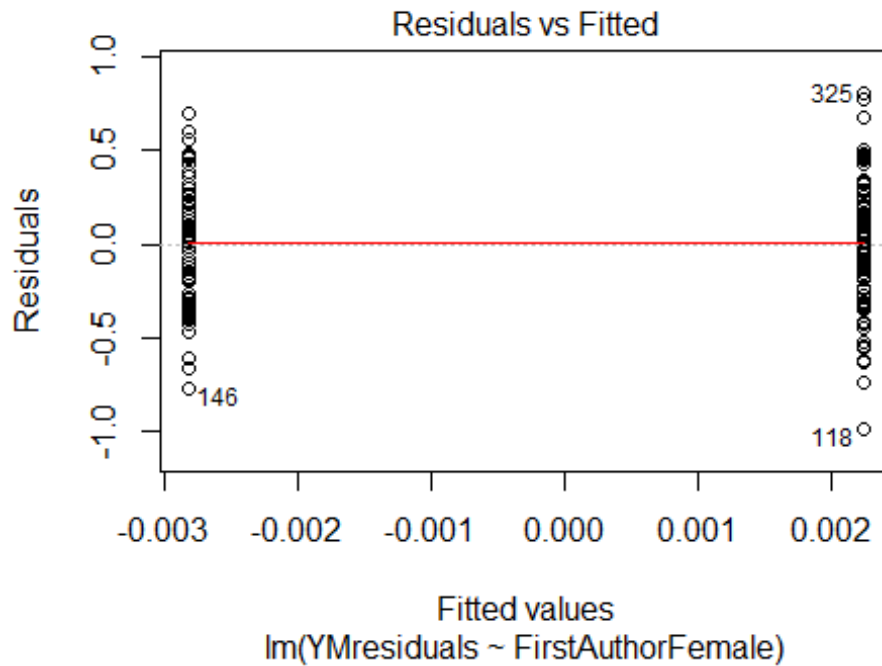
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.054, df = 1, p-value = 0.8

## [1] "Female first author team size 2018 geometric mean: 5.77558079474419"
## [1] "Male first author team size 2018 geometric mean: 4.1872508178068"

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

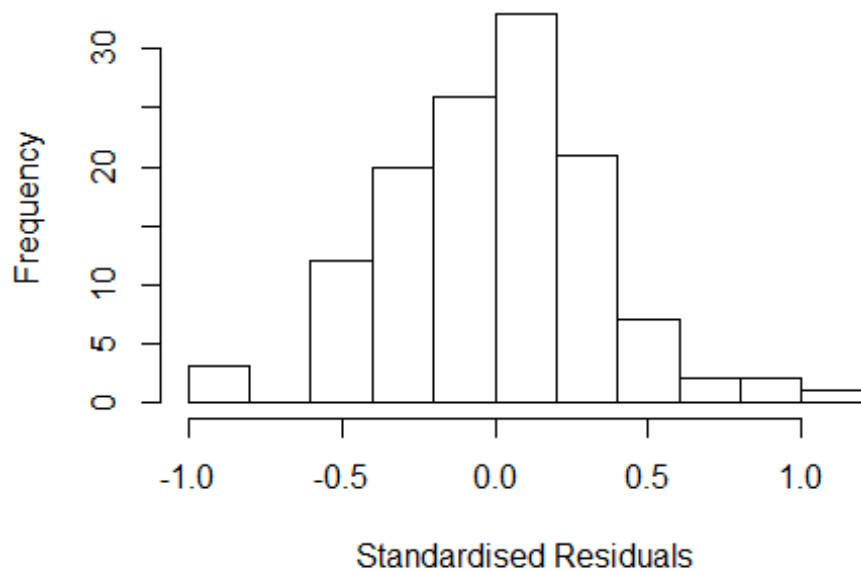
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 78, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 7.0935078069034"
## [1] "Male last author team size 2018 geometric mean: 4.04558425266621"

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



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 98, 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  6.984  1      2.643
## LastAuthorFemale  6.794  1      2.606
## UniqueAuthors    126.550  4      1.831
## Year              375.390 14      1.236
```

Residuals from first and last author and team size



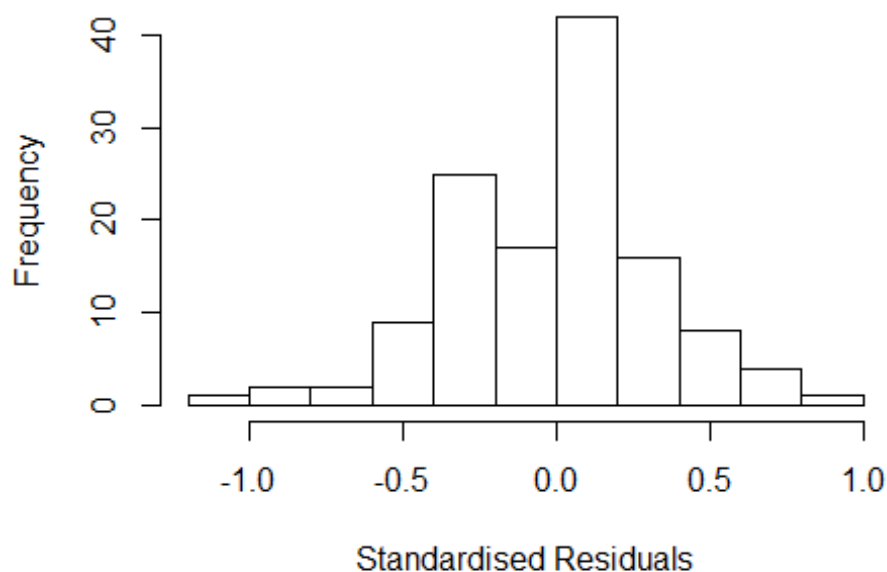
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.97342 -0.20486 0.00258 0.20786 1.00240
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.94465 0.10104 9.35 1.6e-15 ***
## FirstAuthorFemale1 0.00533 0.06672 0.08 0.936
## LastAuthorFemale1 -0.06860 0.06754 -1.02 0.312
## UniqueAuthors2 0.33473 0.16853 1.99 0.050 *
## UniqueAuthors3 0.27832 0.16799 1.66 0.101
## UniqueAuthors4 0.14670 0.16785 0.87 0.384
## UniqueAuthors5 0.34532 0.16063 2.15 0.034 *
## Year1999 -0.01605 0.17165 -0.09 0.926
## Year2000 0.16966 0.14452 1.17 0.243
## Year2001 0.02684 0.10277 0.26 0.794
```

```

## Year2002          0.09091    0.12953    0.70    0.484
## Year2003         -0.09880    0.15274   -0.65    0.519
## Year2004         -0.32131    0.17892   -1.80    0.075 .
## Year2005         -0.13222    0.20106   -0.66    0.512
## Year2006         -0.02088    0.14390   -0.15    0.885
## Year2007          0.00758    0.14589    0.05    0.959
## Year2008         -0.14905    0.16129   -0.92    0.358
## Year2009         -0.04420    0.16752   -0.26    0.792
## Year2010         -0.10323    0.14405   -0.72    0.475
## Year2011         -0.04555    0.12712   -0.36    0.721
## Year2012         -0.09353    0.13568   -0.69    0.492
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.315
## Multiple R-squared:  0.157, Adjusted R-squared:  -0.00235
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 115 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.290  0.873  0.954  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      7.87e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 4.044 1      2.011
## LastAuthorFemale  5.155 1      2.270
## Year              9.895 14      1.085

```


Residuals from first and last author



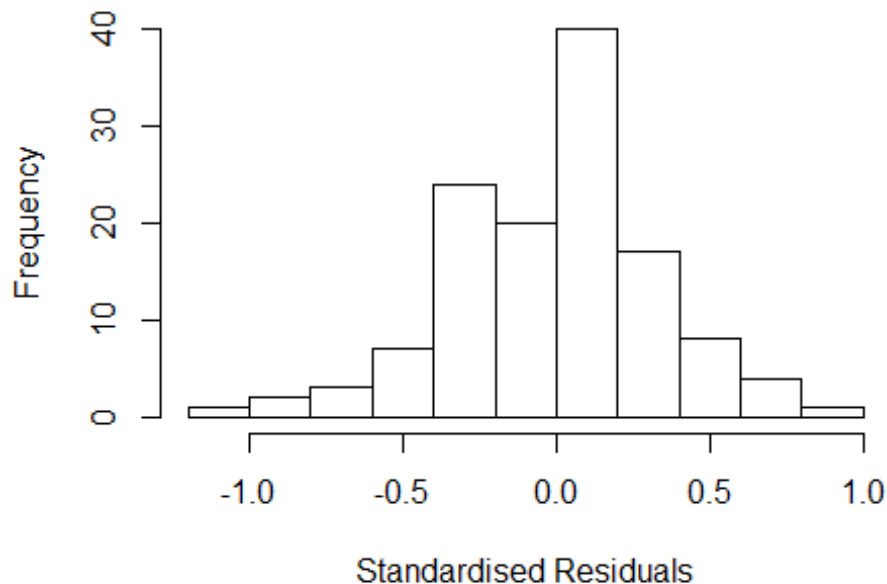
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1033 -0.2296  0.0363  0.1587  0.8583
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.01800    0.11197   9.09 4.7e-15 ***
## FirstAuthorFemale1  0.00199    0.06702   0.03  0.976
## LastAuthorFemale1 -0.06279    0.07302  -0.86  0.392
## Year1999         0.12633    0.18063   0.70  0.486
## Year2000         0.24362    0.16728   1.46  0.148
## Year2001         0.17885    0.17315   1.03  0.304
## Year2002         0.16179    0.14757   1.10  0.275
## Year2003         0.09835    0.15457   0.64  0.526
## Year2004        -0.09781    0.18041  -0.54  0.589
## Year2005         0.08534    0.16164   0.53  0.599
## Year2006         0.11558    0.19612   0.59  0.557
## Year2007         0.26966    0.14052   1.92  0.058 .
```

```

## Year2008          0.08893    0.16002    0.56    0.580
## Year2009          0.16804    0.15592    1.08    0.284
## Year2010          0.11710    0.14774    0.79    0.430
## Year2011          0.13236    0.13982    0.95    0.346
## Year2012          0.13272    0.15612    0.85    0.397
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.326
## Multiple R-squared:  0.0668, Adjusted R-squared:  -0.069
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 117 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.228  0.874  0.955  0.904  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      7.87e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.897 1          1.974
## Year              3.897 14          1.050

```

Residuals from first author



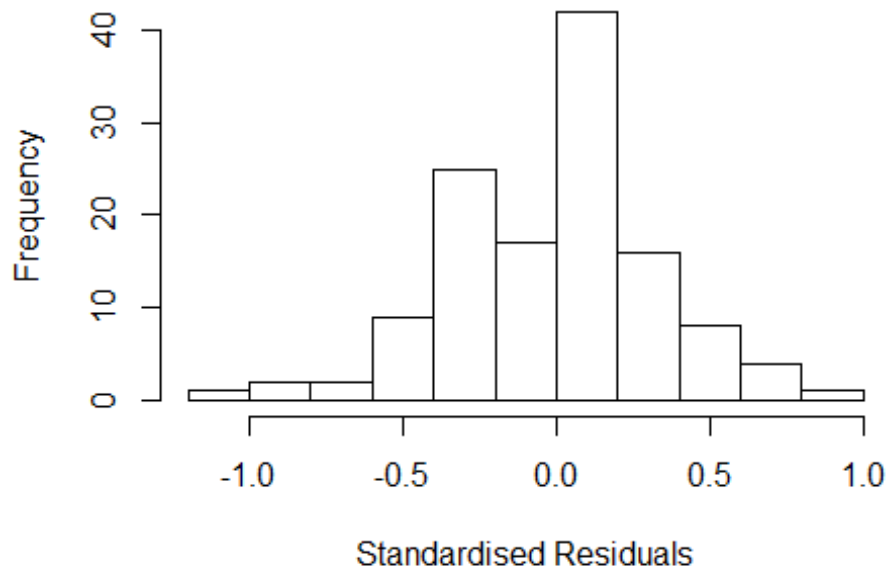
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0861 -0.2275 0.0395 0.1759 0.8680
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.01800 0.11195 9.09 4.4e-15 ***
## FirstAuthorFemale1 -0.00792 0.06731 -0.12 0.907
## Year1999 0.11888 0.17712 0.67 0.504
## Year2000 0.23782 0.16902 1.41 0.162
## Year2001 0.15543 0.16949 0.92 0.361
## Year2002 0.10892 0.13062 0.83 0.406
## Year2003 0.06413 0.14838 0.43 0.666
## Year2004 -0.11781 0.18587 -0.63 0.527
## Year2005 0.06809 0.16107 0.42 0.673
## Year2006 0.11061 0.19416 0.57 0.570
## Year2007 0.25541 0.13840 1.85 0.068 .
## Year2008 0.08714 0.16338 0.53 0.595
```

```

## Year2009          0.16653    0.15383    1.08    0.281
## Year2010          0.10994    0.15006    0.73    0.465
## Year2011          0.12266    0.13972    0.88    0.382
## Year2012          0.12303    0.15232    0.81    0.421
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.326
## Multiple R-squared:  0.0612, Adjusted R-squared:  -0.0657
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 119 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.246  0.886  0.958  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      7.87e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 5.204 1          2.281
## Year              5.204 14          1.061

```

Residuals from last author



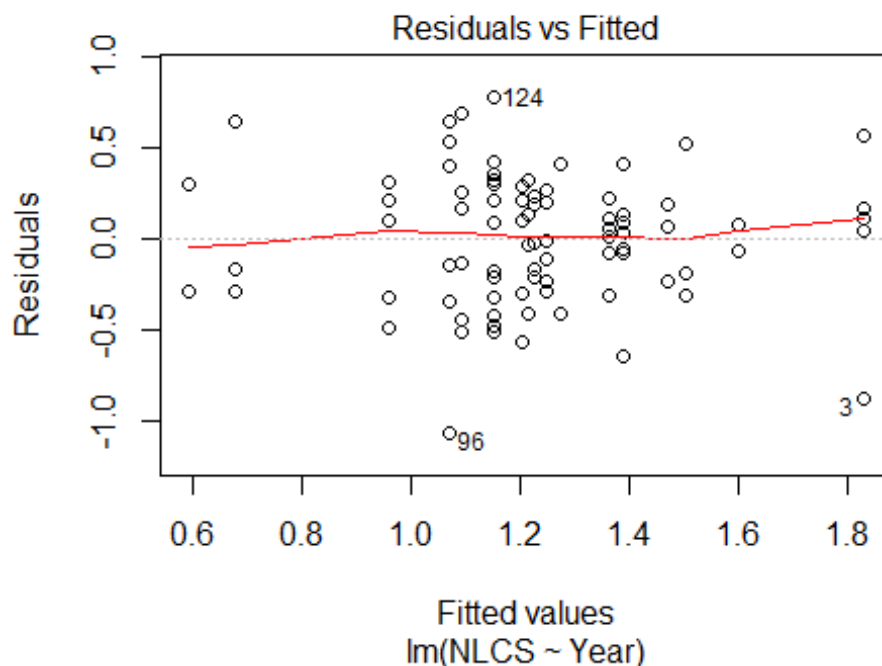
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1068 -0.2292 0.0382 0.1559 0.8582
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0180 0.1121 9.08 4.6e-15 ***
## LastAuthorFemale1 -0.0626 0.0737 -0.85 0.397
## Year1999 0.1258 0.1714 0.73 0.464
## Year2000 0.2444 0.1630 1.50 0.137
## Year2001 0.1788 0.1734 1.03 0.305
## Year2002 0.1636 0.1341 1.22 0.225
## Year2003 0.0993 0.1516 0.65 0.514
## Year2004 -0.0985 0.1804 -0.55 0.586
## Year2005 0.0888 0.1611 0.55 0.583
## Year2006 0.1185 0.1944 0.61 0.543
## Year2007 0.2707 0.1391 1.95 0.054 .
## Year2008 0.0909 0.1573 0.58 0.565
```

```

## Year2009          0.1688      0.1576      1.07      0.286
## Year2010          0.1188      0.1474      0.81      0.422
## Year2011          0.1327      0.1383      0.96      0.340
## Year2012          0.1328      0.1523      0.87      0.385
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.321
## Multiple R-squared:  0.0674, Adjusted R-squared:  -0.0586
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 117 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.211  0.870  0.954  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      7.87e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 127"
## [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
##   10   12   8    6    8    9    4    3    3   10   10   15   12   12   11
## 2011 2012
##   10    7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    3    4    3    7    4    2    2    2    6    6   14    5    6    8
## 2011 2012
##    6    6
##

```

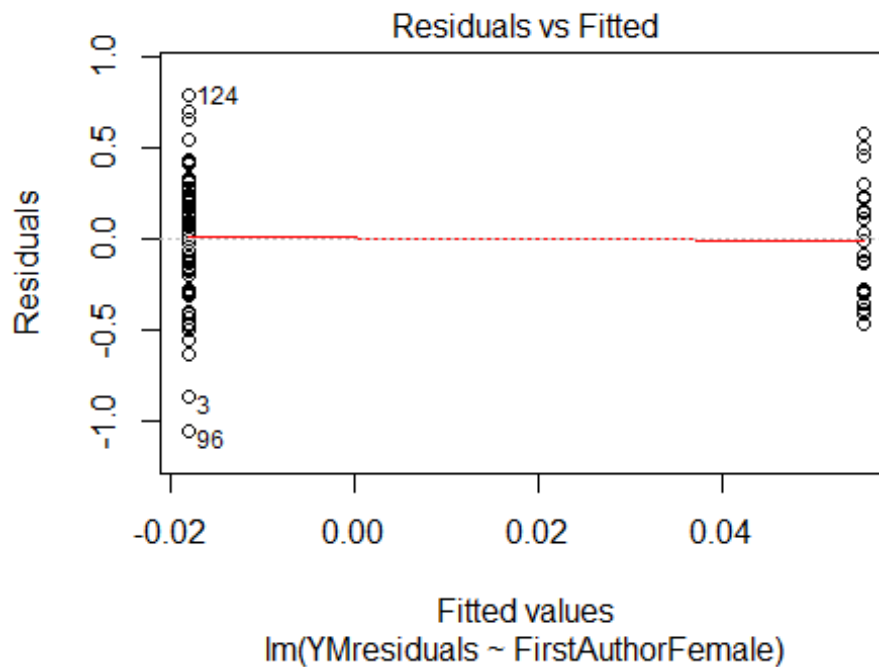
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      5      3      3      2      6      4      2      1      2      5      5     14      4      6      8
## 2011 2012
##      6      6
## [1] "Heteroscedasticity checks, confirming that there are problems with
##      these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 16, df = 16, p-value = 0.4
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.77, df = 1, p-value = 0.4
##
## [1] "Female first author team size 2018 geometric mean: 3.89805984091619"
## [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 = 3.5, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.48140474655716"
## [1] "Male last author team size 2018 geometric mean: 3.10723250595386"

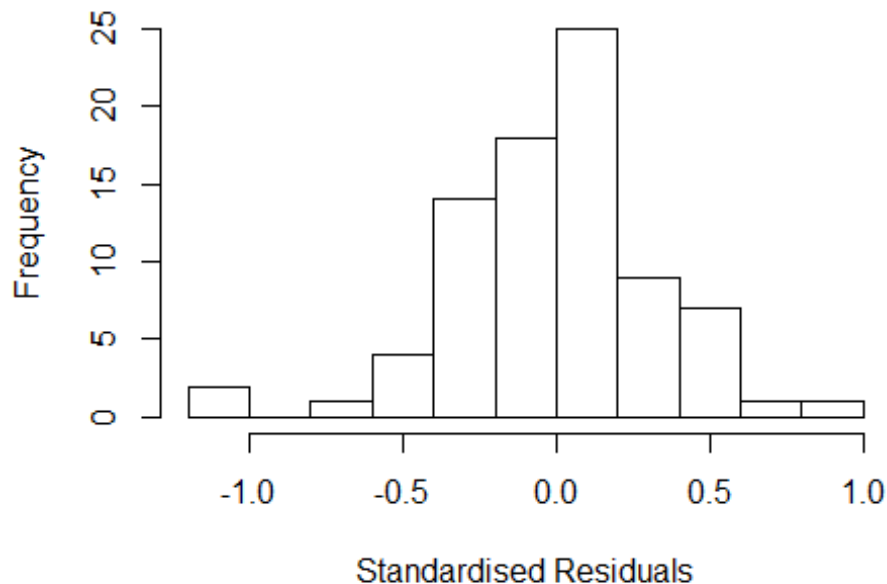
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7, p-value = 0.4
## 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 | 154.44 | 1 | 12.428 |
| LastAuthorFemale | 50.45 | 1 | 7.103 |
| UniqueAuthors | 36461.01 | 4 | 3.717 |
| Year | 319979.41 | 16 | 1.486 |

Residuals from first and last author and team size



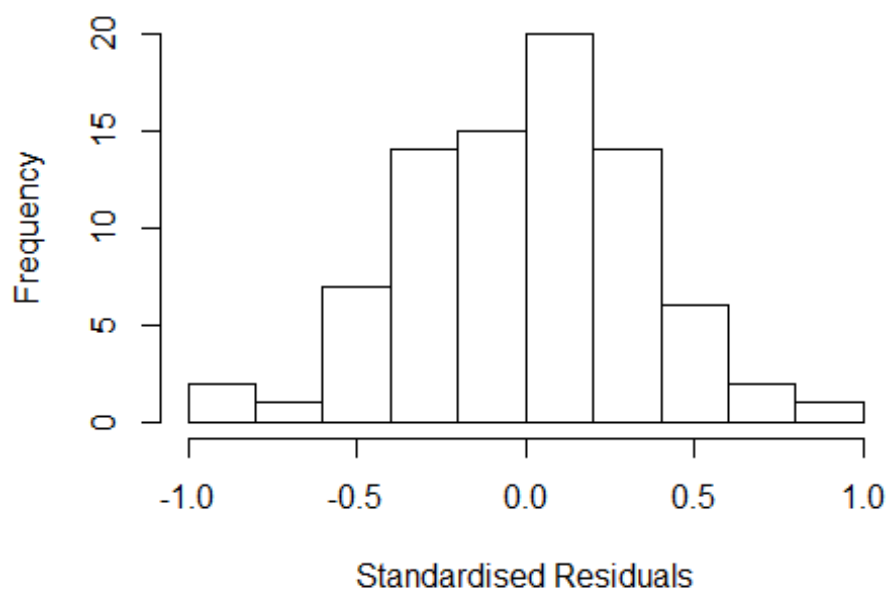
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1051 -0.2058  0.0151  0.1785  0.9175
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.6771     0.2989   5.61 5.6e-07 ***
## FirstAuthorFemale1  0.1795     0.1387   1.29  0.20052
## LastAuthorFemale1  0.0542     0.1229   0.44  0.66069
## UniqueAuthors2     0.0636     0.2108   0.30  0.76392
## UniqueAuthors3     0.1170     0.1522   0.77  0.44534
## UniqueAuthors4     0.3750     0.2377   1.58  0.11998
## UniqueAuthors5     0.1636     0.1770   0.92  0.35898
## Year1997          -0.3427     0.4587  -0.75  0.45802
## Year1998          -0.6459     0.4272  -1.51  0.13593
## Year1999          -0.3835     0.4370  -0.88  0.38371
```

```

## Year2000          -0.6906      0.3946   -1.75   0.08534 .
## Year2001          -1.1669      0.3759   -3.10   0.00293 **
## Year2002          -1.1955      0.3763   -3.18   0.00237 **
## Year2003          -1.0478      0.2543   -4.12   0.00012 ***
## Year2004          -0.4138      0.4224   -0.98   0.33137
## Year2005          -0.8031      0.5170   -1.55   0.12566
## Year2006          -0.9781      0.3561   -2.75   0.00797 **
## Year2007          -0.8352      0.4350   -1.92   0.05971 .
## Year2008          -0.7097      0.4509   -1.57   0.12080
## Year2009          -0.4685      0.3939   -1.19   0.23904
## Year2010          -0.4410      0.3786   -1.16   0.24874
## Year2011          -0.7769      0.4683   -1.66   0.10244
## Year2012          -0.8054      0.3843   -2.10   0.04041 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.338
## Multiple R-squared:  0.499, Adjusted R-squared:  0.312
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 72 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.264  0.887  0.962  0.909  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          1.22e-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 33.68 1          5.804
## LastAuthorFemale  14.82 1          3.850
## Year             110.49 16          1.158

```

Residuals from first and last author



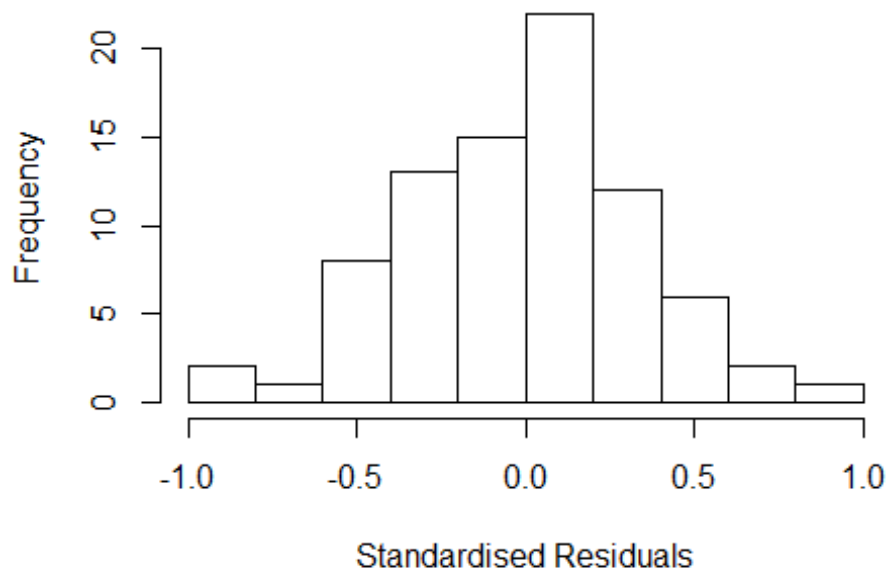
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.9557 -0.2461  0.0248  0.2152  0.8196
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.779      0.240    7.40 3.9e-10 ***
## FirstAuthorFemale1  0.191      0.135    1.42 0.16110
## LastAuthorFemale1 -0.040      0.117   -0.34 0.73396
## Year1997          -0.339      0.302   -1.12 0.26599
## Year1998          -0.541      0.304   -1.78 0.07988 .
## Year1999          -0.169      0.245   -0.69 0.49237
## Year2000          -0.661      0.239   -2.77 0.00737 **
## Year2001          -1.149      0.323   -3.56 0.00072 ***
## Year2002          -1.169      0.356   -3.28 0.00170 **
## Year2003          -1.067      0.188   -5.69 3.6e-07 ***
## Year2004          -0.255      0.230   -1.11 0.27114
## Year2005          -0.824      0.428   -1.92 0.05878 .
```

```

## Year2006          -0.960      0.233   -4.12  0.00011 ***
## Year2007          -0.676      0.250   -2.70  0.00895 **
## Year2008          -0.528      0.263   -2.01  0.04889 *
## Year2009          -0.414      0.248   -1.67  0.10033
## Year2010          -0.398      0.248   -1.60  0.11356
## Year2011          -0.704      0.325   -2.16  0.03431 *
## Year2012          -0.642      0.219   -2.93  0.00477 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.371
## Multiple R-squared:  0.421, Adjusted R-squared:  0.255
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 76 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.487  0.897  0.958  0.923  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] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 18.86 1      4.343
## Year              18.86 16      1.096

```

Residuals from first author



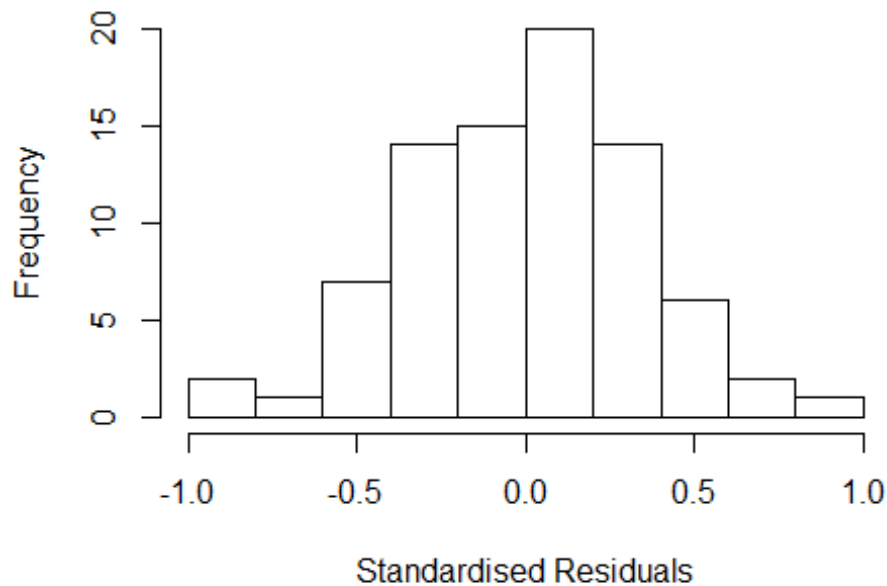
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9662 -0.2421 0.0203 0.2110 0.8308
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.779 0.240 7.43 3.3e-10 ***
## FirstAuthorFemale1 0.180 0.126 1.43 0.15627
## Year1997 -0.347 0.296 -1.17 0.24615
## Year1998 -0.540 0.303 -1.78 0.07968 .
## Year1999 -0.189 0.243 -0.78 0.44043
## Year2000 -0.655 0.235 -2.78 0.00704 **
## Year2001 -1.151 0.323 -3.56 0.00071 ***
## Year2002 -1.188 0.347 -3.43 0.00107 **
## Year2003 -1.096 0.183 -5.98 1.1e-07 ***
## Year2004 -0.269 0.236 -1.14 0.25733
## Year2005 -0.813 0.432 -1.88 0.06465 .
## Year2006 -0.955 0.229 -4.16 9.6e-05 ***
```

```

## Year2007          -0.687      0.267   -2.57  0.01242 *
## Year2008          -0.547      0.262   -2.09  0.04076 *
## Year2009          -0.414      0.247   -1.67  0.09929 .
## Year2010          -0.396      0.247   -1.61  0.11334
## Year2011          -0.704      0.327   -2.16  0.03493 *
## Year2012          -0.654      0.223   -2.93  0.00468 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.363
## Multiple R-squared:  0.423, Adjusted R-squared:  0.27
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 76 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.460  0.888  0.959  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      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 16.31 1          4.038
## Year            16.31 16          1.091

```

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
## -1.0502 -0.2425 0.0163 0.2329 0.7901
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.91541 0.21188 9.04 4.8e-13 ***
## LastAuthorFemale1 0.00975 0.12739 0.08 0.93921
## Year1997 -0.44184 0.32968 -1.34 0.18492
## Year1998 -0.67713 0.28257 -2.40 0.01949 *
## Year1999 -0.33028 0.20144 -1.64 0.10600
## Year2000 -0.70396 0.23156 -3.04 0.00342 **
## Year2001 -1.28564 0.29995 -4.29 6.2e-05 ***
## Year2002 -1.32978 0.29779 -4.47 3.3e-05 ***
## Year2003 -1.06216 0.19997 -5.31 1.5e-06 ***
## Year2004 -0.32078 0.20334 -1.58 0.11959
## Year2005 -0.86516 0.42793 -2.02 0.04739 *
## Year2006 -1.01216 0.24809 -4.08 0.00013 ***
```

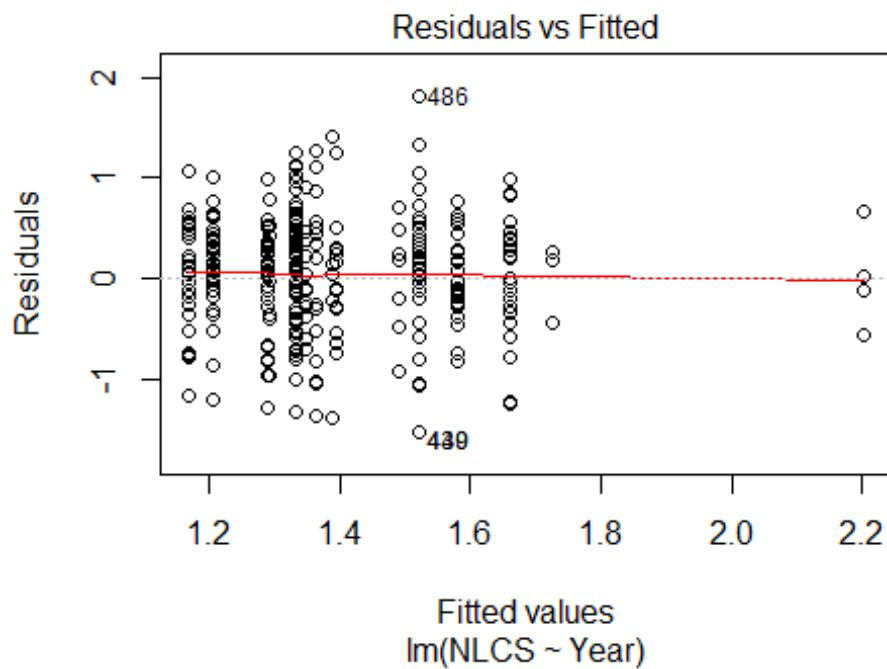
```

## Year2007          -0.78249      0.24775    -3.16  0.00242 **
## Year2008          -0.68839      0.22264    -3.09  0.00294 **
## Year2009          -0.55012      0.22084    -2.49  0.01534 *
## Year2010          -0.50741      0.22841    -2.22  0.02986 *
## Year2011          -0.84023      0.28903    -2.91  0.00501 **
## Year2012          -0.69415      0.23204    -2.99  0.00394 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.368
## Multiple R-squared:  0.408, Adjusted R-squared:  0.25
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 77 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.395  0.899  0.954  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] "Sample size for the above analysis: 82"
## [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    8   11    9    4    8   13   13   13   28   37   52   59   40   40
## 2011 2012
##   33   46
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    3    7    5    4    7   13   13   11   24   32   47   54   37   35
## 2011 2012

```



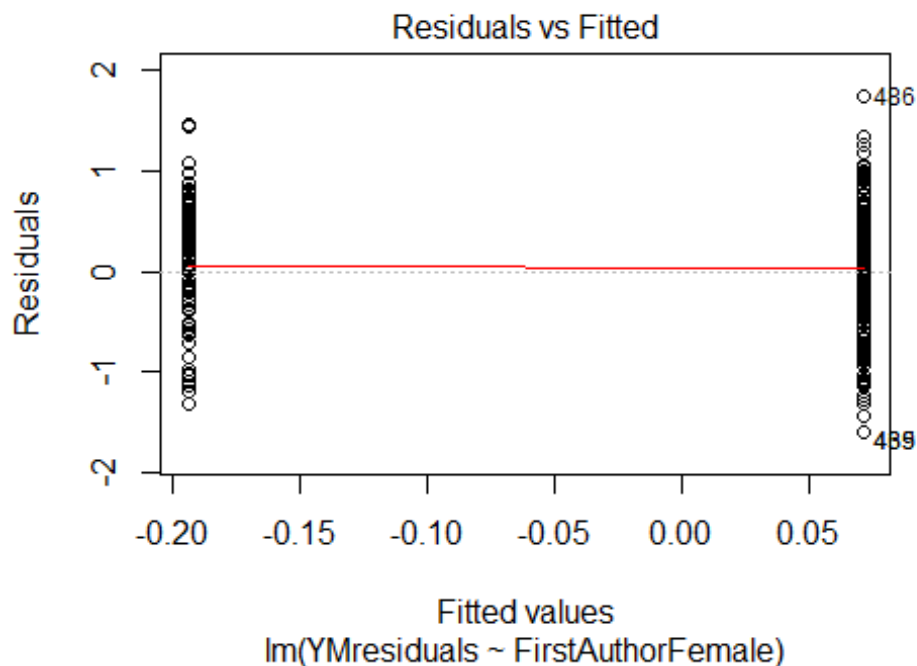
```
## 29 42
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4 3 7 5 4 7 11 11 11 24 30 44 51 35 34
## 2011 2012
## 27 38
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 24, df = 16, p-value = 0.09
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 6.1, df = 1, p-value = 0.01
## [1] "Female first author team size 2018 geometric mean: 3.69361781624108"
## [1] "Male first author team size 2018 geometric mean: 4.75913115793754"
## Warning in wilcox.test.default(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 = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.85958414912572"
## [1] "Male last author team size 2018 geometric mean: 4.03891132618332"

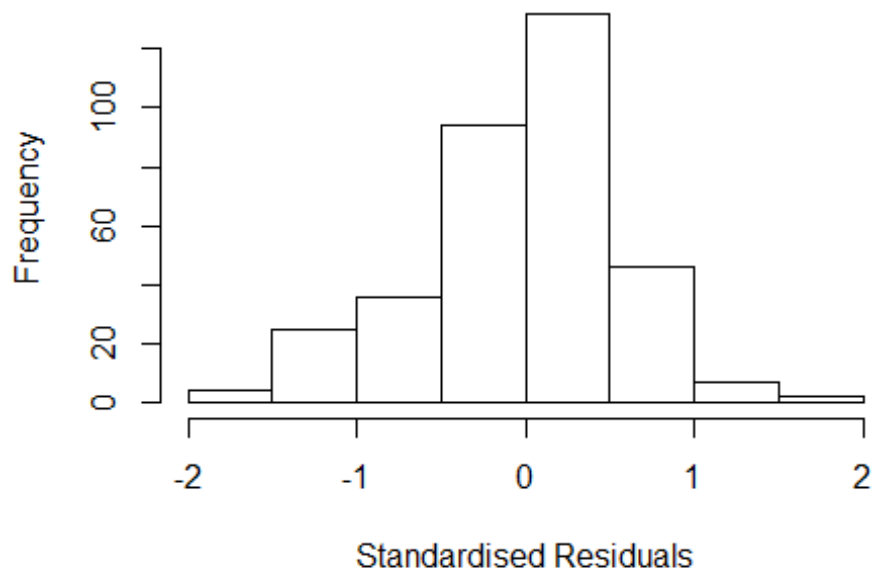
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 100, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
## Year as factors"
##
```

| | GVIF | Df | GVIF ^{1/(2*Df)} |
|-------------------|-------|----|--------------------------|
| FirstAuthorFemale | 1.360 | 1 | 1.166 |
| LastAuthorFemale | 1.285 | 1 | 1.133 |
| UniqueAuthors | 2.812 | 4 | 1.138 |
| Year | 3.760 | 16 | 1.042 |

Residuals from first and last author and team size



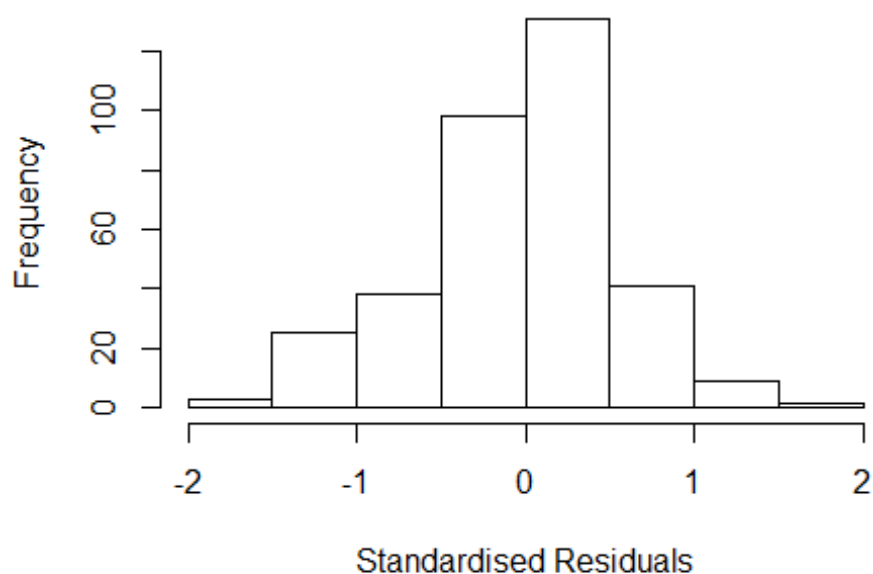
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.547 -0.362 0.050 0.349 1.542
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.05279 0.21000 9.78 < 2e-16 ***
## FirstAuthorFemale1 0.25805 0.09338 2.76 0.00605 **
## LastAuthorFemale1 -0.00988 0.08667 -0.11 0.90927
## UniqueAuthors2 0.07935 0.09090 0.87 0.38335
## UniqueAuthors3 0.24290 0.09975 2.44 0.01542 *
## UniqueAuthors4 0.09464 0.13521 0.70 0.48445
## UniqueAuthors5 0.15251 0.12773 1.19 0.23335
## Year1997 -0.40604 0.25469 -1.59 0.11185
## Year1998 -0.68351 0.26942 -2.54 0.01165 *
## Year1999 -0.91615 0.53235 -1.72 0.08621 .
```

```

## Year2000      -0.97093      0.74703      -1.30      0.19462
## Year2001      -0.97743      0.34220      -2.86      0.00456 **
## Year2002      -1.04507      0.24821      -4.21      3.3e-05 ***
## Year2003      -0.95810      0.27158      -3.53      0.00048 ***
## Year2004      -0.83406      0.41588      -2.01      0.04574 *
## Year2005      -0.74566      0.22522      -3.31      0.00104 **
## Year2006      -0.96037      0.22405      -4.29      2.4e-05 ***
## Year2007      -0.87475      0.21522      -4.06      6.1e-05 ***
## Year2008      -1.03362      0.22023      -4.69      4.0e-06 ***
## Year2009      -0.93195      0.24412      -3.82      0.00016 ***
## Year2010      -1.12649      0.22344      -5.04      7.7e-07 ***
## Year2011      -0.75893      0.23389      -3.24      0.00130 **
## Year2012      -0.75441      0.24798      -3.04      0.00254 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.523
## Multiple R-squared:  0.127, Adjusted R-squared:  0.0671
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 318 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.363  0.856  0.951  0.888  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          2.89e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.381 1      1.175
## LastAuthorFemale  1.292 1      1.137
## Year              1.414 16      1.011

```

Residuals from first and last author



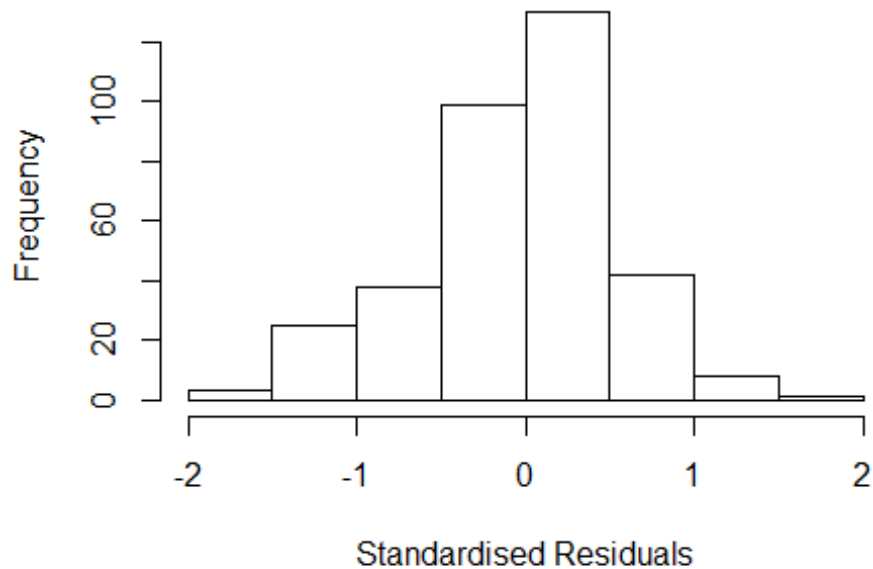
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## --> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.6800 -0.3488  0.0556  0.3505  1.6604
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.05654    0.20938   9.82  < 2e-16 ***
## FirstAuthorFemale1 0.26695    0.09583   2.79  0.00566 **
## LastAuthorFemale1 0.00933    0.08821   0.11  0.91584
## Year1997        -0.41982    0.25219  -1.66  0.09694 .
## Year1998        -0.66540    0.27957  -2.38  0.01788 *
## Year1999        -0.87622    0.49398  -1.77  0.07703 .
## Year2000        -1.00121    0.73217  -1.37  0.17242
## Year2001        -0.96943    0.37630  -2.58  0.01043 *
## Year2002        -1.00509    0.24146  -4.16  4.0e-05 ***
## Year2003        -0.90862    0.27011  -3.36  0.00086 ***
## Year2004        -0.84940    0.42011  -2.02  0.04401 *
## Year2005        -0.72438    0.22565  -3.21  0.00146 **
```

```

## Year2006      -0.90150    0.21881   -4.12  4.8e-05 ***
## Year2007      -0.86087    0.21614   -3.98  8.4e-05 ***
## Year2008      -0.97490    0.21804   -4.47  1.1e-05 ***
## Year2009      -0.87868    0.23983   -3.66  0.00029 ***
## Year2010      -1.07357    0.22162   -4.84  2.0e-06 ***
## Year2011      -0.67104    0.23123   -2.90  0.00396 **
## Year2012      -0.65285    0.23661   -2.76  0.00612 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.53
## Multiple R-squared:  0.113, Adjusted R-squared:  0.0637
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 321 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.294  0.858  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      2.89e-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 1.242 1      1.114
## Year      1.242 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
## -1.6803 -0.3499 0.0575 0.3499 1.6507
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.062 0.202 10.19 < 2e-16 ***
## FirstAuthorFemale1 0.271 0.092 2.95 0.00339 **
## Year1997 -0.423 0.251 -1.69 0.09265 .
## Year1998 -0.668 0.279 -2.40 0.01718 *
## Year1999 -0.878 0.495 -1.77 0.07697 .
## Year2000 -1.002 0.734 -1.36 0.17328
## Year2001 -0.969 0.379 -2.56 0.01096 *
## Year2002 -1.008 0.241 -4.18 3.8e-05 ***
## Year2003 -0.908 0.271 -3.35 0.00089 ***
## Year2004 -0.851 0.420 -2.03 0.04360 *
## Year2005 -0.726 0.226 -3.21 0.00145 **
## Year2006 -0.904 0.218 -4.14 4.5e-05 ***
```

```

## Year2007          -0.862      0.217   -3.97  9.0e-05 ***
## Year2008          -0.977      0.218   -4.47  1.1e-05 ***
## Year2009          -0.880      0.241   -3.65  0.00031 ***
## Year2010          -1.074      0.222   -4.83  2.1e-06 ***
## Year2011          -0.672      0.232   -2.89  0.00405 **
## Year2012          -0.653      0.238   -2.75  0.00637 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.529
## Multiple R-squared:  0.113, Adjusted R-squared:  0.0665
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 323 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.291  0.858  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      2.89e-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 1.217  1      1.103
## Year              1.217 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
## -1.5991 -0.3771  0.0271  0.3543  1.8714

```



```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.0842    0.2201   9.47 < 2e-16 ***
## LastAuthorFemale1  0.1395    0.0901   1.55  0.12241
## Year1997         -0.3993    0.2690  -1.48  0.13876
## Year1998         -0.6318    0.2939  -2.15  0.03231 *
## Year1999         -0.8106    0.4947  -1.64  0.10224
## Year2000         -0.8918    0.7308  -1.22  0.22323
## Year2001         -0.9160    0.3269  -2.80  0.00537 **
## Year2002         -0.8954    0.2410  -3.72  0.00024 ***
## Year2003         -0.8255    0.2744  -3.01  0.00283 **
## Year2004         -0.7641    0.3907  -1.96  0.05135 .
## Year2005         -0.6154    0.2262  -2.72  0.00686 **
## Year2006         -0.8010    0.2262  -3.54  0.00046 ***
## Year2007         -0.7869    0.2235  -3.52  0.00049 ***
## Year2008         -0.8908    0.2231  -3.99  8.1e-05 ***
## Year2009         -0.8107    0.2533  -3.20  0.00151 **
## Year2010         -1.0136    0.2292  -4.42  1.3e-05 ***
## Year2011         -0.5742    0.2380  -2.41  0.01638 *
## Year2012         -0.6246    0.2468  -2.53  0.01186 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.531
## Multiple R-squared:  0.084, Adjusted R-squared:  0.0366
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 317 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.189  0.869  0.947  0.886  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      2.89e-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: 346"
## [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]"
##
## 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    2    2    3    1    2    1    1    9
##
## 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    0    1    1    0    0    1    0    8
##
## 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    0    1    1    0    0    1    0    8
## [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: 2.82842712474619"

## 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 = 1
## 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: 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 = 1.5, 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"
## [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 2902"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2005 2006 2007 2008 2009 2010 2011
##    2    2    4    3    3    2    4    3    5    6    6   11   14   17   15
## 2012
##    10
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2005 2006 2007 2008 2009 2010 2011
##    2    1    3    2    1    1    3    3    5    4    6   11   11   10   10
## 2012
##     8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2005 2006 2007 2008 2009 2010 2011
##    2    1    2    2    1    1    2    3    5    4    6    9   11   10   10
## 2012
##     7
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.34034731932072"
## [1] "Male first author team size 2018 geometric mean: 4.24264068711929"

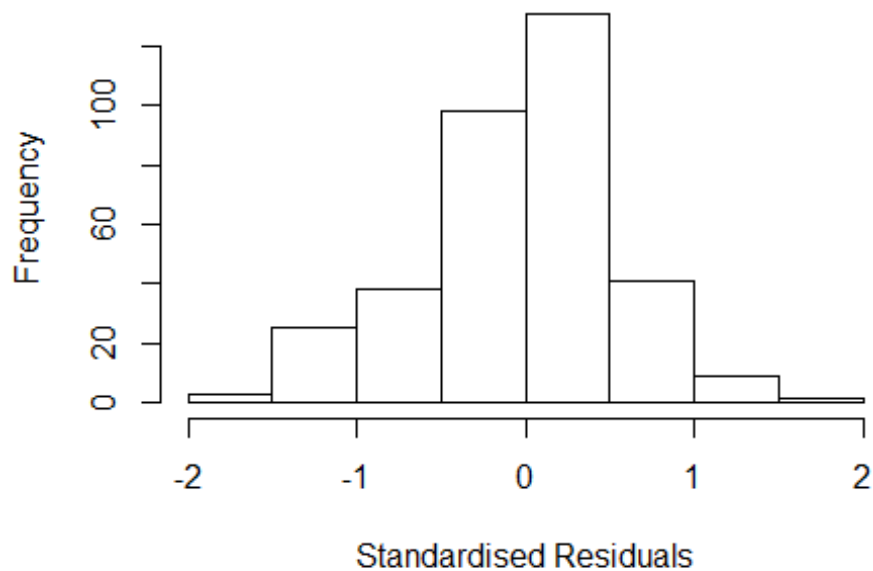
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## 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: 2.15443469003188"
## [1] "Male last author team size 2018 geometric mean: 3.77976314968462"

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

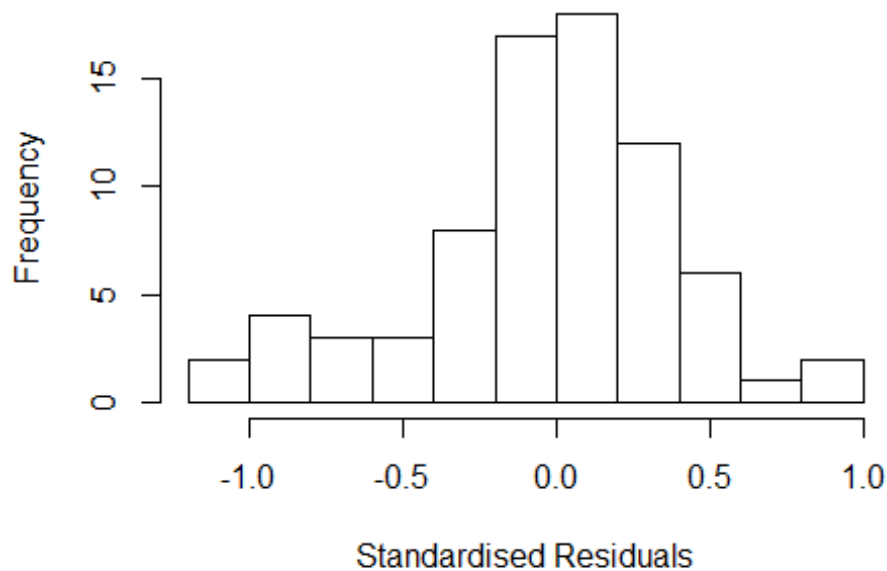
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2.5, p-value = 0.5
## 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 5.124e+14  1      2.264e+07
## LastAuthorFemale  1.182e+15  1      3.437e+07
## UniqueAuthors    6.198e+31  4      9.419e+03
## Year              8.424e+45 15      3.395e+01
```

Residuals from first and last author and team size



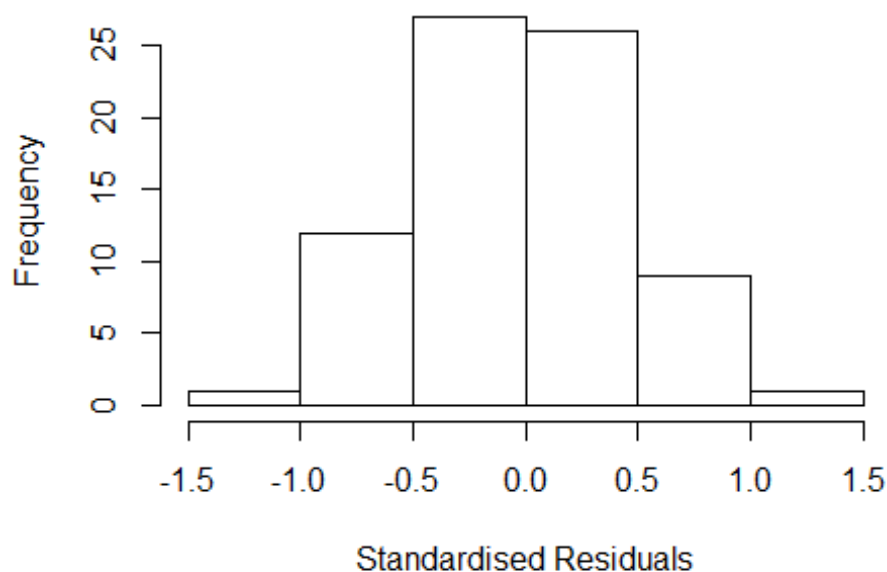
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0419 -0.2148 0.0215 0.2168 0.8516
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.1499 0.1747 0.86 0.39468
## FirstAuthorFemale1 0.0551 0.1648 0.33 0.73953
## LastAuthorFemale1 0.1877 0.1650 1.14 0.26032
## UniqueAuthors2 0.0642 0.1456 0.44 0.66108
## UniqueAuthors3 0.4490 0.2921 1.54 0.13012
## UniqueAuthors4 0.6253 0.1174 5.33 2e-06 ***
## UniqueAuthors5 0.6509 0.2177 2.99 0.00419 **
## Year1997 -0.0558 0.2606 -0.21 0.83125
## Year1998 0.2310 0.4448 0.52 0.60557
## Year1999 -0.1499 0.1747 -0.86 0.39468
```

```

## Year2000          0.5831      0.3291      1.77  0.08210 .
## Year2001          0.9391      0.2706      3.47  0.00103 **
## Year2002          0.7815      0.1965      3.98  0.00021 ***
## Year2003          0.1309      0.2130      0.61  0.54127
## Year2005          0.1220      0.2743      0.44  0.65824
## Year2006          0.8293      0.5618      1.48  0.14567
## Year2007          0.4455      0.2952      1.51  0.13711
## Year2008          0.3772      0.2612      1.44  0.15445
## Year2009          0.4751      0.2951      1.61  0.11318
## Year2010         -0.0373      0.2149     -0.17  0.86297
## Year2011          0.5001      0.2976      1.68  0.09870 .
## Year2012          0.4430      0.4191      1.06  0.29511
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.425
## Multiple R-squared:  0.476, Adjusted R-squared:  0.272
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 66 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.528  0.903   0.970   0.911   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      1.32e-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 -3.379e+15  1      NaN
## LastAuthorFemale  -5.670e+14  1      NaN
## Year              -8.948e+28 15      NaN

```

Residuals from first and last author



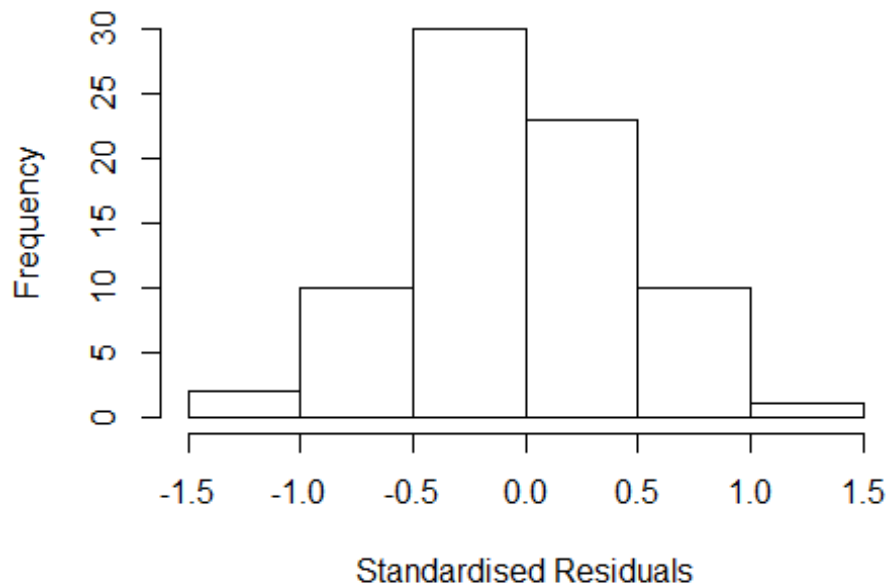
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.04e+00 -2.86e-01 2.91e-16 3.11e-01 1.30e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.1820 0.1317 1.38 0.17214
## FirstAuthorFemale1 -0.0792 0.1988 -0.40 0.69181
## LastAuthorFemale1 0.1404 0.1731 0.81 0.42045
## Year1997 0.5630 0.1317 4.28 7.2e-05 ***
## Year1998 0.5565 0.8876 0.63 0.53314
## Year1999 -0.1820 0.1317 -1.38 0.17214
## Year2000 1.3362 0.2384 5.60 6.1e-07 ***
## Year2001 1.1528 0.2583 4.46 3.8e-05 ***
## Year2002 0.8486 0.2264 3.75 0.00041 ***
## Year2003 0.7773 0.1561 4.98 6.0e-06 ***
## Year2005 0.4222 0.2064 2.04 0.04541 *
## Year2006 1.1555 0.6057 1.91 0.06140 .
```

```

## Year2007          0.7385      0.2860      2.58  0.01235 *
## Year2008          0.6526      0.2666      2.45  0.01744 *
## Year2009          0.8004      0.2967      2.70  0.00913 **
## Year2010          0.3194      0.2657      1.20  0.23416
## Year2011          0.9739      0.3076      3.17  0.00246 **
## Year2012          0.7546      0.3525      2.14  0.03653 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.52
## Multiple R-squared:  0.293, Adjusted R-squared:  0.0862
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 61 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.514  0.896  0.950  0.915  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      1.32e-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 -74.83 1      NaN
## Year              -74.83 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
## -1.03e+00 -3.21e-01 -4.44e-16 2.76e-01 1.31e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.1820 0.1318 1.38 0.17253
## FirstAuthorFemale1 -0.0173 0.2045 -0.08 0.93299
## Year1997 0.5630 0.1318 4.27 7.2e-05 ***
## Year1998 0.5565 0.9213 0.60 0.54814
## Year1999 -0.1820 0.1318 -1.38 0.17253
## Year2000 1.2743 0.2433 5.24 2.3e-06 ***
## Year2001 1.2313 0.2433 5.06 4.4e-06 ***
## Year2002 0.8176 0.2124 3.85 0.00029 ***
## Year2003 0.7567 0.1530 4.95 6.6e-06 ***
## Year2005 0.4402 0.2052 2.15 0.03608 *
## Year2006 1.2115 0.5823 2.08 0.04181 *
## Year2007 0.7933 0.2763 2.87 0.00567 **
```

```

## Year2008          0.7073      0.2684      2.64  0.01071 *
## Year2009          0.8653      0.2976      2.91  0.00512 **
## Year2010          0.3839      0.2643      1.45  0.15176
## Year2011          1.0084      0.3046      3.31  0.00159 **
## Year2012          0.8187      0.3646      2.25  0.02849 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.507
## Multiple R-squared:  0.29,   Adjusted R-squared:  0.0976
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 63 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.485  0.886  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      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 1.019e+14  1      1.009e+07
## Year              1.019e+14 15      2.930e+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
## -1.04e+00 -3.05e-01  4.72e-16  3.23e-01  1.28e+00
##

```

```

## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.182      0.132   1.38  0.17243
## LastAuthorFemale1 0.114      0.166   0.69  0.49366
## Year1997          0.563      0.132   4.27  7.1e-05 ***
## Year1998          0.557      0.914   0.61  0.54495
## Year1999         -0.182      0.132  -1.38  0.17243
## Year2000          1.257      0.132   9.54  1.5e-13 ***
## Year2001          1.100      0.212   5.20  2.6e-06 ***
## Year2002          0.809      0.182   4.45  3.8e-05 ***
## Year2003          0.751      0.136   5.51  8.3e-07 ***
## Year2005          0.411      0.202   2.03  0.04680 *
## Year2006          1.121      0.646   1.73  0.08803 .
## Year2007          0.681      0.238   2.86  0.00578 **
## Year2008          0.601      0.212   2.83  0.00633 **
## Year2009          0.746      0.253   2.95  0.00461 **
## Year2010          0.284      0.245   1.16  0.25172
## Year2011          0.914      0.256   3.57  0.00072 ***
## Year2012          0.726      0.345   2.10  0.03957 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.51
## Multiple R-squared:  0.294, Adjusted R-squared:  0.102
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 62 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.510  0.886  0.951  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      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 2903"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 2003 2005 2010 2011 2012
##    2    1    1    2    1    1    2    1
##
## 1996 1997 1998 2003 2005 2010 2011 2012
##    2    0    0    2    1    1    1    0
##
## 1996 1997 1998 2003 2005 2010 2011 2012
##    2    0    0    2    1    1    1    0
## [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: 7"
## [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]"
##
## 2009 2011 2012
##    2    1    3
##
## 2009 2011 2012
##    2    1    2
##
## 2009 2011 2012
##    2    1    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: 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: 5"
## [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]"
##
## 2007 2008 2009 2010 2011 2012
##    1    2    4    7    9   14
##
## 2007 2008 2009 2010 2011 2012
##    1    2    4    6    8   11
##
## 2007 2008 2009 2010 2011 2012
##    1    2    4    5    8   10
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.21336383940064"
## [1] "Male first author team size 2018 geometric mean: 1.7099759466767"

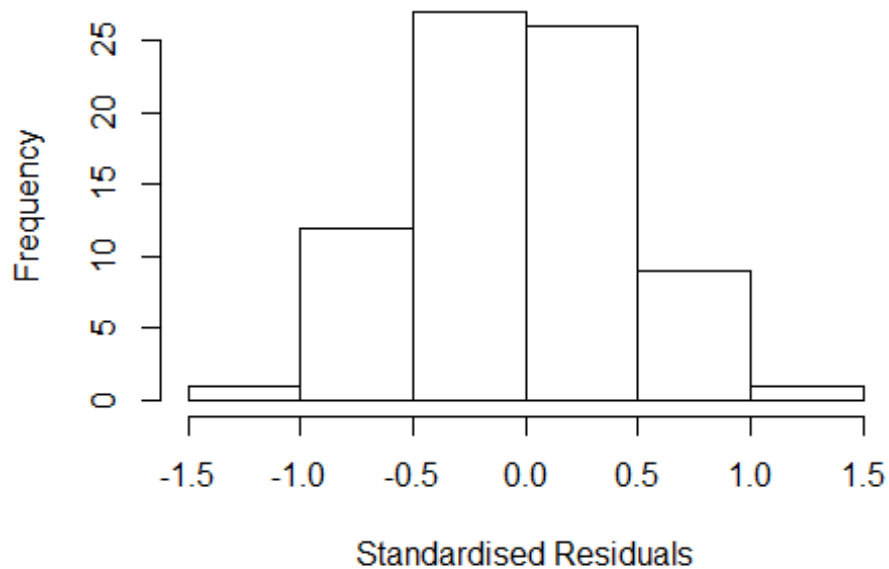
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## 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] "Female last author team size 2018 geometric mean: 2.60517108469735"
## [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

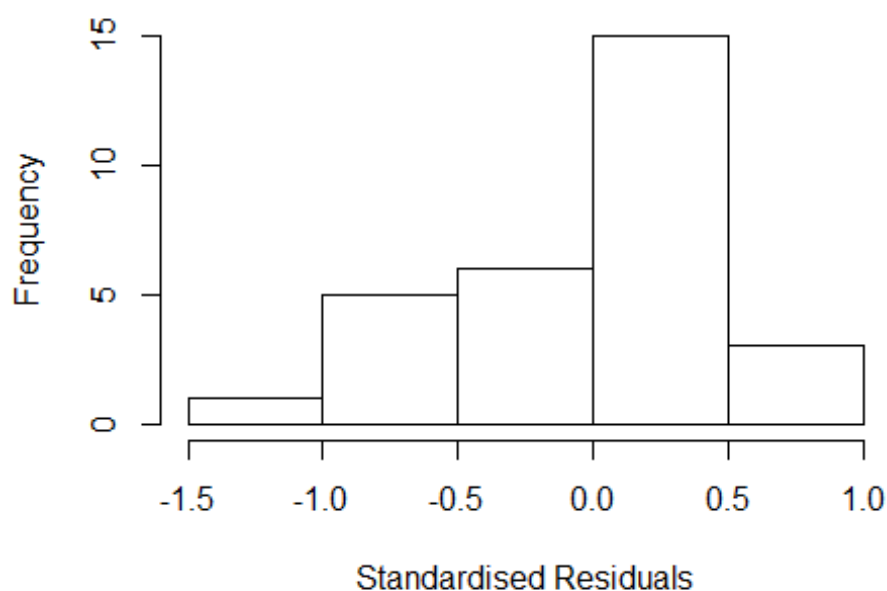
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 10, 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"
## [1] "Regression 2: First author gender, Last author gender, Year as
##      factors"
##               GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale  9.426  1          3.070
## LastAuthorFemale  29.797  1          5.459
## Year               38.924  5          1.442
```

Residuals from first and last author



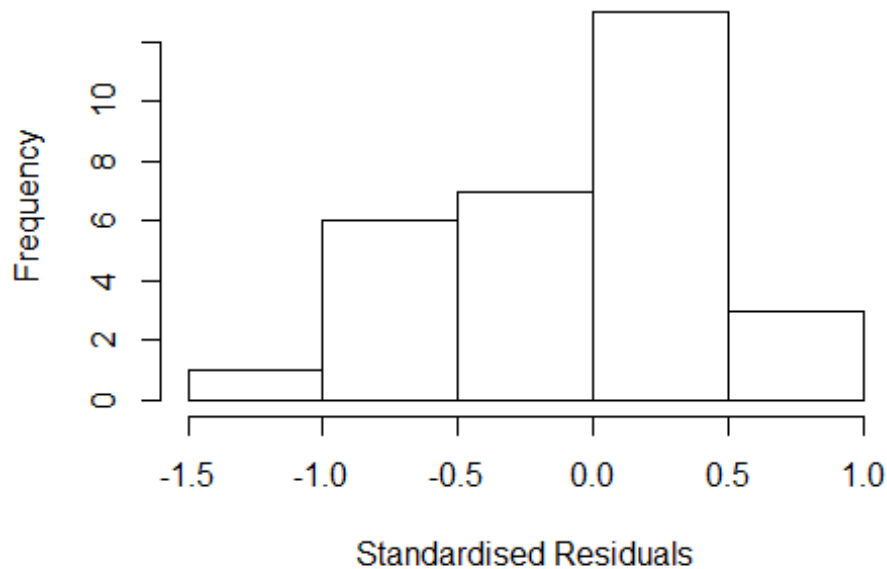
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3876 -0.4102 0.0637 0.3231 0.7234
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6731 0.1993 3.38 0.0027 **
## FirstAuthorFemale1 -0.0831 0.4468 -0.19 0.8542
## LastAuthorFemale1 0.5780 0.4049 1.43 0.1674
## Year2008 0.7425 0.2078 3.57 0.0017 **
## Year2009 -0.1874 0.2515 -0.75 0.4640
## Year2010 -0.2834 0.2393 -1.18 0.2489
## Year2011 0.1279 0.2137 0.60 0.5555
## Year2012 0.2196 0.2294 0.96 0.3489
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.54
```

```

## Multiple R-squared:  0.269, Adjusted R-squared:  0.0362
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 3 weights are ~= 1. The remaining 27 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.490  0.919  0.944  0.924  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           3.33e-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 2.07 1          1.439
## Year              2.07 5          1.075

```


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
## -1.3011 -0.4621 0.0898 0.4109 0.8099
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8562 0.2193 3.90 0.00071 ***
## FirstAuthorFemale1 0.3118 0.2193 1.42 0.16855
## Year2008 0.4535 0.2655 1.71 0.10112
## Year2009 -0.2368 0.2408 -0.98 0.33560
## Year2010 -0.2086 0.2455 -0.85 0.40416
## Year2011 0.0662 0.2043 0.32 0.74875
## Year2012 0.1331 0.2240 0.59 0.55808
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.662
## Multiple R-squared: 0.164, Adjusted R-squared: -0.0544
```

```

## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 3 weights are ~= 1. The remaining 27 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.679  0.930  0.951   0.944   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      3.33e-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 4.308 1      2.076
## Year      4.308 5      1.157

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.3854 -0.4184  0.0884  0.3735  0.7256
##
## Coefficients:
##      Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.655      0.197   3.32   0.003 **
## LastAuthorFemale1  0.513      0.197   2.60   0.016 *
## Year2008          0.710      0.121   5.86 5.7e-06 ***
## Year2009         -0.182      0.249  -0.73   0.473
## Year2010         -0.267      0.218  -1.22   0.234
## Year2011          0.123      0.203   0.61   0.549
## Year2012          0.217      0.226   0.96   0.347
## ---

```

```

## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.572
## Multiple R-squared:  0.257, Adjusted R-squared:  0.0629
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 3 weights are ~= 1. The remaining 27 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.537  0.922  0.949  0.932  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      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 2906"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1997 1998 1999 2002 2004 2005 2008 2010 2011 2012
##    1    1    2    2    3    3    3    1    4    1
##
## 1997 1998 1999 2002 2004 2005 2008 2010 2011 2012
##    1    1    1    2    2    3    3    1    3    0
##
## 1997 1998 1999 2002 2004 2005 2008 2010 2011 2012
##    1    1    1    2    2    2    3    1    3    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: 16"
## [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]"
##
## 1997 1999 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    2    3    1    1    6    6    4    3    4
##
## 1997 1999 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    0    2    3    1    1    5    5    4    3    3
##
## 1997 1999 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    0    2    2    1    1    5    5    4    3    3
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.5874010519682"
## [1] "Male first author team size 2018 geometric mean: 3.47602664488645"

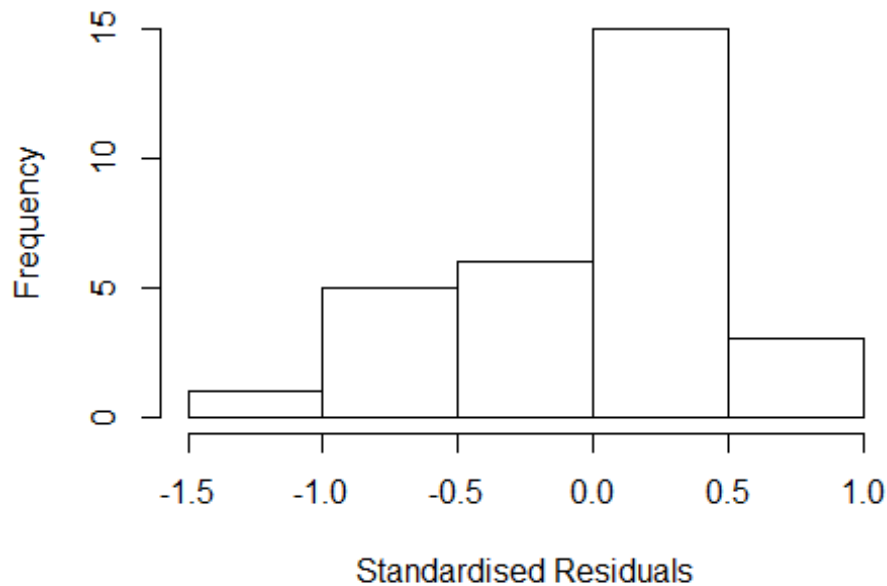
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2.5, p-value = 0.5
## 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: 2.78651802273122"

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

```

Residuals from last author



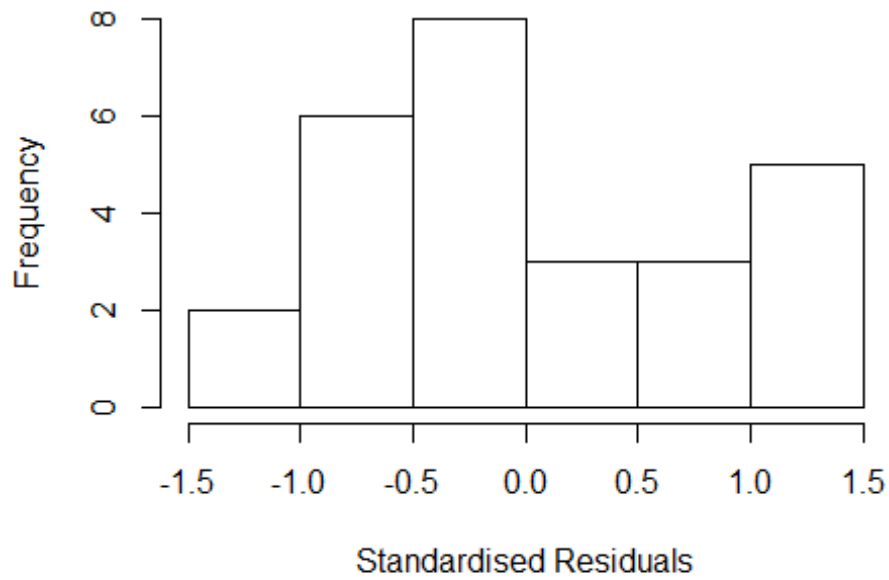
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0.5, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
##      Year as factors"
## [1] "Regression 2: First author gender, Last author gender, Year as
##      factors"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -3.876e+15  1      NaN
## LastAuthorFemale -2.612e+15  1      NaN
## Year              -6.093e+29  9      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
```

```
## -1.40e+00 -5.63e-01 -5.11e-15 6.20e-01 1.26e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      3.736      0.462   8.08 7.6e-07 ***
## FirstAuthorFemale1 -0.129      0.511  -0.25 0.80340
## LastAuthorFemale1  0.608      0.498   1.22 0.24120
## Year2004          -2.954      1.109  -2.66 0.01769 *
## Year2005          -3.477      0.611  -5.69 4.3e-05 ***
## Year2006          -2.435      0.498  -4.89 0.00020 ***
## Year2007          -2.400      0.462  -5.19 0.00011 ***
## Year2008          -2.784      0.421  -6.61 8.2e-06 ***
## Year2009          -1.962      0.466  -4.21 0.00076 ***
## Year2010          -2.203      0.530  -4.15 0.00085 ***
## Year2011          -2.757      0.587  -4.70 0.00029 ***
## Year2012          -2.346      0.318  -7.37 2.3e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 1.24
## Multiple R-squared:  0.422, Adjusted R-squared:  -0.00157
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 3 weights are ~= 1. The remaining 24 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.887 0.917 0.968 0.959 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           3.70e-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"

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

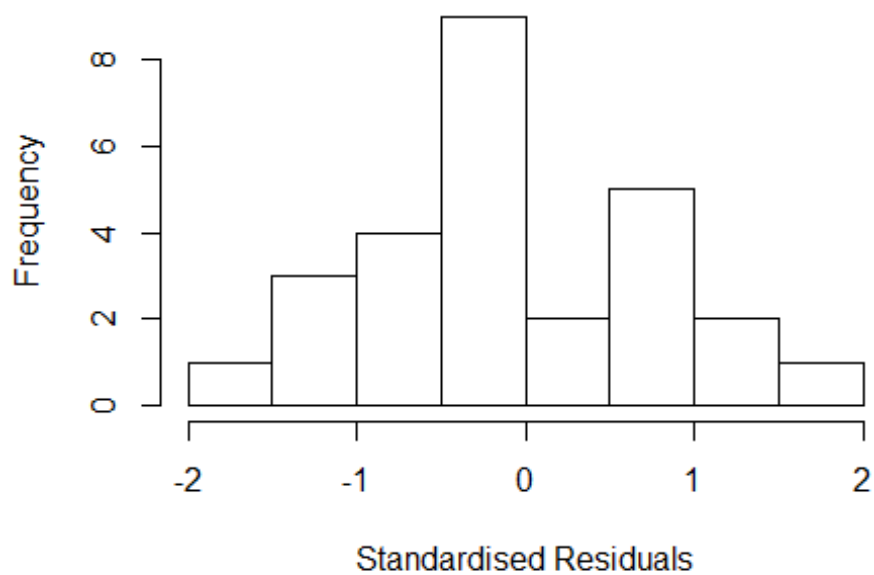
## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

Residuals from first and last author



| | | | | |
|----|-------------------|------|----|---------------------|
| ## | | GVIF | Df | $GVIF^{(1/(2*Df))}$ |
| ## | FirstAuthorFemale | NaN | 1 | NaN |
| ## | Year | NaN | 9 | NaN |

Residuals from first author

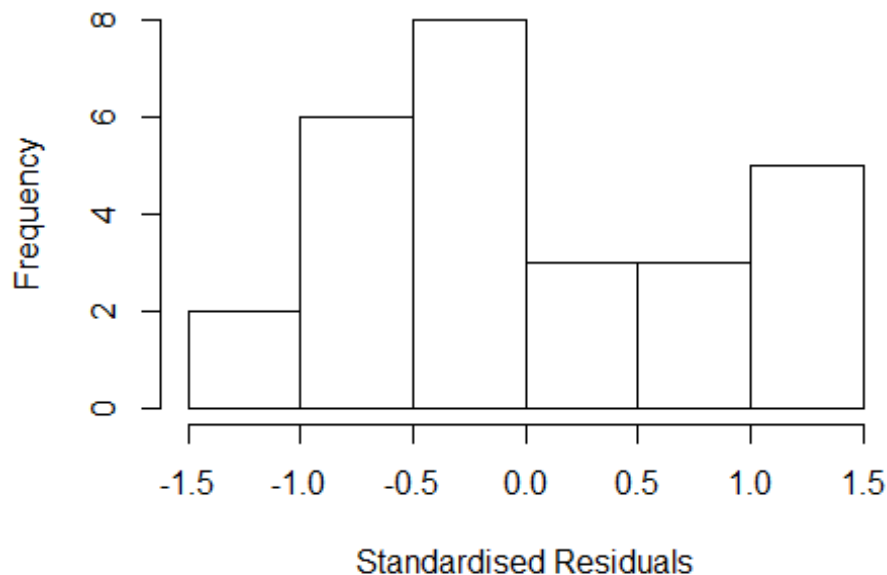


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.6695 -0.5512 -0.0185 0.5737 1.7449
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 4.054 0.508 7.97 5.8e-07 ***
## FirstAuthorFemale1 0.161 0.508 0.32 0.75616
## Year2004 -2.954 1.123 -2.63 0.01817 *
## Year2005 -3.781 0.314 -12.05 1.9e-09 ***
## Year2006 -3.043 0.000 -Inf < 2e-16 ***
## Year2007 -2.718 0.508 -5.35 6.5e-05 ***
## Year2008 -3.011 0.469 -6.43 8.4e-06 ***
## Year2009 -1.970 0.495 -3.98 0.00108 **
## Year2010 -2.546 0.591 -4.31 0.00054 ***
## Year2011 -2.859 0.673 -4.25 0.00062 ***
## Year2012 -2.346 0.319 -7.36 1.6e-06 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 1.21
## Multiple R-squared: 0.39, Adjusted R-squared: 0.00913
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~ 1. The remaining 21 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.821 0.914 0.972 0.947 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 3.70e-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"
##          GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 8.717e+14 1 2.952e+07
## Year 8.717e+14 9 6.761e+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
## -1.46e+00 -6.03e-01 -2.22e-15 6.43e-01 1.32e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 3.659 0.427 8.56 2.3e-07 ***
## LastAuthorFemale1 0.556 0.427 1.30 0.21212
## Year2004 -2.954 1.170 -2.53 0.02249 *
## Year2005 -3.504 0.583 -6.01 1.8e-05 ***
```

```

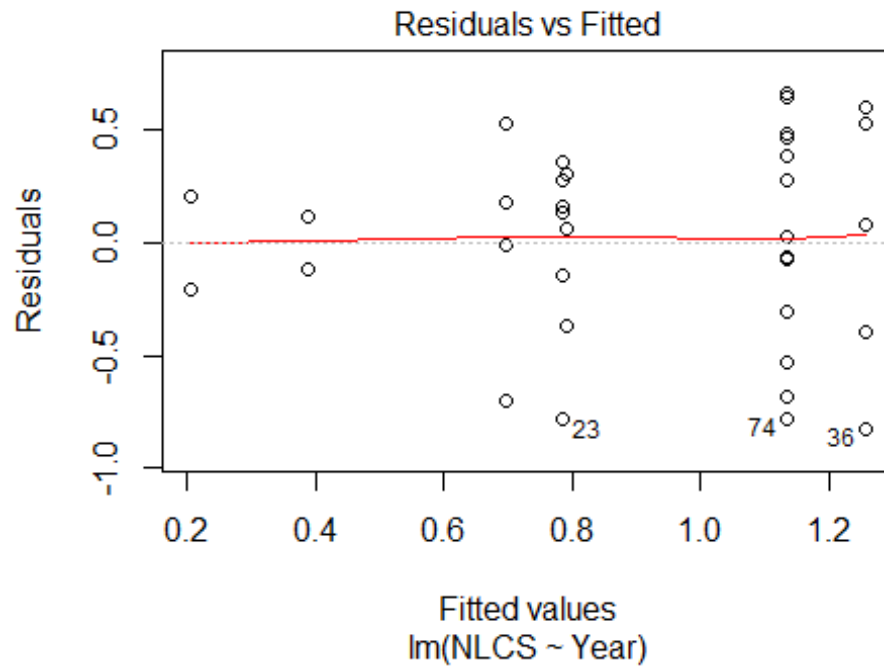
## Year2006          -2.487          0.427      -5.82  2.6e-05 ***
## Year2007          -2.323          0.427      -5.44  5.5e-05 ***
## Year2008          -2.747          0.508      -5.41  5.7e-05 ***
## Year2009          -1.881          0.381      -4.94  0.00015 ***
## Year2010          -2.199          0.540      -4.07  0.00089 ***
## Year2011          -2.736          0.582      -4.70  0.00024 ***
## Year2012          -2.347          0.320      -7.33  1.7e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 1.13
## Multiple R-squared:  0.424, Adjusted R-squared:  0.0644
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 3 weights are ~= 1. The remaining 24 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.855  0.906  0.965  0.950  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          3.70e-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: 27"
## [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 2005 2006 2007 2008 2009 2011 2012
##    1    2    1    1    2    2    1    2
##
## 2002 2005 2006 2007 2008 2009 2011 2012
##    1    0    0    0    1    1    1    1
##
## 2002 2005 2006 2007 2008 2009 2011 2012

```

```

##      1      0      0      0      1      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: 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 2909"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 2000 2001 2002 2006 2007 2008 2009 2010 2011 2012
##      1      1      2      2      2      6      3      13      16      18
##
## 2000 2001 2002 2006 2007 2008 2009 2010 2011 2012
##      0      0      0      2      2      4      3      6      5      14
##
## 2000 2001 2002 2006 2007 2008 2009 2010 2011 2012
##      0      0      0      2      2      4      3      6      5      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 = 2.4, df = 6, p-value = 0.9

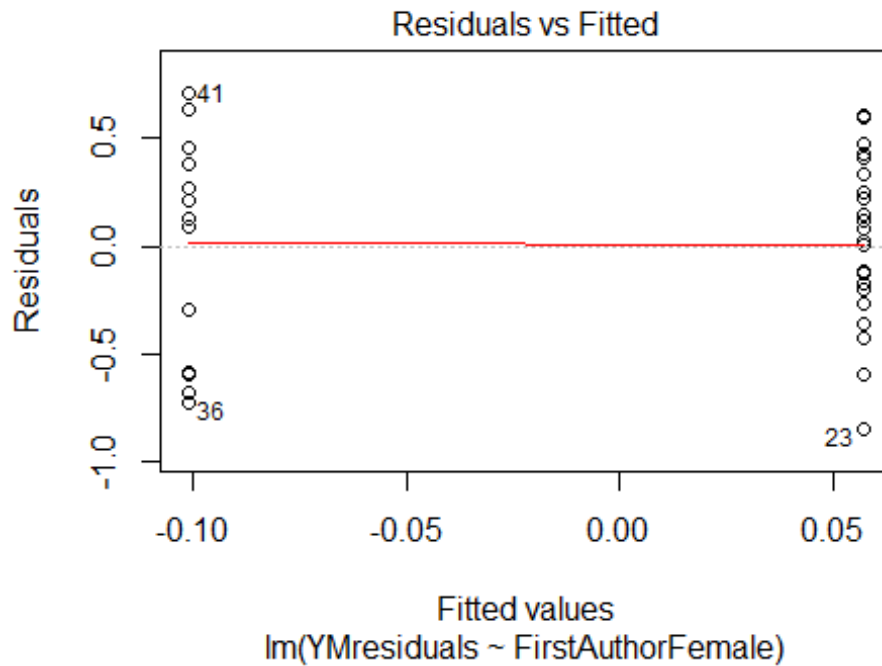
```



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

## [1] "Female first author team size 2018 geometric mean: 5.74456264653803"
## [1] "Male first author team size 2018 geometric mean: 3.46410161513775"

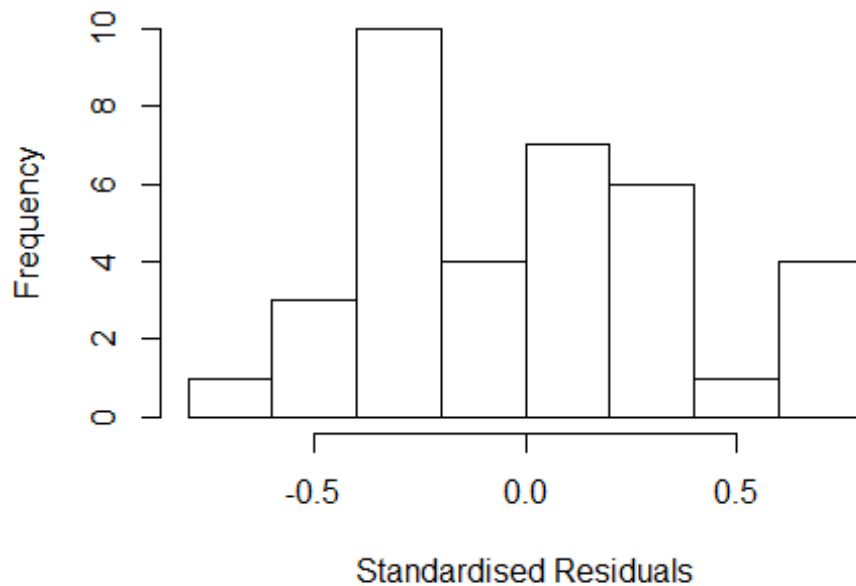
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 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: 4.46091344257344"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
```

| | GVIF | Df | GVIF ^{1/(2*Df)} |
|-------------------|--------|----|--------------------------|
| FirstAuthorFemale | 7.428 | 1 | 2.725 |
| LastAuthorFemale | 10.869 | 1 | 3.297 |
| UniqueAuthors | 39.029 | 4 | 1.581 |
| Year | 28.955 | 6 | 1.324 |

Residuals from first and last author and team size



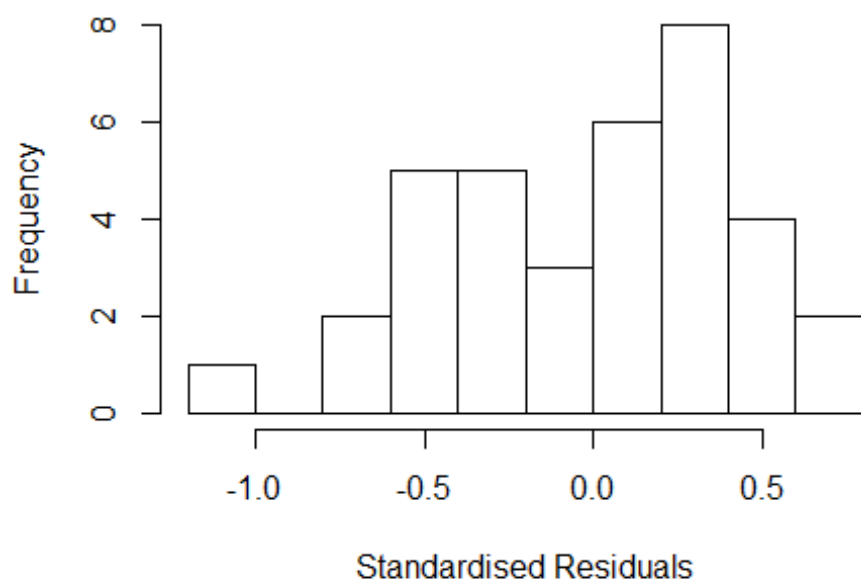
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.7173 -0.3129 0.0135 0.2569 0.7659
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.2569 0.2546 1.01 0.324
## FirstAuthorFemale1 0.4680 0.3181 1.47 0.155
## LastAuthorFemale1 -0.2881 0.3321 -0.87 0.395
## UniqueAuthors2 -0.2038 0.2341 -0.87 0.393
## UniqueAuthors3 0.1723 0.2504 0.69 0.498
## UniqueAuthors4 -0.5511 0.6228 -0.88 0.385
## UniqueAuthors5 0.0175 0.4691 0.04 0.971
## Year2007 -0.1299 0.3022 -0.43 0.671
## Year2008 0.4937 0.3816 1.29 0.209
## Year2009 0.3696 0.3535 1.05 0.307
```

```

## Year2010          0.5435      0.2998      1.81      0.083 .
## Year2011          0.9838      0.4262      2.31      0.030 *
## Year2012          0.6793      0.3410      1.99      0.058 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.503
## Multiple R-squared:  0.436, Adjusted R-squared:  0.142
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 3 weights are ~= 1. The remaining 33 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.800  0.909  0.965  0.944  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      2.78e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.352 1      1.533
## LastAuthorFemale  5.145 1      2.268
## Year              4.588 6      1.135

```

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
## -1.0758 -0.3575 0.0733 0.3061 0.7350
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.2313 0.2356 0.98 0.3351
## FirstAuthorFemale1 0.3154 0.2400 1.31 0.1997
## LastAuthorFemale1 -0.2463 0.2872 -0.86 0.3988
## Year2007 -0.0955 0.3364 -0.28 0.7787
## Year2008 0.4648 0.3454 1.35 0.1895
## Year2009 0.4934 0.3469 1.42 0.1664
## Year2010 0.5291 0.3710 1.43 0.1653
## Year2011 0.9845 0.3182 3.09 0.0046 **
## Year2012 0.7465 0.2460 3.04 0.0053 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

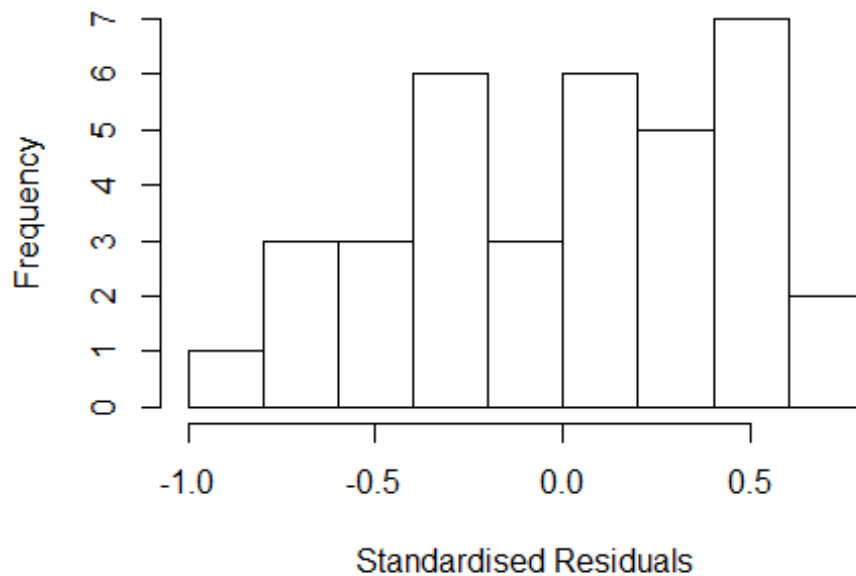


```

## Robust residual standard error: 0.509
## Multiple R-squared: 0.341, Adjusted R-squared: 0.146
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 2 weights are ~= 1. The remaining 34 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.634  0.913   0.960   0.936   0.985   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      2.78e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.285 1      1.134
## Year              1.285 6      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
## -0.8763 -0.3310 0.0366 0.3476 0.6792
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.292 0.177 1.65 0.1107
## FirstAuthorFemale1 0.193 0.180 1.07 0.2919
## Year2007 -0.281 0.232 -1.21 0.2356
## Year2008 0.364 0.273 1.33 0.1934
## Year2009 0.308 0.244 1.26 0.2168
## Year2010 0.391 0.241 1.62 0.1156
## Year2011 0.890 0.297 3.00 0.0056 **
## Year2012 0.707 0.211 3.35 0.0023 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.556
```

```

## Multiple R-squared:  0.315, Adjusted R-squared:  0.144
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 31 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.786  0.900  0.952   0.939   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           2.78e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01                5.00e-01
## nResample      max.it    best.r.s    k.fast.s      k.max maxit.scale
##           500           50         2         1         1000         200
## trace.lev      mts    compute.rd
##           0           1000         0
##           psi                subsampling                cov
##           "bisquare"                "nonsingular"                ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 4.145  1         2.036
## Year              4.145  6         1.126

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.8497 -0.3387  0.0495  0.2963  0.6921
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.3890    0.0813   4.78   5e-05 ***
## LastAuthorFemale1 -0.0698    0.2196  -0.32  0.75285
## Year2007        -0.1142    0.2776  -0.41  0.68396
## Year2008         0.3475    0.2810   1.24  0.22655
## Year2009         0.4745    0.2886   1.64  0.11139
## Year2010         0.4607    0.2447   1.88  0.07012 .
## Year2011         0.9094    0.2822   3.22  0.00321 **

```

```

## Year2012          0.7708      0.1759      4.38  0.00015 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.528
## Multiple R-squared:  0.299, Adjusted R-squared:  0.124
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 32 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.778  0.901  0.957  0.933  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          2.78e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01          5.00e-01
##   nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500          50          2          1          1000          200
##   trace.lev      mts      compute.rd
##           0          1000          0
##           psi          subsampling          cov
##           "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 36"
## [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]"
##
## 2000 2001 2002 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    1    1    1    1    1    4    3    3    2    5
##
## 2000 2001 2002 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    1    0    1    1    0    1    1    4    1    2    2    5
##
## 2000 2001 2002 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    1    0    1    1    0    1    1    3    1    2    2    5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.30975091964687"
## [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 = 3.5, p-value = 0.5
## 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: 2.23606797749979"

## Warning in wilcox.test.default(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"

## 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"
## [1] "Regression 3: First 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

## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 18"
## [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]"
##

```

```

## 1998 2000 2001 2002 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    3    1    2    3    1    1    2    6    3    7    8    4    3
##
## 1998 2000 2001 2002 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    0    0    2    1    1    2    6    3    7    7    3    3
##
## 1998 2000 2001 2002 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    0    0    2    1    1    2    5    3    7    7    3    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

## Warning in lmrob.S(x, y, control = control, mf = mf): find_scale() did not
## converge in 'maxit.scale' (= 200) iterations

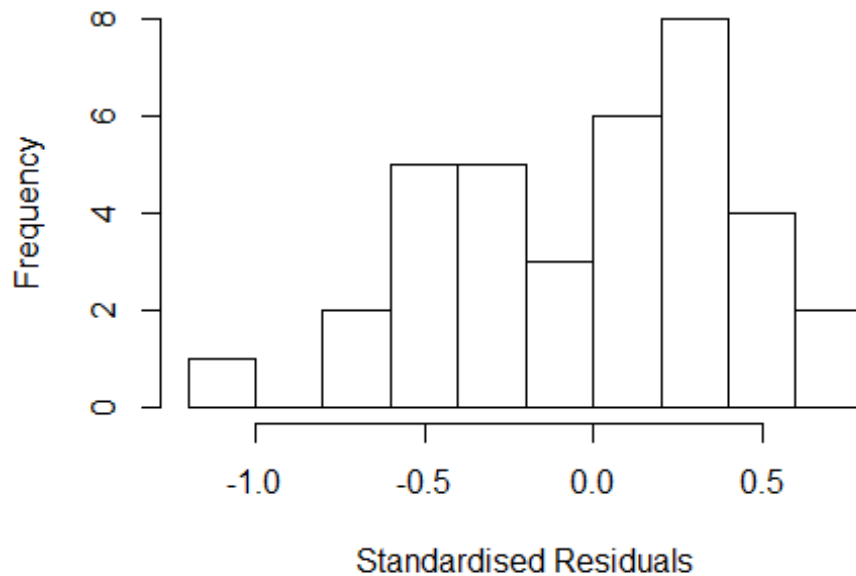
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"

## Warning in 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 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
## -0.8184 -0.2741 -0.0902  0.4192  1.1157
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.2460    0.3318   3.76  0.0011 **
## FirstAuthorFemale1  0.4399    0.3329   1.32  0.1999
## LastAuthorFemale1 -0.2729    0.2145  -1.27  0.2166
## Year2002        -0.6460    0.5874  -1.10  0.2833
## Year2004         0.8470    0.0000   Inf <2e-16 ***
## Year2005        -0.3000    0.3318  -0.90  0.3757
## Year2006         1.4615    0.4220   3.46  0.0022 **
```

```

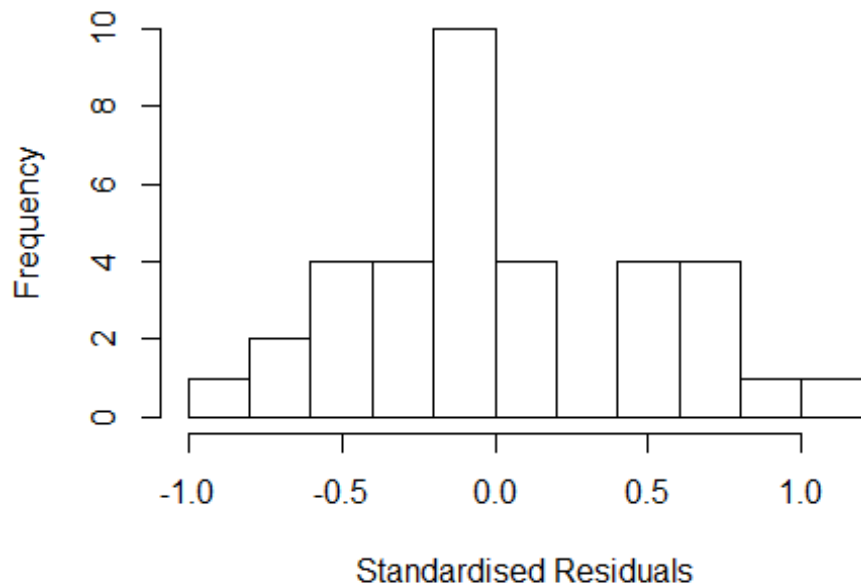
## Year2007          0.8802      0.2493      3.53      0.0019 **
## Year2008          0.7534      0.3254      2.32      0.0303 *
## Year2009          0.0959      0.0996      0.96      0.3461
## Year2010         -0.1717      0.3920     -0.44      0.6656
## Year2011         -0.0826      0.5086     -0.16      0.8725
## Year2012          0.3218      0.1986      1.62      0.1194
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.524
## Multiple R-squared:  0.563, Adjusted R-squared:  0.325
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 30 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.630  0.884  0.934  0.917  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      2.86e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          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

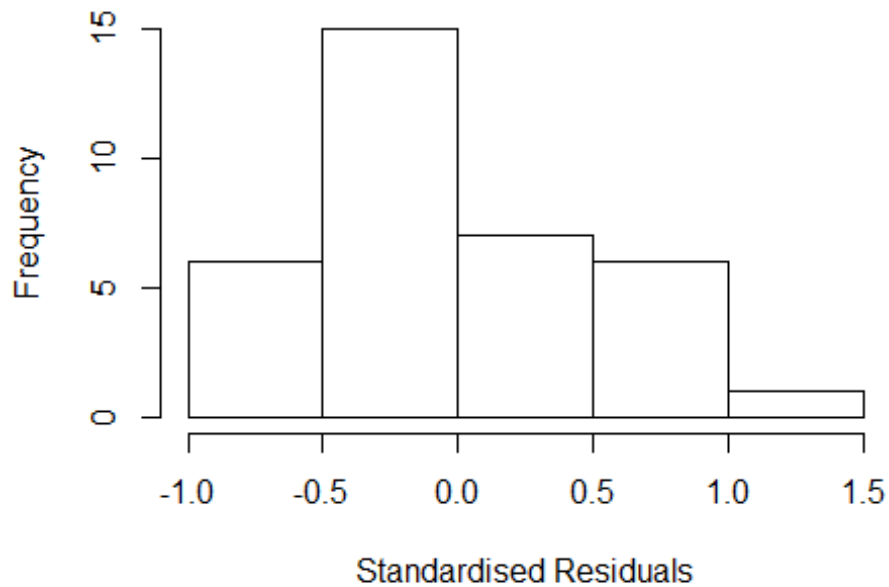
```


Residuals from first and last author



```
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale  NaN  1          NaN
## Year              NaN 10          NaN
```

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
## -0.9474 -0.2957 -0.0226 0.4771 1.0899
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1902 0.2923 4.07 0.00047 ***
## FirstAuthorFemale1 0.2228 0.2923 0.76 0.45370
## Year2002 -0.6181 0.5090 -1.21 0.23692
## Year2004 0.8470 0.0000 Inf < 2e-16 ***
## Year2005 -0.2442 0.2923 -0.84 0.41216
## Year2006 1.4894 0.4262 3.49 0.00195 **
## Year2007 0.9556 0.2155 4.43 0.00019 ***
## Year2008 0.8595 0.2435 3.53 0.00179 **
## Year2009 0.1778 0.1175 1.51 0.14382
## Year2010 -0.1459 0.3367 -0.43 0.66881
## Year2011 0.0464 0.4755 0.10 0.92310
## Year2012 0.3405 0.2059 1.65 0.11182
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.607
## Multiple R-squared: 0.544, Adjusted R-squared: 0.325
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 29 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.728 0.920 0.940 0.934 0.988 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 2.86e-03 1.82e-12
## warn.limit.reject warn.limit.meanrw
## 5.00e-01 5.00e-01
## nResample max.it best.r.s k.fast.s k.max maxit.scale
## 500 50 2 1 1000 200
## trace.lev mts compute.rd
## 0 1000 0
## psi subsampling cov
```

```

##          "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##          "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##          GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 8.963e+16 1      2.994e+08
## Year              8.963e+16 10      7.041e+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
## -0.9194 -0.2724 -0.0301  0.3220  1.1533
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.48e+00   2.74e-01  5.43e+00 1.6e-05 ***
## LastAuthorFemale1 -7.17e-02   2.74e-01 -2.60e-01  0.7955
## Year2002         -7.65e-01   9.27e-01 -8.30e-01  0.4172
## Year2004          8.47e-01   1.46e-08  5.79e+07 < 2e-16 ***
## Year2005         -5.39e-01   2.74e-01 -1.97e+00  0.0611 .
## Year2006          1.34e+00   3.04e-01  4.41e+00  0.0002 ***
## Year2007          9.23e-01   2.51e-01  3.68e+00  0.0012 **
## Year2008          8.25e-01   3.10e-01  2.67e+00  0.0138 *
## Year2009          1.18e-01   1.25e-01  9.40e-01  0.3576
## Year2010         -2.09e-01   4.61e-01 -4.50e-01  0.6539
## Year2011         -6.66e-02   5.03e-01 -1.30e-01  0.8957
## Year2012          2.47e-01   1.43e-01  1.72e+00  0.0987 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.483
## Multiple R-squared:  0.55, Adjusted R-squared:  0.335
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~ = 1. The remaining 29 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.547  0.868  0.952  0.895  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          2.86e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##          500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##          0          1000          0
##          psi          subsampling          cov
##          "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##          "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 35"
## [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]"
##
## 2002 2006 2008 2009 2010 2011 2012
##    1    2    2    1    3    2    1
##
## 2002 2006 2008 2009 2010 2011 2012
##    1    2    2    1    2    1    1
##
## 2002 2006 2008 2009 2010 2011 2012
##    1    2    2    1    2    1    1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.30192724889463"
## [1] "Male first 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 = 0, p-value = 0.3
## 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"

## Warning in wilcox.test.default(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"
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 10"
## [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]"
##
## 1999 2001 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    1    2    1    1    1    3    3    5    8    8    5
##
## 1999 2001 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    1    2    1    1    1    3    3    5    7    7    5
##
## 1999 2001 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    1    2    1    1    1    3    3    5    7    7    5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.6855783722714"
## [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 = 9, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.66091424666239"
## [1] "Male last author team size 2018 geometric mean: 2.44948974278318"

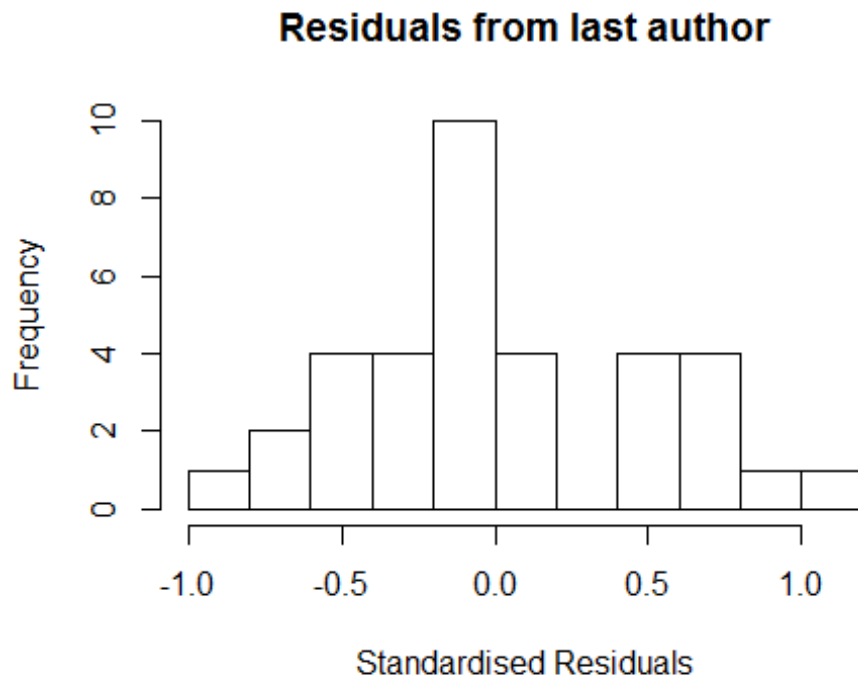
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 14, 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

## 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 3          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
## -1.11e+00 -2.80e-01  6.66e-16  2.49e-01  1.09e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.95370    0.33537   2.84  0.01004 *
## FirstAuthorFemale1 0.37759    0.50007   0.76  0.45901
## LastAuthorFemale1 0.34171    0.54094   0.63  0.53474
## UniqueAuthors2    0.85566    0.26481   3.23  0.00419 **
## UniqueAuthors3    1.86182    0.54247   3.43  0.00264 **
## UniqueAuthors5    1.29317    0.55220   2.34  0.02965 *
## Year2001         -0.01450    0.00672  -2.16  0.04322 *
## Year2003         -0.95395    0.48883  -1.95  0.06514 .
## Year2005         -1.34466    0.26481  -5.08  5.7e-05 ***
## Year2006         -1.67300    0.00000   -Inf < 2e-16 ***
## Year2007         -0.57210    0.30510  -1.88  0.07545 .
## Year2008         -0.65863    0.64161  -1.03  0.31690
## Year2009         -1.98486    0.44218  -4.49  0.00022 ***
## Year2010         -1.10174    0.23114  -4.77  0.00012 ***
## Year2011         -0.56763    0.27664  -2.05  0.05350 .
## Year2012         -1.23696    0.37864  -3.27  0.00386 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.658
## Multiple R-squared:  0.57,   Adjusted R-squared:  0.248
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 28 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.759  0.901  0.974  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          2.78e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500            50          2            1            1000      200
##      trace.lev      mts      compute.rd
##      0              1000          0

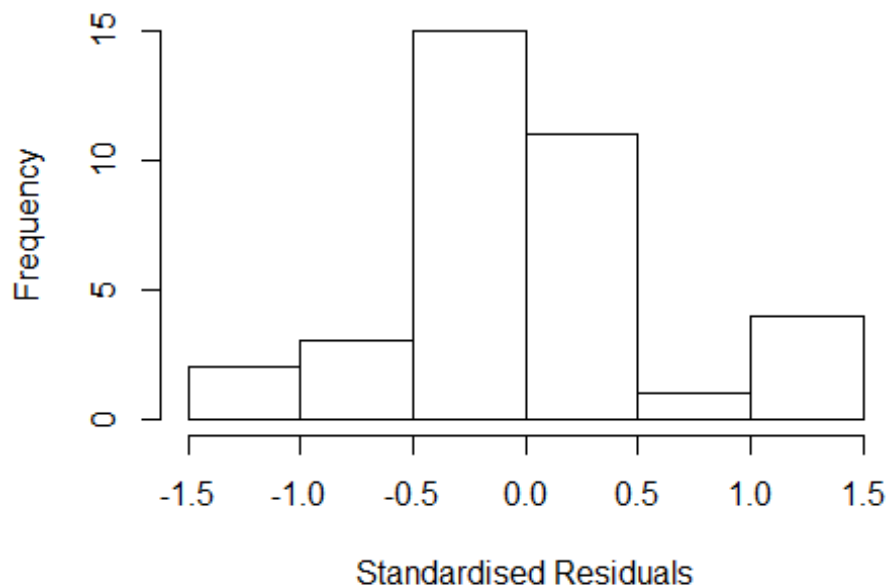
```

```
##           psi           subsampling           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

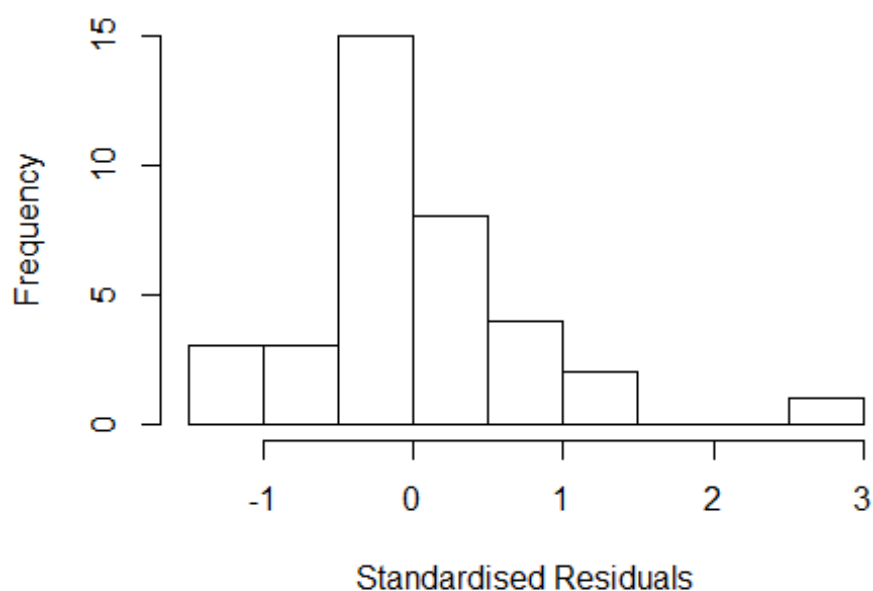
## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

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
```


Residuals from first and last author



```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 36 78651371342 2.586 2011      2913      1      2.534
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.35e+00 -2.43e-01  6.94e-16  2.69e-01  2.53e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.00695    0.26324   3.83  0.00087 ***
## FirstAuthorFemale1  1.29303    0.62540   2.07  0.05012 .
## LastAuthorFemale1 -0.62698    0.62551  -1.00  0.32660
## Year2001        -0.01450    0.00672  -2.16  0.04157 *
## Year2003        -1.06698    0.62551  -1.71  0.10152
## Year2005        -0.48900    0.00000   -Inf < 2e-16 ***
## Year2006        -1.67300    0.00000   -Inf < 2e-16 ***
## Year2007        -0.62831    0.28171  -2.23  0.03577 *
## Year2008        -0.65759    0.61324  -1.07  0.29470
## Year2009        -0.64307    0.30901  -2.08  0.04875 *
## Year2010        -0.89353    0.30351  -2.94  0.00728 **
## Year2011        -0.32764    0.37203  -0.88  0.38761
```

```

## Year2012          -0.94251    0.50239    -1.88    0.07340 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.709
## Multiple R-squared:  0.403, Adjusted R-squared:  0.0914
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 27 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.175  0.862  0.965   0.898   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           2.78e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01             5.00e-01
##   nResample      max.it    best.r.s    k.fast.s      k.max maxit.scale
##           500          50         2         1         1000         200
##   trace.lev      mts    compute.rd
##           0          1000         0
##           psi             subsampling             cov
##           "bisquare"      "nonsingular"             ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.949  1           1.717
## Year              2.949 10           1.056

## [1] "List of 1 outliers with residuals above 2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 36 78651371342 2.586 2011      2913      1      2.534
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##   control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.48944 -0.34435 -0.00475  0.34506  1.94877
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    8.21e-01   3.91e-01  2.10e+00  0.0462 *
## FirstAuthorFemale1 8.52e-01   3.91e-01  2.18e+00  0.0391 *

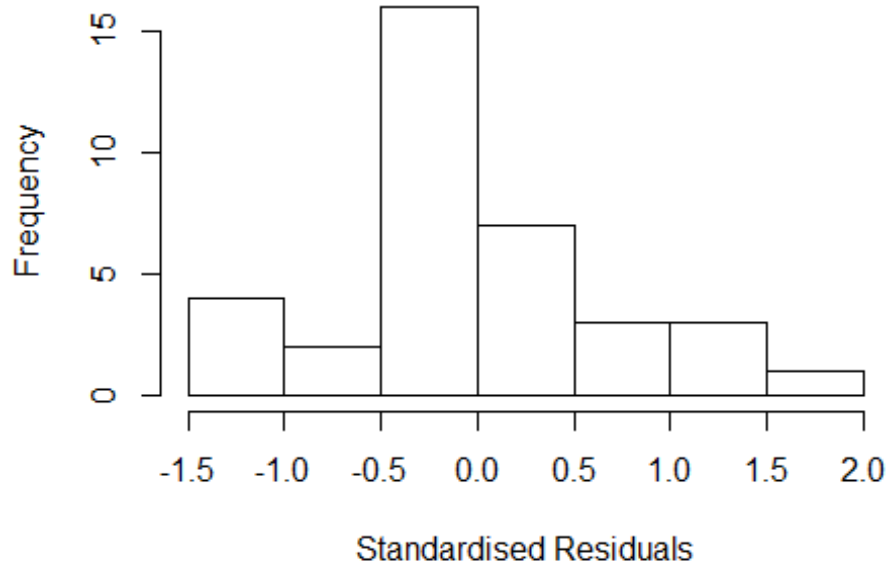
```

```

## Year2001      -1.45e-02   6.72e-03 -2.16e+00   0.0411 *
## Year2003      -4.40e-01   4.09e-08 -1.07e+07   <2e-16 ***
## Year2005      -4.89e-01   3.86e-08 -1.27e+07   <2e-16 ***
## Year2006      -1.67e+00   3.86e-08 -4.33e+07   <2e-16 ***
## Year2007      -3.58e-01   3.37e-01 -1.06e+00   0.2985
## Year2008      -6.56e-01   5.67e-01 -1.16e+00   0.2585
## Year2009      -6.59e-01   3.77e-01 -1.75e+00   0.0935 .
## Year2010      -7.69e-01   2.38e-01 -3.24e+00   0.0035 **
## Year2011      -1.84e-01   3.48e-01 -5.30e-01   0.6028
## Year2012      -5.12e-01   3.78e-01 -1.36e+00   0.1879
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.839
## Multiple R-squared:  0.282, Adjusted R-squared:  -0.0473
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  11 weights are ~= 1. The remaining 25 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.569  0.871  0.962   0.917   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      2.78e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##  nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##  trace.lev      mts      compute.rd
##      0          1000          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

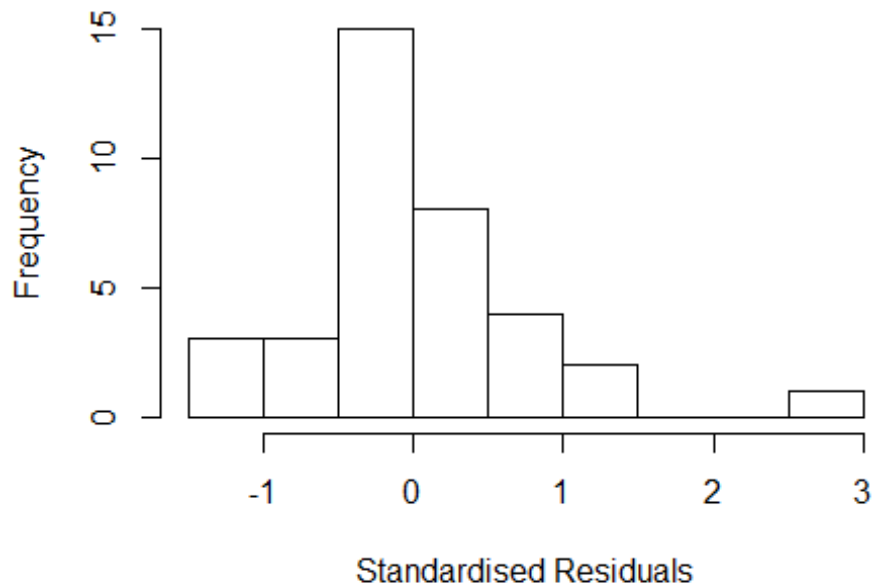
```

Residuals from first author



| | | | |
|---------------------|------|----|-------------------------|
| ## | GVIF | Df | $GVIF^{1/(2 \cdot Df)}$ |
| ## LastAuthorFemale | NaN | 1 | NaN |
| ## Year | NaN | 10 | NaN |

Residuals from last author



```

## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 36 78651371342 2.586 2011      2913      1      2.534
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.37e+00 -4.13e-01  7.22e-16  5.04e-01  1.42e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.72e+00   5.30e-01  3.24e+00  0.0035 **
## LastAuthorFemale1 -4.40e-02   5.30e-01 -8.00e-02  0.9346
## Year2001        -1.45e-02   6.72e-03 -2.16e+00  0.0411 *
## Year2003        -4.84e-01   5.30e-01 -9.10e-01  0.3702
## Year2005        -4.89e-01   1.08e-08 -4.54e+07  <2e-16 ***
## Year2006        -1.67e+00   0.00e+00    -Inf  <2e-16 ***
## Year2007        -6.82e-01   4.54e-01 -1.50e+00  0.1459
## Year2008        -6.55e-01   5.34e-01 -1.23e+00  0.2317
## Year2009        -8.41e-01   3.83e-01 -2.19e+00  0.0381 *
## Year2010        -9.00e-01   3.06e-01 -2.94e+00  0.0072 **
## Year2011        -3.43e-01   3.76e-01 -9.10e-01  0.3711
## Year2012        -7.32e-01   6.18e-01 -1.18e+00  0.2478
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 1.04
## Multiple R-squared:  0.144, Adjusted R-squared:  -0.248
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 26 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.838  0.913  0.944  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      2.78e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
##      nResample  max.it  best.r.s  k.fast.s  k.max maxit.scale
##      500      50      2      1      1000      200
##      trace.lev      mts  compute.rd
##      0      1000      0
##      psi      subsampling      cov

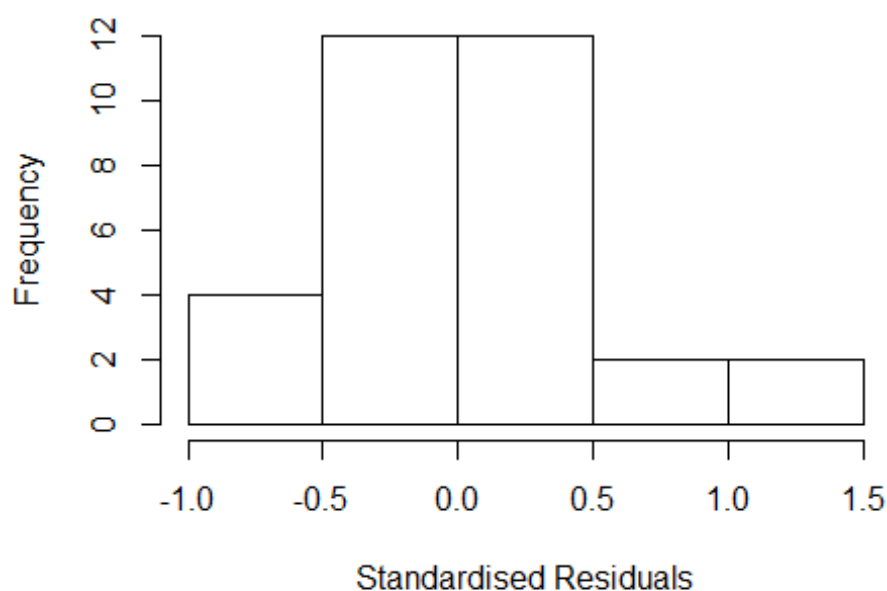
```

```

##          "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##          "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 36"
## [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]"
##
## 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2    1    1    3    2    5   10    5    4    5
##
## 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2    0    1    3    2    4   10    3    3    5
##
## 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2    0    1    3    2    4    9    3    3    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: 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 10.31 1          3.211
## LastAuthorFemale   7.32 1          2.705
## Year               11.77 8          1.167

```

Residuals from first and last author



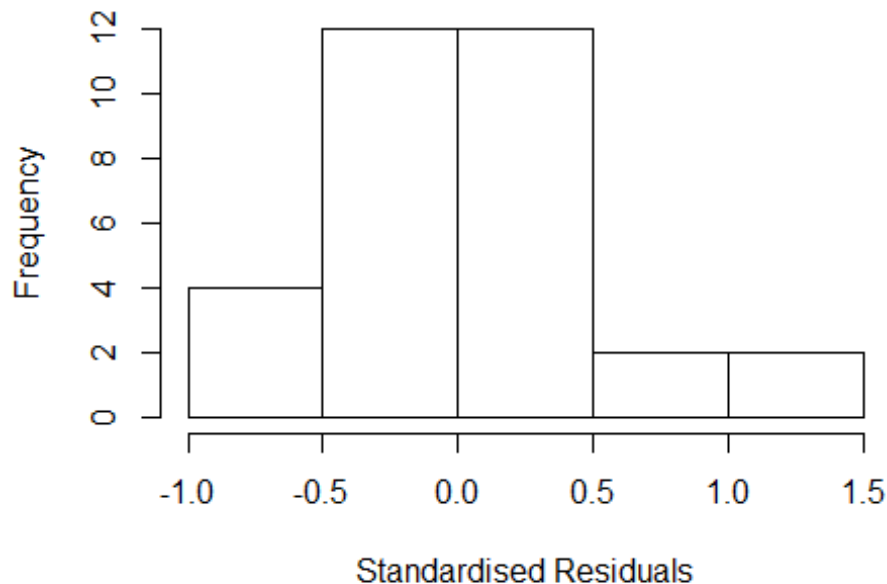
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.94476 -0.31927 0.00502 0.30132 1.18924
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9396 0.3127 3.00 0.0068 **
## FirstAuthorFemale1 0.3876 0.2553 1.52 0.1439
## LastAuthorFemale1 -0.0123 0.1827 -0.07 0.9467
## Year2005 -1.3148 0.2553 -5.15 4.2e-05 ***
## Year2006 -0.3701 0.7808 -0.47 0.6404
## Year2007 0.0025 0.2557 0.01 0.9923
## Year2008 -0.2639 0.2830 -0.93 0.3617
## Year2009 -0.5760 0.3568 -1.61 0.1213
## Year2010 -1.0236 0.2748 -3.72 0.0013 **
## Year2011 -0.8891 0.6430 -1.38 0.1813
## Year2012 -0.7748 0.3118 -2.48 0.0215 *
## ---
```

```

## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.612
## Multiple R-squared:  0.355, Adjusted R-squared:  0.0478
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 28 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.686  0.926  0.975  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      3.13e-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 4.836 1          2.199
## Year              4.836 8          1.104

```


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
## -0.93184 -0.31545  0.00459  0.29914  1.20216
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.9427     0.3025   3.12 0.00503 **
## FirstAuthorFemale1 0.3783     0.1908   1.98 0.05996 .
## Year2005        -1.3210     0.2348  -5.63 1.2e-05 ***
## Year2006        -0.3892     0.8095  -0.48 0.63544
## Year2007         0.0025     0.2507   0.01 0.99213
## Year2008        -0.2692     0.2673  -1.01 0.32494
## Year2009        -0.5827     0.3427  -1.70 0.10320
## Year2010        -1.0256     0.2646  -3.88 0.00082 ***
## Year2011        -0.9145     0.6240  -1.47 0.15696
## Year2012        -0.7770     0.3060  -2.54 0.01869 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Robust residual standard error: 0.58
## Multiple R-squared: 0.361, Adjusted R-squared: 0.0996
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 27 ones are summarized as
##   Min. 1st Qu. Median   Mean 3rd Qu.   Max.
## 0.646 0.911 0.971 0.927 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          3.13e-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 21.37 1          4.623
## Year          21.37 8          1.211

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -0.92180 -0.28088 -0.00747 0.19777 1.21256
##
## Coefficients:
##      Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2190     0.1672    7.29 2.7e-07 ***
## LastAuthorFemale1 0.2040     0.1304    1.56 0.1319
## Year2005          -1.4230     0.1672   -8.51 2.1e-08 ***
## Year2006          -0.5012     0.8324   -0.60 0.5532
## Year2007           0.0025     0.1547    0.02 0.9873
```

```

## Year2008          -0.4165      0.1927    -2.16    0.0418 *
## Year2009          -0.7787      0.2465    -3.16    0.0045 **
## Year2010          -1.0610      0.1694    -6.26    2.6e-06 ***
## Year2011          -1.0305      0.6435    -1.60    0.1236
## Year2012          -0.9191      0.2460    -3.74    0.0011 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.56
## Multiple R-squared:  0.34,   Adjusted R-squared:  0.0699
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 23 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.619  0.879  0.956  0.904  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      3.13e-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: 32"
## [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
##    4    6   13   17   12    9   12    5    8   12   14   11   11   18   17
## 2011 2012
##   31   27
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    3    7   10    3    1    5    3    1    9   10    7    5   11   13
## 2011 2012
##   15   14

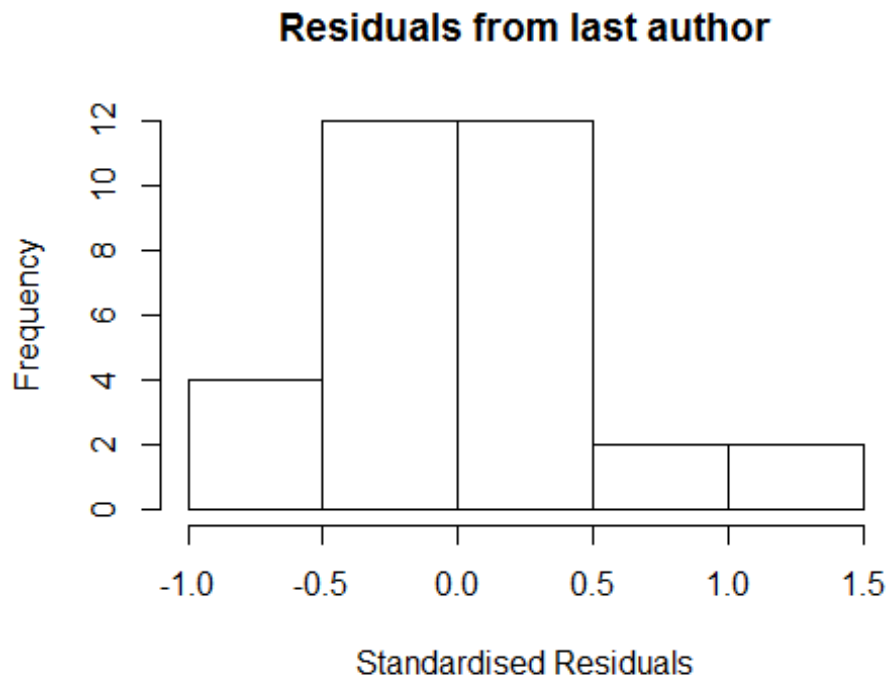
```

```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    3    6    8    3    1    4    1    1    8    8    7    5    9    9
## 2011 2012
##   14   12
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 5.39382720425877"
## [1] "Male first author team size 2018 geometric mean: 4.55901411390956"

## Warning in wilcox.test.default(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 = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.96571618608088"
## [1] "Male last author team size 2018 geometric mean: 4.66689988970022"

## Warning in wilcox.test.default(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 = 0.2
## 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 6.714e-01 1      8.194e-01
## LastAuthorFemale -3.810e+15 1      NaN
## UniqueAuthors    8.464e+43 4      3.097e+05
## Year              6.534e+44 16     2.515e+01

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##      Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##      k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.14e+00 -2.07e-01  6.11e-16  1.90e-01  1.40e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -0.1000     0.1508   -0.66  0.50906
## FirstAuthorFemale1  0.0699     0.1400    0.50  0.61871
## LastAuthorFemale1  0.0301     0.1118    0.27  0.78861
## UniqueAuthors2    0.3658     0.2105    1.74  0.08623 .
## UniqueAuthors3    0.4862     0.2174    2.24  0.02826 *
## UniqueAuthors4    0.3071     0.2842    1.08  0.28330
## UniqueAuthors5    0.4701     0.2284    2.06  0.04294 *
## Year1997          0.5683     0.1500    3.79  0.00030 ***
## Year1998          1.0116     0.3093    3.27  0.00161 **
## Year1999          0.7597     0.2826    2.69  0.00880 **
## Year2000          0.7554     0.2590    2.92  0.00464 **
## Year2001          1.3832     0.2428    5.70  2.1e-07 ***
## Year2002          0.4056     0.3156    1.29  0.20258
## Year2003          0.5629     0.2411    2.33  0.02219 *
## Year2004          1.2909     0.2842    4.54  2.0e-05 ***
## Year2005          0.8701     0.2229    3.90  0.00020 ***
## Year2006          0.9130     0.2888    3.16  0.00225 **
## Year2007          0.8640     0.2059    4.20  7.2e-05 ***
```

```

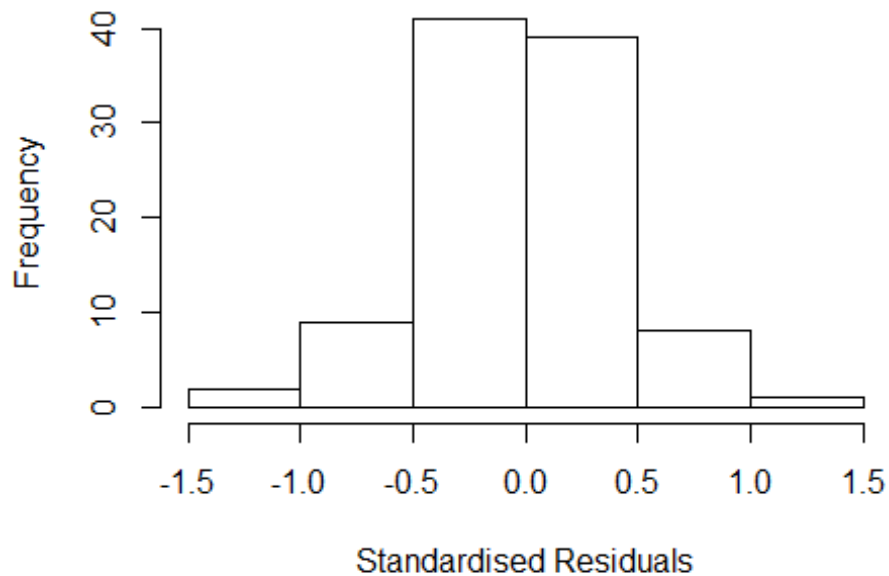
## Year2008          0.9052      0.2455      3.69  0.00042 ***
## Year2009          0.8310      0.2825      2.94  0.00431 **
## Year2010          1.0884      0.2709      4.02  0.00014 ***
## Year2011          0.9461      0.2419      3.91  0.00020 ***
## Year2012          1.0459      0.4350      2.40  0.01859 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.351
## Multiple R-squared:  0.35,   Adjusted R-squared:  0.164
## Convergence in 28 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 82 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0744 0.8400 0.9510 0.8770 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      1.00e-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"

## 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

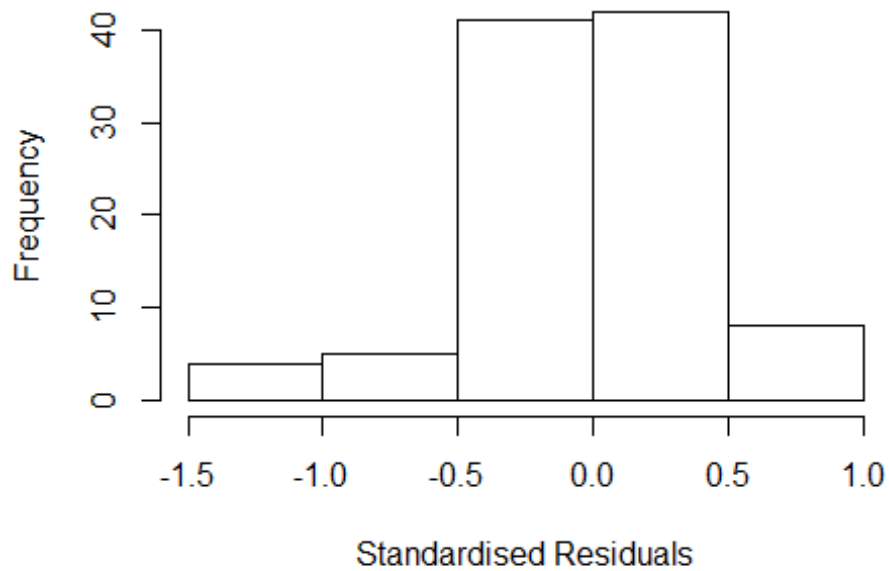
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.37156 -0.22280  0.00093  0.18809  0.99785
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -0.2033    0.1604   -1.27   0.2088
## FirstAuthorFemale1  0.1692    0.0972    1.74   0.0856 .
## LastAuthorFemale1  0.0341    0.1201    0.28   0.7774
## Year1997         0.8649    0.1791    4.83 6.4e-06 ***
## Year1998         1.5449    0.2659    5.81 1.2e-07 ***
## Year1999         1.2674    0.1664    7.62 4.3e-11 ***
## Year2000         1.0406    0.3427    3.04  0.0032 **
```

```

## Year2001      1.7531      0.1201      14.60 < 2e-16 ***
## Year2002      0.9173      0.2113      4.34 4.1e-05 ***
## Year2003      1.0531      0.1201      8.77 2.3e-13 ***
## Year2004      1.5980      0.0000      Inf < 2e-16 ***
## Year2005      1.2982      0.1097     11.83 < 2e-16 ***
## Year2006      1.4187      0.1687      8.41 1.2e-12 ***
## Year2007      1.2138      0.1982      6.12 3.1e-08 ***
## Year2008      1.3172      0.0845     15.58 < 2e-16 ***
## Year2009      1.2832      0.1700      7.55 5.8e-11 ***
## Year2010      1.5070      0.1606      9.38 1.4e-14 ***
## Year2011      1.3763      0.1362     10.10 5.4e-16 ***
## Year2012      1.3716      0.2853      4.81 7.0e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.367
## Multiple R-squared:  0.268, Adjusted R-squared:  0.106
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 84 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.131  0.861  0.953   0.881  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"
## 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

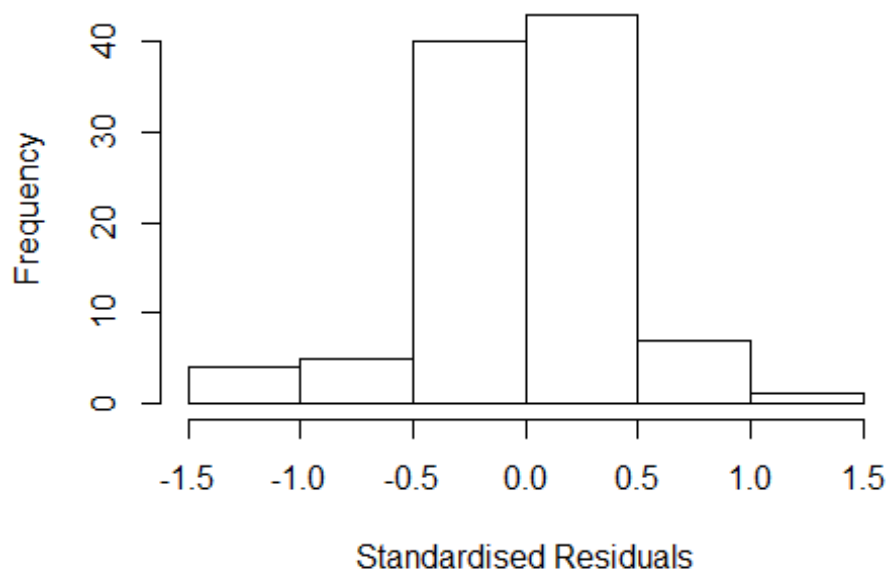
```


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
## -1.35635 -0.23447 0.00686 0.18799 1.00564
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.1778 0.1020 -1.74 0.0850 .
## FirstAuthorFemale1 0.1778 0.1020 1.74 0.0850 .
## Year1997 0.8652 0.1794 4.82 6.5e-06 ***
## Year1998 1.5137 0.1965 7.70 2.7e-11 ***
## Year1999 1.2341 0.1220 10.11 4.5e-16 ***
## Year2000 1.0308 0.3342 3.08 0.0028 **
## Year2001 1.7190 0.0000 Inf < 2e-16 ***
## Year2002 0.8999 0.2030 4.43 2.8e-05 ***
## Year2003 1.0190 0.0000 Inf < 2e-16 ***
## Year2004 1.5980 0.0000 Inf < 2e-16 ***
## Year2005 1.2783 0.0711 17.98 < 2e-16 ***
## Year2006 1.4012 0.1601 8.75 2.2e-13 ***
## Year2007 1.1923 0.1642 7.26 2.0e-10 ***
## Year2008 1.2984 0.0361 35.99 < 2e-16 ***
## Year2009 1.2629 0.1541 8.20 2.8e-12 ***
## Year2010 1.4796 0.1069 13.84 < 2e-16 ***
## Year2011 1.3665 0.1256 10.88 < 2e-16 ***
## Year2012 1.3564 0.2544 5.33 8.4e-07 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.363
## Multiple R-squared: 0.268, Adjusted R-squared: 0.117
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~ 1. The remaining 84 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.133 0.854 0.953 0.879 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 1.00e-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 2.096e+15 1 4.579e+07
## Year 2.096e+15 16 3.012e+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
## -1.29051 -0.22925  0.00526  0.24737  0.86710
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)  -6.03e-02  1.23e-01 -4.90e-01  0.62682
## LastAuthorFemale1  6.03e-02  1.23e-01  4.90e-01  0.62682
## Year1997      8.64e-01  1.77e-01  4.87e+00  5.4e-06 ***
## Year1998      1.45e+00  2.44e-01  5.94e+00  6.8e-08 ***
## Year1999      1.26e+00  1.72e-01  7.29e+00  1.7e-10 ***
## Year2000      1.04e+00  3.50e-01  2.99e+00  0.00370 **
## Year2001      1.78e+00  1.23e-01  1.44e+01 < 2e-16 ***
## Year2002      8.21e-01  2.02e-01  4.06e+00  0.00011 ***
## Year2003      1.08e+00  1.23e-01  8.74e+00  2.4e-13 ***
## Year2004      1.60e+00  1.63e-08  9.80e+07 < 2e-16 ***
## Year2005      1.27e+00  1.08e-01  1.18e+01 < 2e-16 ***
## Year2006      1.37e+00  1.73e-01  7.95e+00  9.0e-12 ***
## Year2007      1.20e+00  1.95e-01  6.16e+00  2.6e-08 ***
## Year2008      1.30e+00  8.75e-02  1.48e+01 < 2e-16 ***
## Year2009      1.30e+00  1.70e-01  7.67e+00  3.2e-11 ***
## Year2010      1.50e+00  1.70e-01  8.80e+00  1.8e-13 ***
## Year2011      1.32e+00  1.23e-01  1.07e+01 < 2e-16 ***
## Year2012      1.29e+00  2.92e-01  4.42e+00  2.9e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

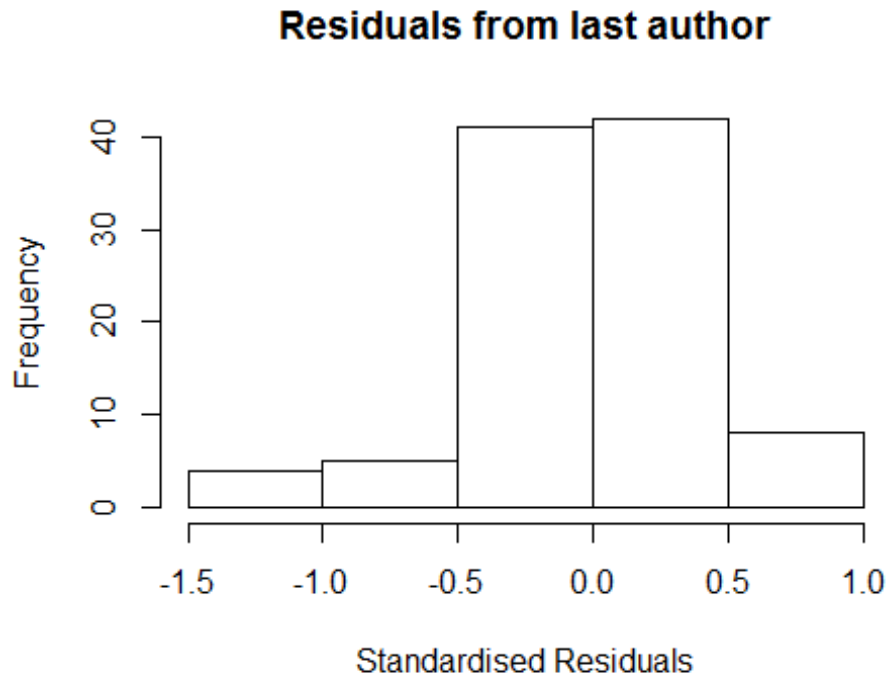
```

##
## Robust residual standard error: 0.386
## Multiple R-squared: 0.225, Adjusted R-squared: 0.0644
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 84 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.240  0.850  0.949  0.890  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           1.00e-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 2917"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 2002 2005 2006 2008 2009 2010 2011 2012
##    1    1    2    3    4    8    6    6
##
## 2002 2005 2006 2008 2009 2010 2011 2012
##    0    1    2    2    3    8    6    6
##
## 2002 2005 2006 2008 2009 2010 2011 2012
##    0    1    2    2    3    8    6    6
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.30192724889463"
## [1] "Male first author team size 2018 geometric mean: NaN"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"

```

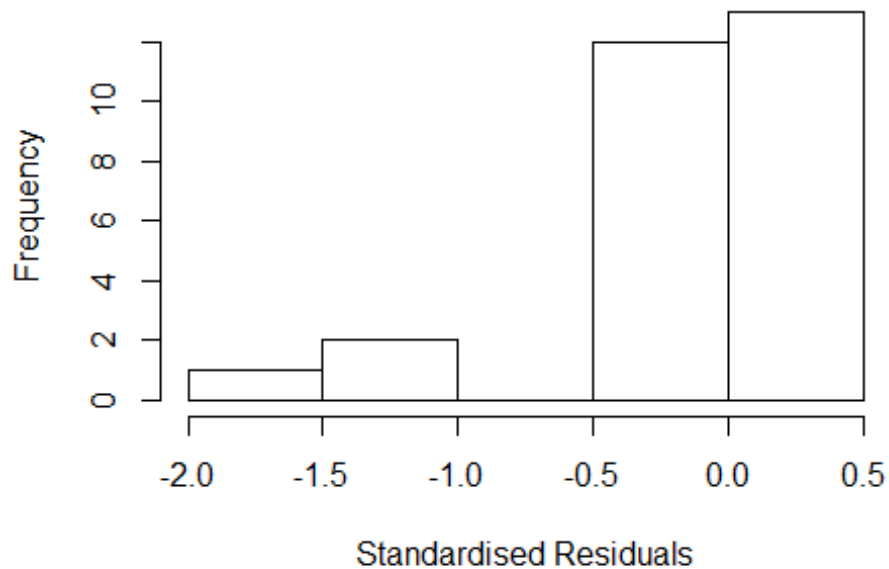
```
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

## 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 | 6 | NaN |

Residuals from first and last author and team size



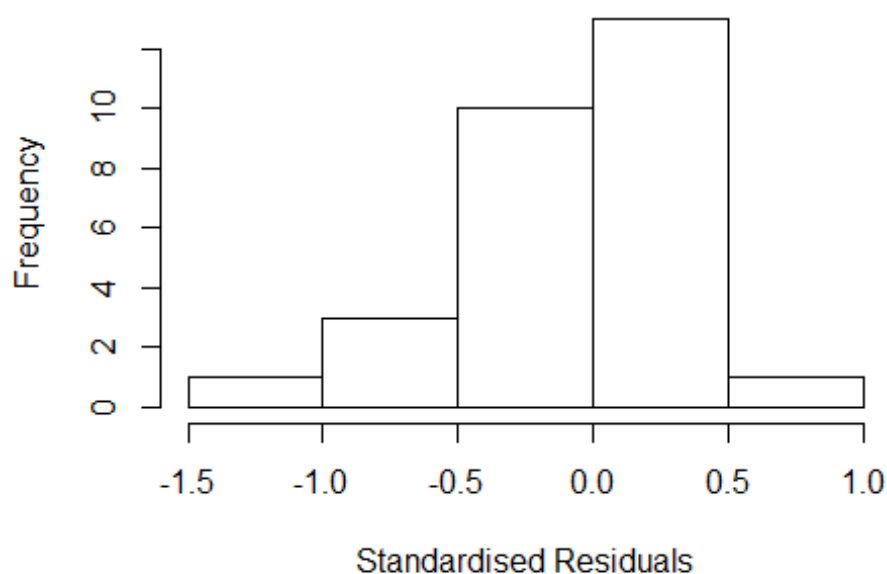
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.53e+00 -9.16e-02 -1.94e-16 3.79e-02 2.27e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8536 0.1834 4.65 0.00031 ***
## FirstAuthorFemale1 0.0104 0.0760 0.14 0.89341
## LastAuthorFemale1 0.8488 0.0658 12.90 1.6e-09 ***
## UniqueAuthors2 0.6859 0.0964 7.11 3.6e-06 ***
## UniqueAuthors3 0.4983 0.0962 5.18 0.00011 ***
## UniqueAuthors4 -0.3087 0.1382 -2.23 0.04118 *
## UniqueAuthors5 0.5530 0.1276 4.33 0.00059 ***
## Year2006 -0.5947 0.1704 -3.49 0.00329 **
## Year2008 -0.1367 0.1591 -0.86 0.40401
## Year2009 -0.4415 0.1702 -2.59 0.02034 *
```

```

## Year2010          -1.2839      0.1224   -10.49   2.7e-08 ***
## Year2011          -0.8650      0.0000     -Inf   < 2e-16 ***
## Year2012          -0.5879      0.1381    -4.26   0.00069 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.234
## Multiple R-squared:  0.935, Adjusted R-squared:  0.883
## Convergence in 9 IRWLS iterations
##
## Robustness weights:
## 3 observations c(19,20,24) are outliers with |weight| = 0 ( < 0.0036);
## 7 weights are ~= 1. The remaining 18 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.886  0.960  0.988  0.971  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 2: First author gender, Last author gender, Year as
factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 376341 1          613.467
## LastAuthorFemale  417567 1          646.194
## Year              455164 6          2.962

```

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
## -1.34e+00 -1.55e-01 9.88e-05 1.61e-01 5.46e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.4866 0.1025 4.75 0.00014 ***
## FirstAuthorFemale1 0.0742 0.1714 0.43 0.66971
## LastAuthorFemale1 0.8431 0.1805 4.67 0.00017 ***
## Year2006 0.3971 0.1064 3.73 0.00142 **
## Year2008 0.2012 0.0512 3.93 0.00090 ***
## Year2009 -0.0937 0.0919 -1.02 0.32073
## Year2010 -0.4178 0.0989 -4.22 0.00046 ***
## Year2011 -0.4106 1.3269 -0.31 0.76034
## Year2012 -0.0689 0.2086 -0.33 0.74494
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```



```

## Robust residual standard error: 0.326
## Multiple R-squared: 0.585, Adjusted R-squared: 0.411
## Convergence in 48 IRWLS iterations
##
## Robustness weights:
## 3 weights are ~= 1. The remaining 25 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0563 0.8310 0.9700 0.8510 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           3.57e-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 13.31 1           3.649
## Year              13.31 6           1.241

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.9521 -0.3012  0.0405  0.2760  0.7191
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.88021    0.27728   3.17   0.0048 **
## FirstAuthorFemale1 0.52379    0.27728   1.89   0.0735 .
## Year2006        -0.02450    0.44043  -0.06   0.9562
## Year2008         0.00439    0.19931   0.02   0.9826
## Year2009        -0.35678    0.19992  -1.78   0.0895 .
## Year2010        -0.45191    0.18034  -2.51   0.0210 *

```

```

## Year2011          -0.58413    0.39529   -1.48    0.1551
## Year2012          -0.51798    0.28145   -1.84    0.0806 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.494
## Multiple R-squared:  0.247, Adjusted R-squared:  -0.0159
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 2 weights are ~= 1. The remaining 26 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.690  0.874  0.956  0.914  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           3.57e-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 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: 28"
## [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]"
##
## 2002 2003 2004 2006 2007 2009 2010 2011 2012
##    1    1    1    2    2   13    5    4    8
##
## 2002 2003 2004 2006 2007 2009 2010 2011 2012
##    1    1    0    2    2   10    5    3    7
##
## 2002 2003 2004 2006 2007 2009 2010 2011 2012

```

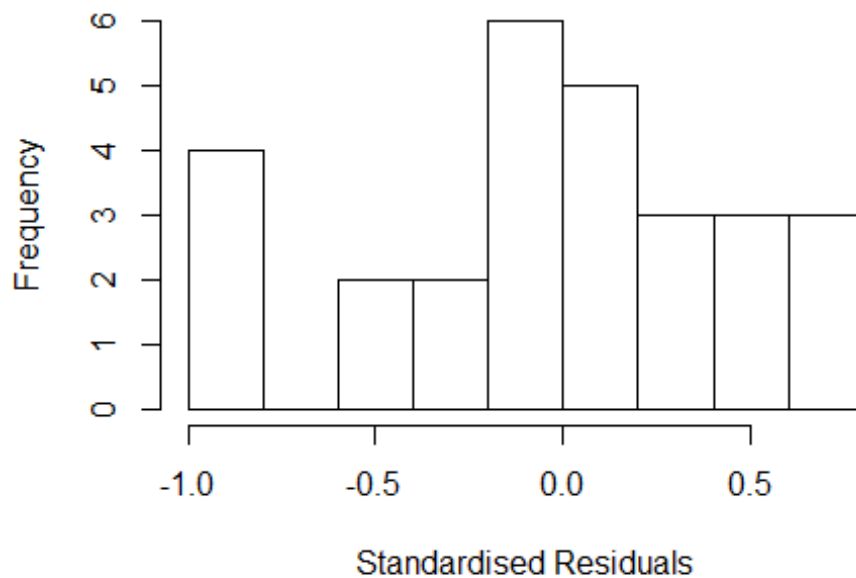
```
##      1      1      0      2      1     10      4      3      7
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.44948974278318"
## [1] "Male first author team size 2018 geometric mean: 2.26793315526605"

## Warning in wilcox.test.default(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: 2.7123532530899"
## [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
```

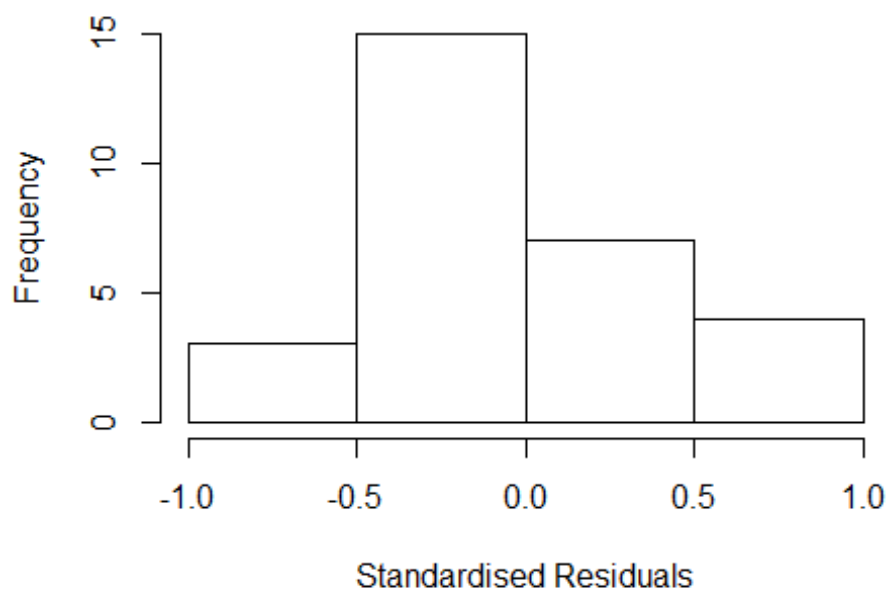
Residuals from first author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 18, p-value = 0.6
```

```
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.932e+00 1      1.390e+00
## LastAuthorFemale 2.390e+14 1      1.546e+07
## UniqueAuthors    4.593e+15 3      4.077e+02
## Year             9.589e+15 7      1.385e+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
## -0.789 -0.266 -0.036  0.252  0.986
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
```

```

## (Intercept)          2.536      0.431      5.88 2.3e-05 ***
## FirstAuthorFemale1   -0.484      0.235     -2.06 0.0558 .
## LastAuthorFemale1    0.527      0.272      1.94 0.0704 .
## UniqueAuthors2       -0.169      0.263     -0.64 0.5294
## UniqueAuthors3        0.993      0.316      3.14 0.0064 **
## UniqueAuthors5        0.400      0.135      2.96 0.0091 **
## Year2003             -2.579      0.298     -8.67 1.9e-07 ***
## Year2006             -0.898      0.915     -0.98 0.3409
## Year2007             -0.989      0.298     -3.32 0.0043 **
## Year2009             -1.617      0.272     -5.94 2.1e-05 ***
## Year2010             -2.098      0.301     -6.96 3.2e-06 ***
## Year2011             -0.984      0.426     -2.31 0.0348 *
## Year2012             -0.958      0.387     -2.47 0.0250 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.591
## Multiple R-squared:  0.544, Adjusted R-squared:  0.201
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 23 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.763  0.905  0.960  0.934  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          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 2: First author gender, Last author gender, Year as
factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.494e+00 1          1.869e+00
## LastAuthorFemale  1.699e+14 1          1.304e+07
## Year              2.777e+14 7          1.076e+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)

```

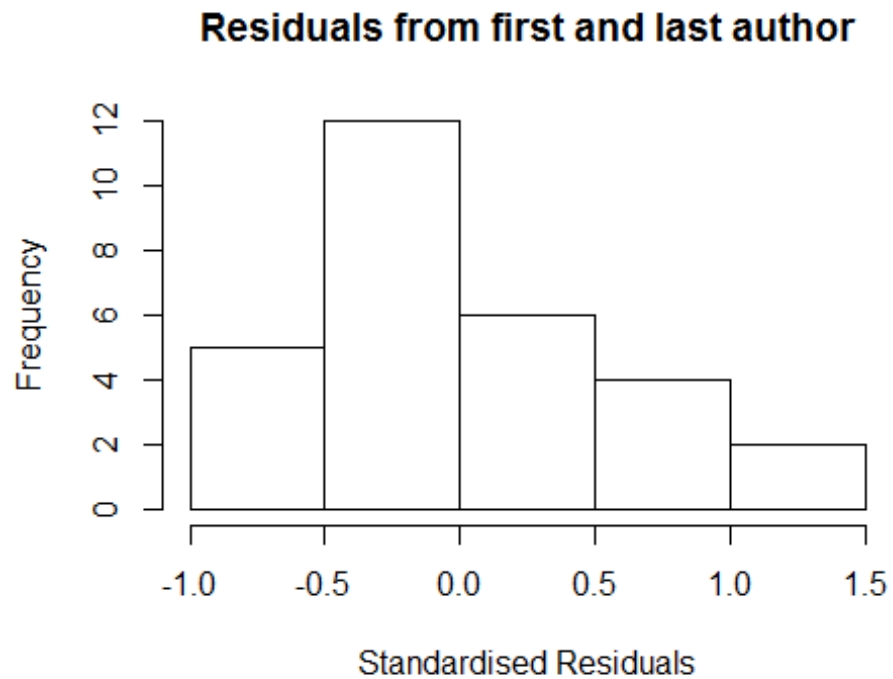
```

##
## 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
## -0.986 -0.260  0.000  0.262  1.046
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.196      0.341    6.44 3.6e-06 ***
## FirstAuthorFemale1 -0.313      0.341   -0.92 0.37037
## LastAuthorFemale1  0.642      0.389    1.65 0.11480
## Year2003          -2.525      0.389   -6.50 3.2e-06 ***
## Year2006          -0.844      0.927   -0.91 0.37386
## Year2007          -0.935      0.389   -2.41 0.02644 *
## Year2009          -1.293      0.422   -3.06 0.00641 **
## Year2010          -1.944      0.463   -4.20 0.00049 ***
## Year2011          -0.961      0.414   -2.32 0.03166 *
## Year2012          -0.877      0.415   -2.11 0.04807 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.61
## Multiple R-squared:  0.411, Adjusted R-squared:  0.131
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 24 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.750  0.866  0.955  0.923  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      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"

```

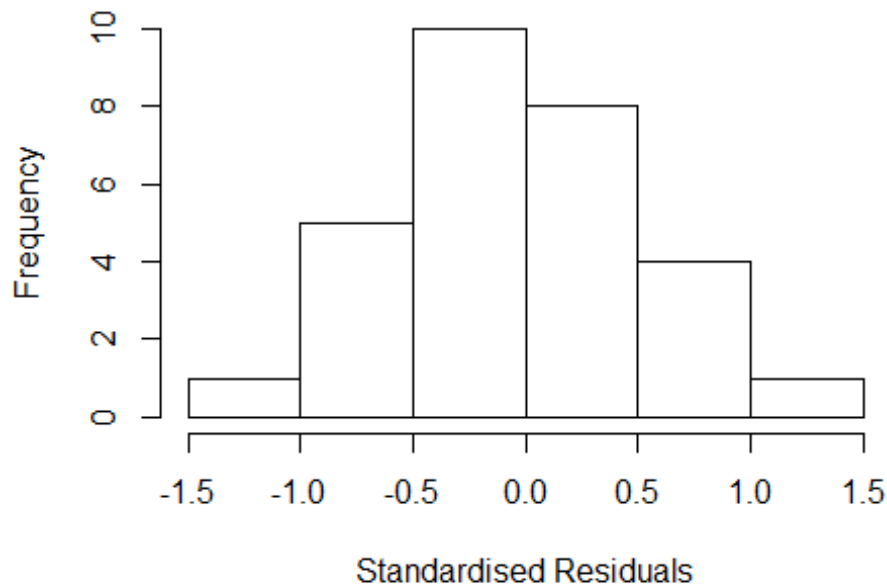
```
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

## 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  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
## -1.17e+00 -3.03e-01 2.22e-16 4.25e-01 1.11e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.8075 0.2858 6.33 3.6e-06 ***
## FirstAuthorFemale1 0.0755 0.2858 0.26 0.794
## Year2003 -1.8830 0.0000 -Inf < 2e-16 ***
## Year2006 -0.2020 0.7079 -0.29 0.778
## Year2007 -0.2930 0.0000 -Inf < 2e-16 ***
## Year2009 -0.7174 0.2549 -2.81 0.011 *
## Year2010 -1.2056 0.1690 -7.14 6.5e-07 ***
## Year2011 -0.7158 0.3460 -2.07 0.052 .
## Year2012 -0.3094 0.2354 -1.31 0.204
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```



```

## Robust residual standard error: 0.746
## Multiple R-squared: 0.322, Adjusted R-squared: 0.0512
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 24 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.790 0.901 0.966 0.941 0.987 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 4: Last author gender, Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale -2.924e+14 1 NaN
## Year -2.924e+14 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
## -1.02e+00 -2.85e-01 -3.11e-15  3.48e-01  1.07e+00
##
## Coefficients:
##           Estimate Std. Error  t value Pr(>|t|)
## (Intercept)  1.88e+00  1.87e-08  1.00e+08 < 2e-16 ***
## LastAuthorFemale1 4.44e-01  3.17e-01  1.40e+00 0.17641
## Year2003 -2.33e+00  3.17e-01 -7.35e+00 4.2e-07 ***
## Year2006 -6.46e-01  8.66e-01 -7.50e-01 0.46464
## Year2007 -7.37e-01  3.17e-01 -2.33e+00 0.03058 *
## Year2009 -1.04e+00  3.45e-01 -3.03e+00 0.00658 **

```

```

## Year2010          -1.67e+00   3.52e-01 -4.73e+00   0.00013 ***
## Year2011          -8.81e-01   3.69e-01 -2.39e+00   0.02681 *
## Year2012          -6.45e-01   3.33e-01 -1.94e+00   0.06684 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.633
## Multiple R-squared:  0.389, Adjusted R-squared:  0.145
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 23 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.759  0.864  0.964  0.923  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           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 2920"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 2001 2008 2011
##    1    1    1
##
## 2001 2008 2011
##    0    1    1
##
## 2001 2008 2011
##    0    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: 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]"
##
## 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2    2    2    3    5    4    9    8   18   14   15   16
##
## 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    1    0    3    2    2    5   12    9    9   11
##
## 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    1    0    3    2    1    2   10    9    9    9
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.97920782685131"
## [1] "Male first author team size 2018 geometric mean: 3.6889272467625"

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

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 31, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.57474738039426"
## [1] "Male last author team size 2018 geometric mean: 4.89897948556636"

## Warning in wilcox.test.default(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 = 0.2
## alternative hypothesis: true location shift is not equal to 0

```

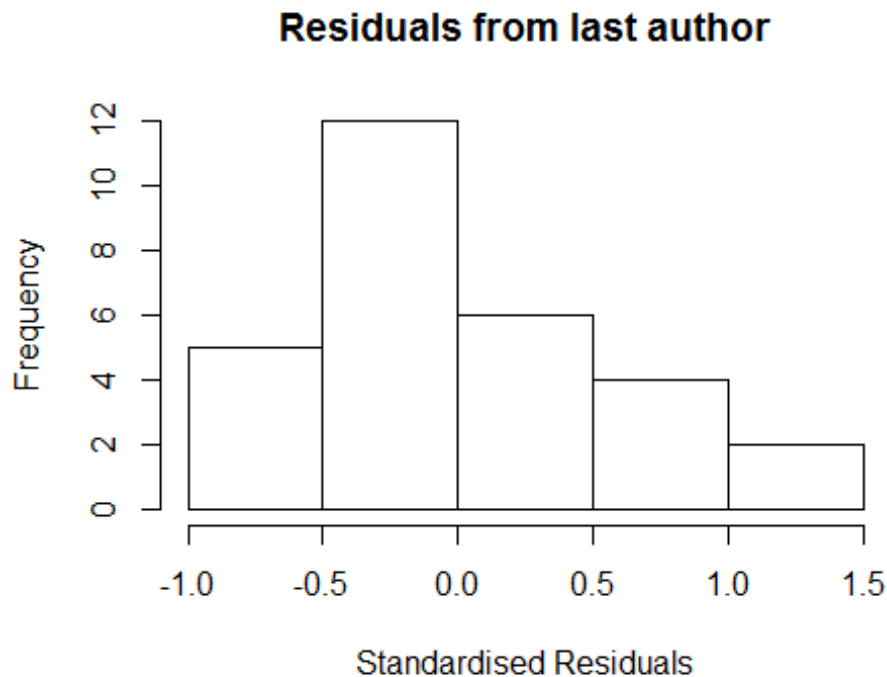
```
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"

## Warning in 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

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

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



```
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -6.747e+13  1      NaN
## LastAuthorFemale  1.860e+15  1    4.312e+07
## Year              3.999e+29 10    3.021e+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:
```

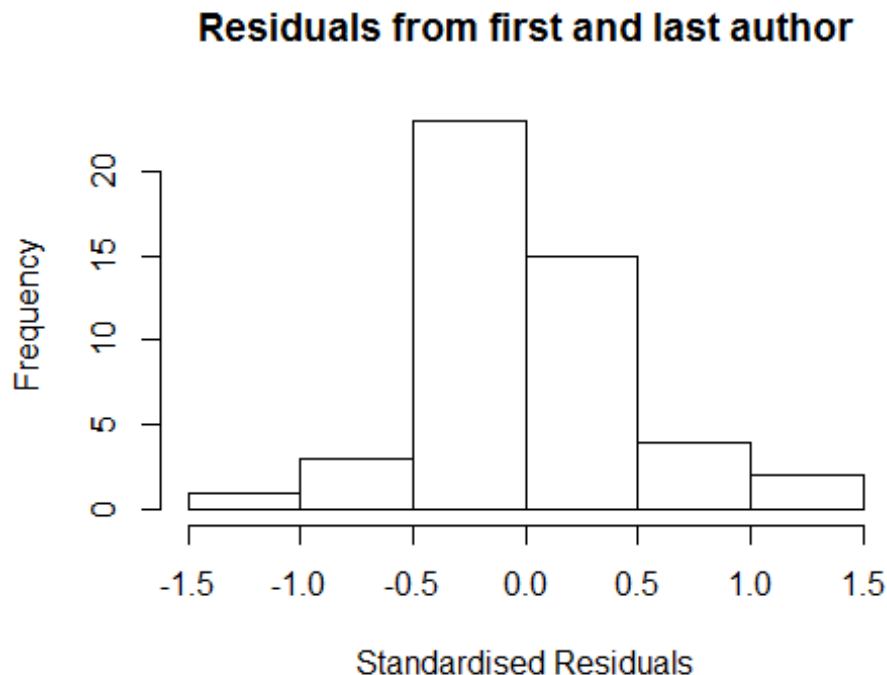
```

## 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
## -1.34e+00 -2.91e-01  5.55e-16  2.83e-01  1.12e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2670      0.0000      Inf < 2e-16 ***
## FirstAuthorFemale1 -0.0909      0.1699     -0.53  0.5961
## LastAuthorFemale1  0.3903      0.2143      1.82  0.0771 .
## Year2002          0.7069      0.1699      4.16  0.0002 ***
## Year2003         -1.5664      0.3225     -4.86  2.5e-05 ***
## Year2005         -0.6633      0.3053     -2.17  0.0366 *
## Year2006         -0.5371      0.2814     -1.91  0.0646 .
## Year2007         -1.5664      0.3225     -4.86  2.5e-05 ***
## Year2008         -0.1030      0.3994     -0.26  0.7980
## Year2009         -0.4090      0.4008     -1.02  0.3145
## Year2010         -0.3344      0.2356     -1.42  0.1646
## Year2011          0.1346      0.2049      0.66  0.5153
## Year2012          0.0553      0.1712      0.32  0.7487
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.53
## Multiple R-squared:  0.393, Adjusted R-squared:  0.185
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 41 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.503  0.927  0.960   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      2.08e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          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
```



```
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 4.735e+14  1      2.176e+07
## Year              4.735e+14 10      5.417e+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
## -1.1373 -0.2551 -0.0156  0.3392  1.2131
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2670     0.0000    Inf < 2e-16 ***
## FirstAuthorFemale1 -0.0725     0.1776   -0.41  0.68555
```

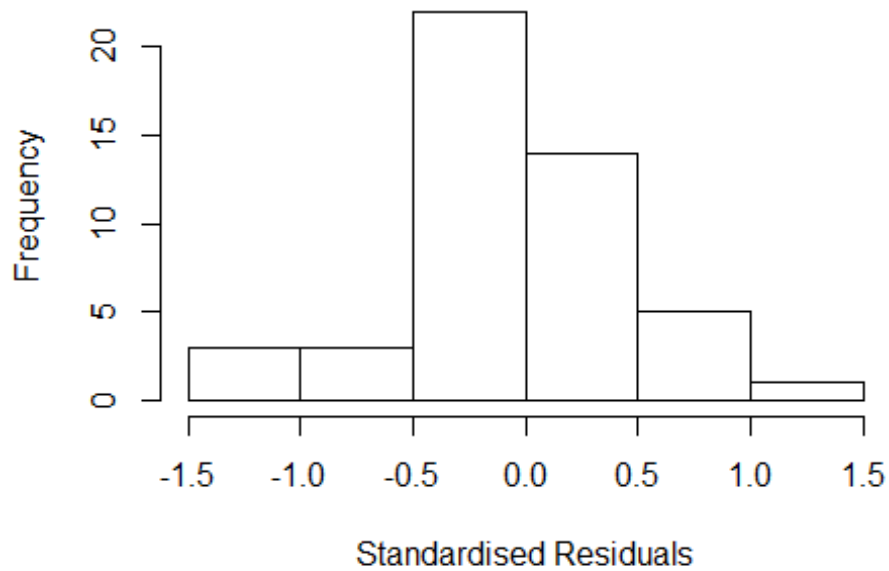
```

## Year2002          0.6885      0.1776      3.88  0.00043 ***
## Year2003         -1.1945      0.1776     -6.73  7.5e-08 ***
## Year2005         -0.5448      0.2146     -2.54  0.01560 *
## Year2006         -0.5463      0.2885     -1.89  0.06634 .
## Year2007         -1.1945      0.1776     -6.73  7.5e-08 ***
## Year2008         -0.1030      0.3877     -0.27  0.79203
## Year2009         -0.1297      0.3283     -0.39  0.69524
## Year2010         -0.1283      0.2128     -0.60  0.55035
## Year2011          0.2229      0.2012      1.11  0.27533
## Year2012          0.0947      0.1613      0.59  0.56083
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.579
## Multiple R-squared:  0.317, Adjusted R-squared:  0.108
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 41 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.640  0.915   0.965   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           2.08e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01                5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500           50           2           1           1000           200
## trace.lev      mts      compute.rd
##           0           1000           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

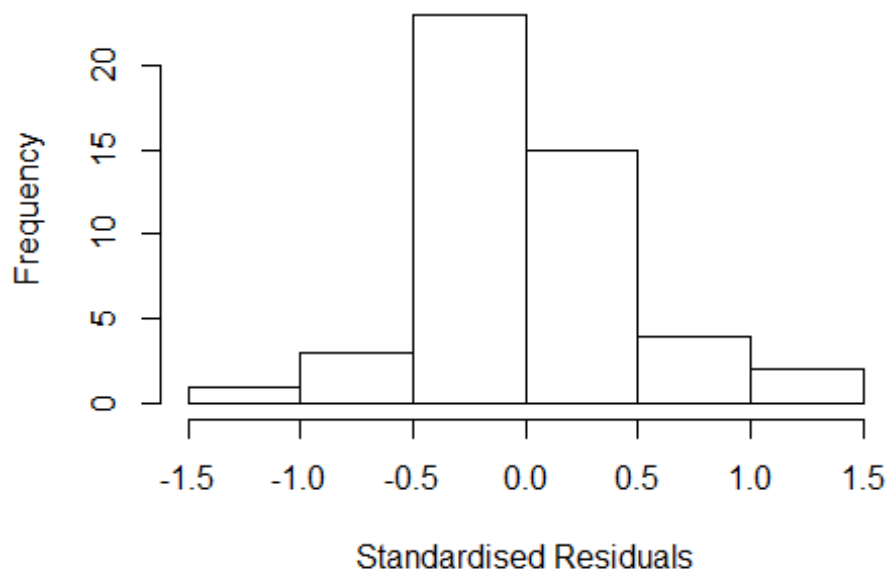
```

Residuals from first author



| | | | | |
|----|------------------|------------|----|-------------------------|
| ## | | GVIF | Df | $GVIF^{1/(2 \cdot Df)}$ |
| ## | LastAuthorFemale | -2.751e+15 | 1 | NaN |
| ## | Year | -2.751e+15 | 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
## -1.36659 -0.30701 -0.00276 0.24601 1.09624
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.27e+00 0.00e+00 Inf < 2e-16 ***
## LastAuthorFemale1 3.82e-01 2.16e-01 1.77e+00 0.085 .
## Year2002 6.16e-01 2.06e-08 2.99e+07 < 2e-16 ***
## Year2003 -1.65e+00 2.16e-01 -7.64e+00 4.9e-09 ***
## Year2005 -7.25e-01 2.79e-01 -2.60e+00 0.013 *
## Year2006 -5.83e-01 3.11e-01 -1.87e+00 0.069 .
## Year2007 -1.65e+00 2.16e-01 -7.64e+00 4.9e-09 ***
## Year2008 -1.03e-01 3.95e-01 -2.60e-01 0.796
## Year2009 -4.59e-01 3.40e-01 -1.35e+00 0.185
## Year2010 -3.84e-01 1.63e-01 -2.35e+00 0.024 *
## Year2011 7.95e-02 1.85e-01 4.30e-01 0.670
## Year2012 5.40e-03 1.29e-01 4.00e-02 0.967
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.546
## Multiple R-squared: 0.387, Adjusted R-squared: 0.2
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 39 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.511 0.918 0.959 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 2.08e-03 1.82e-12
## warn.limit.reject warn.limit.meanrw
## 5.00e-01 5.00e-01
## nResample max.it best.r.s k.fast.s k.max maxit.scale
## 500 50 2 1 1000 200
## trace.lev mts compute.rd
## 0 1000 0
## psi subsampling cov

```

```

##          "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##          "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 48"
## [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]"
##
## 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    1    1    3    2    4    2    2
##
## 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    1    0    1    1    3    2    4    2    2
##
## 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    1    0    1    1    3    1    3    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: 2.23606797749979"
##
## 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: 4.47213595499958"
## [1] "Male last author team size 2018 geometric mean: 1"
##
## Wilcoxon rank sum test
##
## 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"
## [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 2923"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 2002 2006 2009 2010 2011 2012
##    1    2    2    1    2    3
##
## 2002 2006 2009 2010 2011 2012
##    1    0    2    1    2    3
##
## 2002 2006 2009 2010 2011 2012
##    1    0    1    0    2    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: 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 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2010 2011 2012
##    4    2    3    2    2    4    3    4    1    5    6    4    3    4    5
##
## 1996 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2010 2011 2012
##    0    1    0    2    1    1    2    2    1    1    5    3    3    3    2
##
## 1996 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2010 2011 2012
##    0    1    0    2    1    1    1    2    1    1    5    3    3    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"

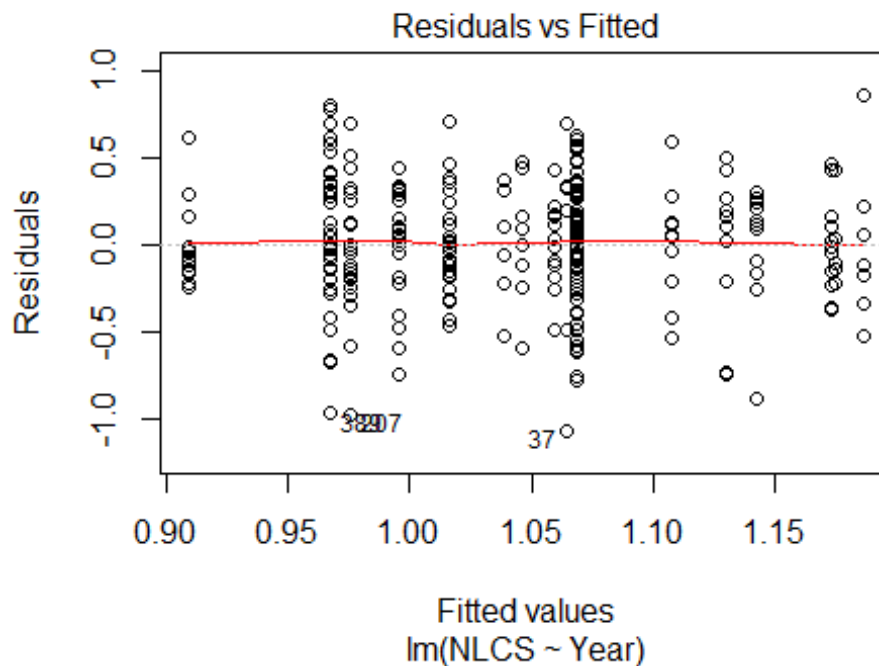
```

```

## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 25"
## [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]"
##
## 1999 2001 2009 2011 2012
##    1    1    3    1    1
##
## 1999 2001 2009 2011 2012
##    0    1    3    1    1
##
## 1999 2001 2009 2011 2012
##    0    1    2    1    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: 5"
## [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
##   14    9    8   13   17   18    8   15   14    7   16   27   33   31   51
## 2011 2012
##   58   34
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   12    7    6   11   10   11    5   10   11    6   12   23   29   23   43
## 2011 2012
##   49   26
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   10    6    4   11    7    9    4    8   11    6   11   18   26   21   36
## 2011 2012

```

```
## 45 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 = 28, df = 16, p-value = 0.03
```



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

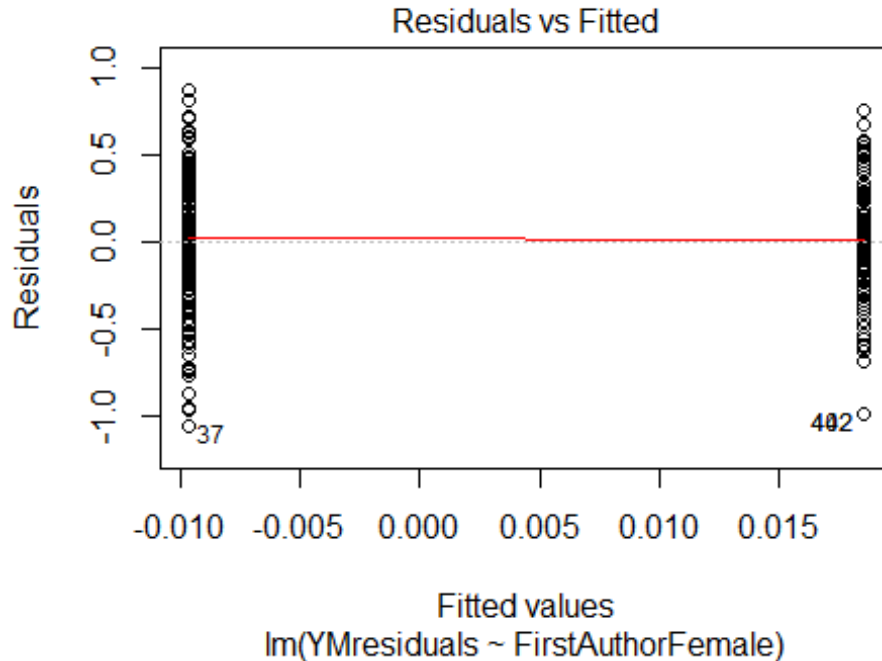
## [1] "Female first author team size 2018 geometric mean: 5.98511147895538"
## [1] "Male first author team size 2018 geometric mean: 3.58162124675998"

## Warning in wilcox.test.default(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] "Female last author team size 2018 geometric mean: 6.24063701205153"
## [1] "Male last author team size 2018 geometric mean: 3.09317070639762"

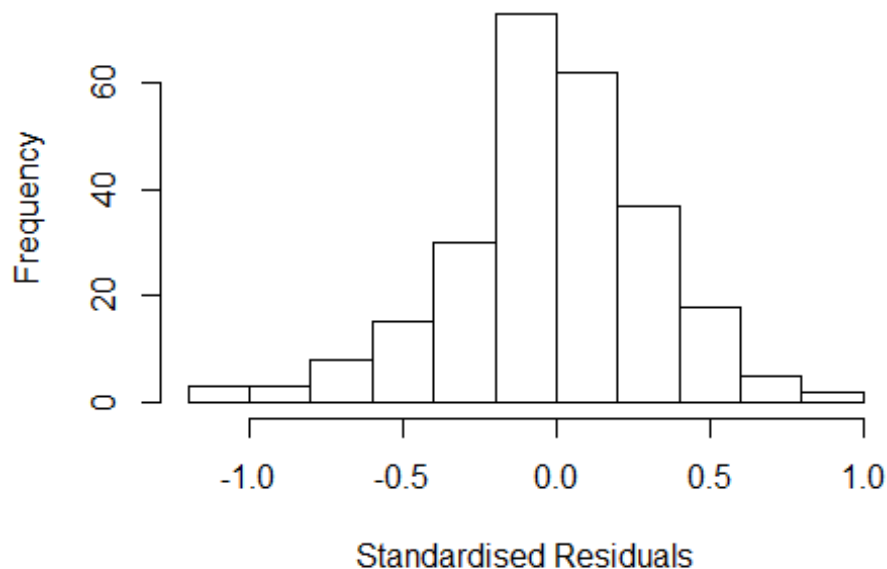
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 50, p-value = 0.05
## 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 | 1.689 | 1 | 1.300 |
| LastAuthorFemale | 1.705 | 1 | 1.306 |
| UniqueAuthors | 8.952 | 4 | 1.315 |
| Year | 20.781 | 16 | 1.099 |

Residuals from first and last author and team size



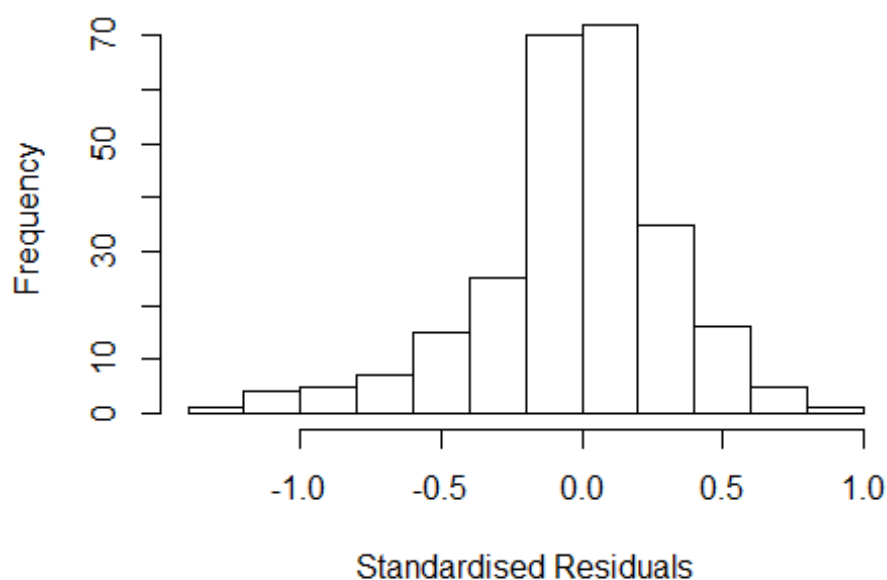
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.10828 -0.17837 -0.00654 0.19178 0.87933
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.78740 0.24533 3.21 0.0015 **
## FirstAuthorFemale1 0.00698 0.04504 0.15 0.8770
## LastAuthorFemale1 0.02500 0.04937 0.51 0.6131
## UniqueAuthors2 0.23561 0.22921 1.03 0.3051
## UniqueAuthors3 0.23429 0.23816 0.98 0.3263
## UniqueAuthors4 0.30784 0.23181 1.33 0.1855
## UniqueAuthors5 0.32173 0.22387 1.44 0.1520
## Year1997 0.07143 0.22770 0.31 0.7540
## Year1998 0.28861 0.15189 1.90 0.0586 .
## Year1999 0.15104 0.11410 1.32 0.1869
```

```

## Year2000      0.20437      0.14419      1.42      0.1577
## Year2001      0.05575      0.13744      0.41      0.6854
## Year2002     -0.00196      0.09068     -0.02      0.9828
## Year2003      0.04361      0.13945      0.31      0.7548
## Year2004      0.13007      0.11554      1.13      0.2614
## Year2005      0.00120      0.23483      0.01      0.9959
## Year2006     -0.22356      0.10094     -2.21      0.0277 *
## Year2007     -0.16002      0.11072     -1.45      0.1498
## Year2008     -0.04176      0.10169     -0.41      0.6817
## Year2009     -0.02284      0.09500     -0.24      0.8102
## Year2010     -0.01784      0.09580     -0.19      0.8524
## Year2011     -0.03283      0.09874     -0.33      0.7398
## Year2012      0.07138      0.12525      0.57      0.5693
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.277
## Multiple R-squared:  0.133, Adjusted R-squared:  0.0513
## Convergence in 32 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 234 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0746 0.8330 0.9500 0.8730 0.9870 0.9990
## Algorithmic parameters:
##      tuning.chi      bb      tuning.psi      refine.tol
##      1.55e+00      5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      3.91e-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 1.583 1      1.258
## LastAuthorFemale 1.692 1      1.301
## Year      2.649 16      1.031

```


Residuals from first and last author



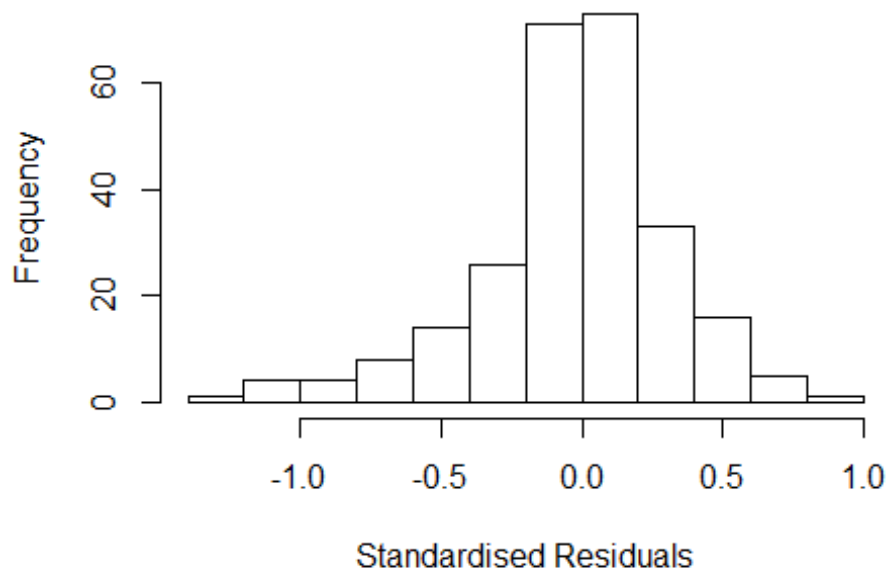
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.34254 -0.17256  0.00323  0.18446  0.93956
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.02510    0.07138   14.36 < 2e-16 ***
## FirstAuthorFemale1 0.01986    0.04607    0.43  0.667
## LastAuthorFemale1 0.03382    0.04906    0.69  0.491
## Year1997         0.08134    0.21469    0.38  0.705
## Year1998         0.31744    0.07922    4.01 8.2e-05 ***
## Year1999         0.16718    0.10840    1.54  0.124
## Year2000         0.16608    0.13878    1.20  0.233
## Year2001         0.01688    0.11079    0.15  0.879
## Year2002         0.03296    0.08520    0.39  0.699
## Year2003         0.07937    0.14057    0.56  0.573
## Year2004         0.12420    0.11055    1.12  0.262
## Year2005         0.00640    0.17162    0.04  0.970
```

```

## Year2006          -0.19102      0.09044      -2.11      0.036 *
## Year2007          -0.10765      0.10481      -1.03      0.305
## Year2008          -0.01227      0.09061      -0.14      0.892
## Year2009           0.00608      0.09357       0.06      0.948
## Year2010           0.00596      0.08491       0.07      0.944
## Year2011          -0.01651      0.09582      -0.17      0.863
## Year2012           0.10475      0.11321       0.93      0.356
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.284
## Multiple R-squared:  0.0763, Adjusted R-squared:  0.00611
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## observation 20 is an outlier with |weight| = 0 ( < 0.00039);
## 24 weights are ~= 1. The remaining 231 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.131  0.846   0.956   0.875   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           3.91e-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 1.587 1           1.260
## Year              1.587 16           1.015

```

Residuals from first author



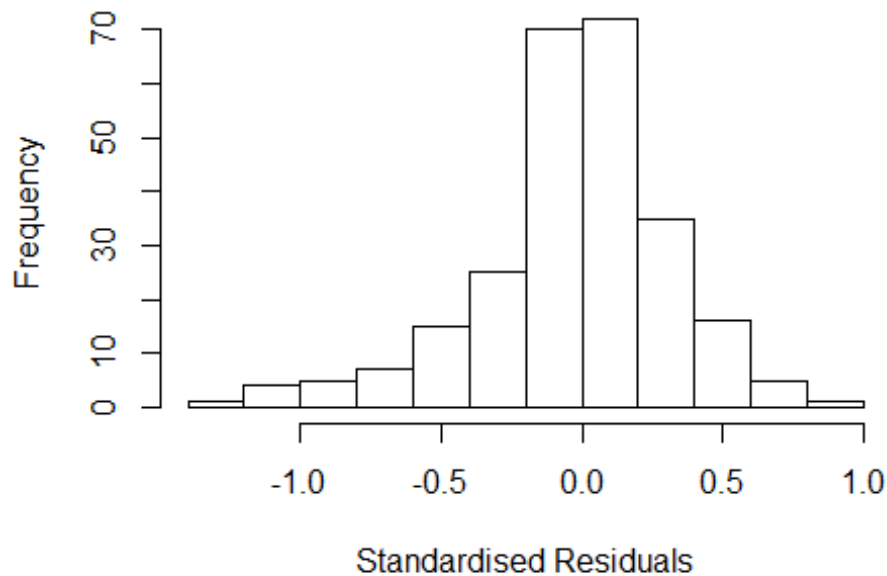
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.341837 -0.177883 -0.000775 0.181979 0.937353
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.03229 0.07015 14.72 < 2e-16 ***
## FirstAuthorFemale1 0.02195 0.04584 0.48 0.633
## Year1997 0.07636 0.20732 0.37 0.713
## Year1998 0.30955 0.07759 3.99 8.8e-05 ***
## Year1999 0.16572 0.10497 1.58 0.116
## Year2000 0.16596 0.14049 1.18 0.239
## Year2001 0.01652 0.11247 0.15 0.883
## Year2002 0.03448 0.08511 0.41 0.686
## Year2003 0.07786 0.13844 0.56 0.574
## Year2004 0.12345 0.11239 1.10 0.273
## Year2005 0.02186 0.16246 0.13 0.893
## Year2006 -0.18534 0.08927 -2.08 0.039 *
```

```

## Year2007          -0.10670      0.10547      -1.01      0.313
## Year2008          -0.01494      0.08994      -0.17      0.868
## Year2009           0.00417      0.09367       0.04      0.965
## Year2010           0.00955      0.08530       0.11      0.911
## Year2011          -0.00731      0.09526      -0.08      0.939
## Year2012           0.11285      0.11022       1.02      0.307
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.283
## Multiple R-squared:  0.0754, Adjusted R-squared:  0.00932
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## observation 20 is an outlier with |weight| = 0 ( < 0.00039);
## 25 weights are ~= 1. The remaining 230 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.140  0.846  0.958  0.873  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      3.91e-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 1.657 1          1.287
## Year              1.657 16          1.016

```

Residuals from last author



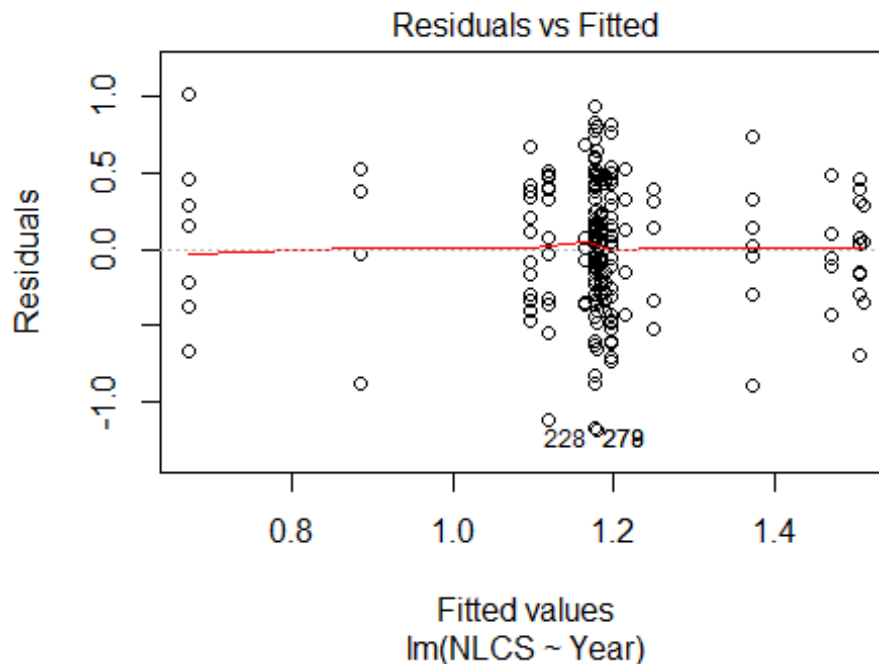
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.34922 -0.18012 0.00228 0.17895 0.93596
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.02920 0.07090 14.52 < 2e-16 ***
## LastAuthorFemale1 0.03493 0.04862 0.72 0.473
## Year1997 0.08084 0.21146 0.38 0.703
## Year1998 0.32003 0.08037 3.98 9.1e-05 ***
## Year1999 0.17465 0.10476 1.67 0.097 .
## Year2000 0.16256 0.13827 1.18 0.241
## Year2001 0.02265 0.11071 0.20 0.838
## Year2002 0.02856 0.08458 0.34 0.736
## Year2003 0.07473 0.13830 0.54 0.589
## Year2004 0.12824 0.10990 1.17 0.244
## Year2005 0.00261 0.16768 0.02 0.988
## Year2006 -0.18782 0.09094 -2.07 0.040 *
```

```

## Year2007          -0.10206      0.10364    -0.98      0.326
## Year2008          -0.00510      0.08912    -0.06      0.954
## Year2009           0.00974      0.09241     0.11      0.916
## Year2010           0.00792      0.08523     0.09      0.926
## Year2011          -0.01266      0.09580    -0.13      0.895
## Year2012           0.10993      0.11406     0.96      0.336
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.282
## Multiple R-squared:  0.0758, Adjusted R-squared:  0.00978
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## observation 20 is an outlier with |weight| = 0 ( < 0.00039);
## 21 weights are ~= 1. The remaining 234 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.135  0.847   0.957   0.875   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      3.91e-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: 256"
## [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
##   16   17   13    5   10    9   11   21   13   12   20   14   17   23   26
## 2011 2012
##   56   39
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    7    5    4    5    3    6   15    9    8   13    8   12   17   16

```

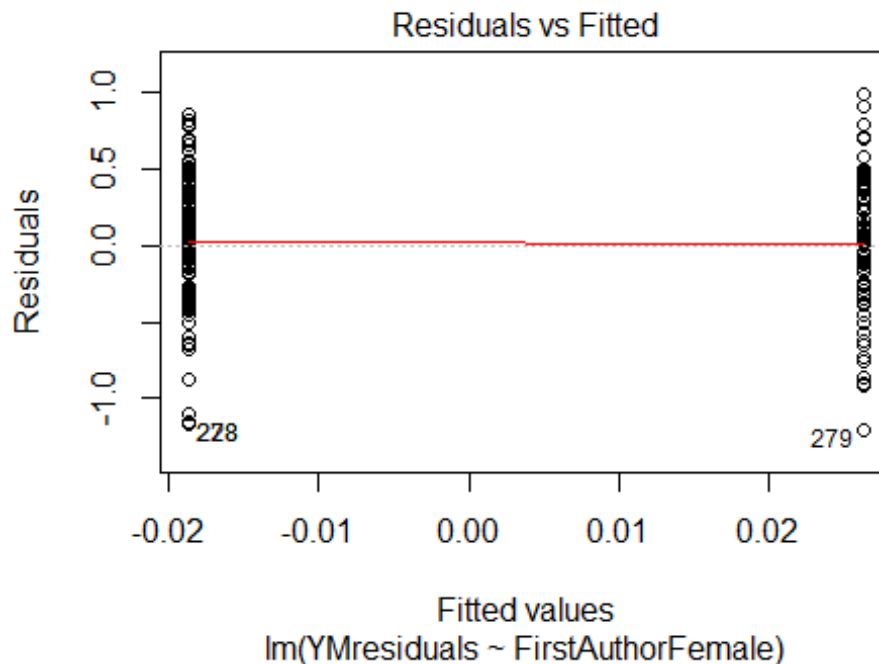
```
## 2011 2012
## 47 28
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 6 7 3 4 4 3 5 12 8 8 12 8 11 16 14
## 2011 2012
## 42 23
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 23, df = 16, p-value = 0.1
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.093, df = 1, p-value = 0.8
## [1] "Female first author team size 2018 geometric mean: 3.44747907076565"
## [1] "Male first author team size 2018 geometric mean: 4.3436839171418"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 89, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.4336011289598"
## [1] "Male last author team size 2018 geometric mean: 4.36530175951803"

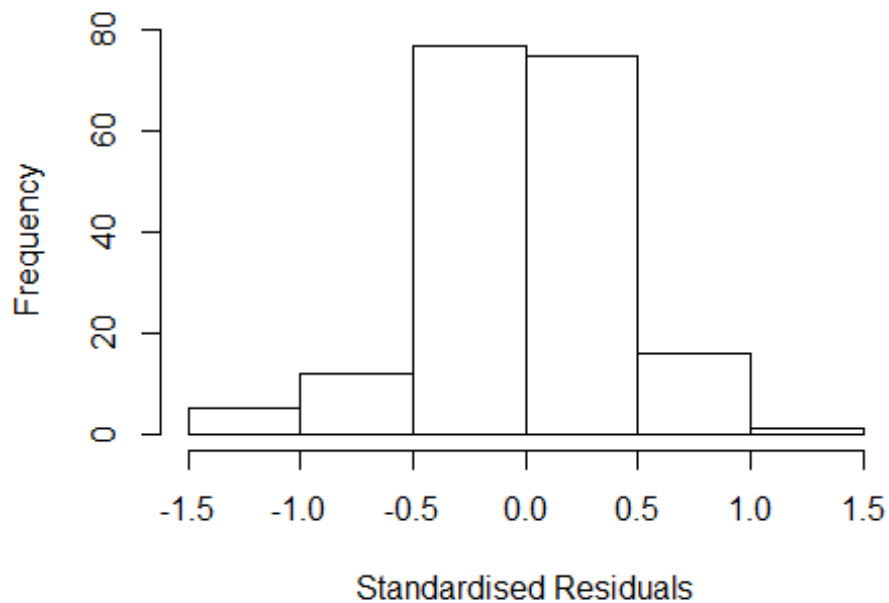
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 90, p-value = 0.5
## 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 | 2.127 | 1 | 1.458 |
| LastAuthorFemale | 1.621 | 1 | 1.273 |
| UniqueAuthors | 169.670 | 4 | 1.900 |
| Year | 256.206 | 16 | 1.189 |

Residuals from first and last author and team size



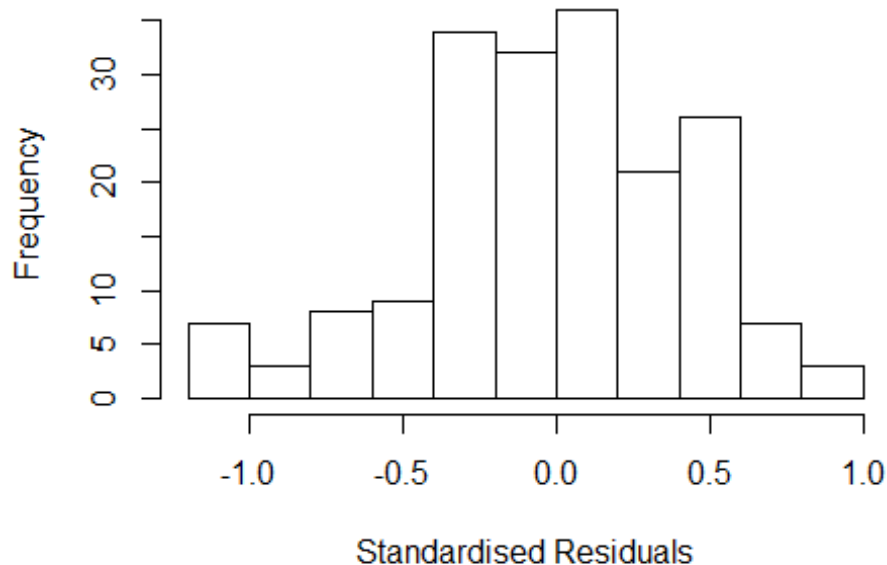
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.19277 -0.28743 -0.00253  0.27848  1.47787
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.50253    0.27332    1.84   0.0678 .
## FirstAuthorFemale1 0.04038    0.08117    0.50   0.6195
## LastAuthorFemale1 0.05005    0.07257    0.69   0.4914
## UniqueAuthors2    0.60965    0.20679    2.95   0.0037 **
## UniqueAuthors3    0.61527    0.22061    2.79   0.0059 **
## UniqueAuthors4    0.56392    0.20976    2.69   0.0079 **
## UniqueAuthors5    0.59693    0.20521    2.91   0.0041 **
## Year1997          0.27837    0.27035    1.03   0.3047
## Year1998          0.84311    0.22827    3.69   0.0003 ***
## Year1999         -0.10344    0.25525   -0.41   0.6858
```

```

## Year2000      0.03583    0.26776    0.13    0.8937
## Year2001      0.34911    0.27826    1.25    0.2114
## Year2002      0.00471    0.26136    0.02    0.9856
## Year2003      0.01096    0.18753    0.06    0.9535
## Year2004      0.38211    0.19712    1.94    0.0543 .
## Year2005     -0.41085    0.26225   -1.57    0.1191
## Year2006      0.01351    0.26523    0.05    0.9594
## Year2007      0.05047    0.22534    0.22    0.8231
## Year2008      0.19034    0.19953    0.95    0.3415
## Year2009      0.04291    0.21417    0.20    0.8415
## Year2010      0.01603    0.18346    0.09    0.9305
## Year2011      0.03459    0.23217    0.15    0.8817
## Year2012      0.05598    0.20844    0.27    0.7886
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.397
## Multiple R-squared:  0.205, Adjusted R-squared:  0.0976
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 168 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.136  0.886  0.947  0.898  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      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 1.665 1      1.290
## LastAuthorFemale  1.600 1      1.265
## Year              2.387 16      1.028

```

Residuals from first and last author



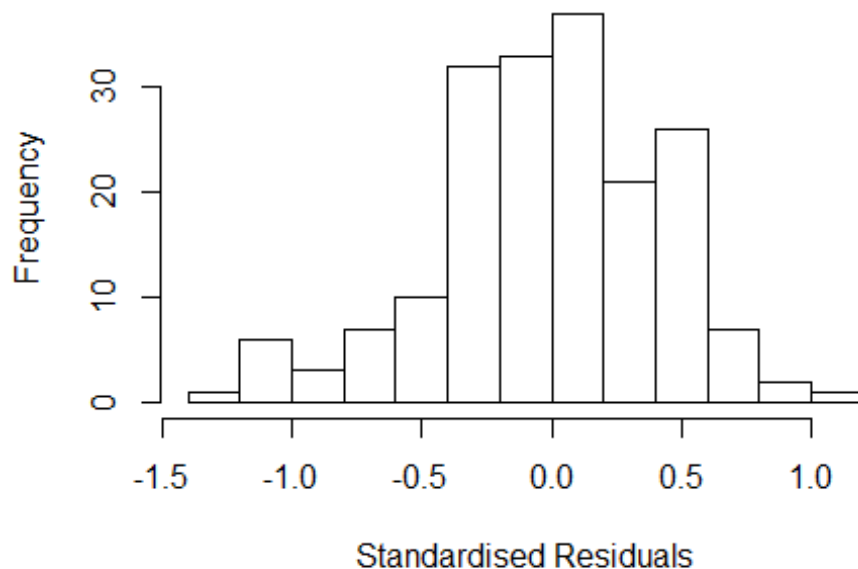
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1844 -0.2954 0.0076 0.3053 0.9761
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.11075 0.14954 7.43 5.4e-12 ***
## FirstAuthorFemale1 0.05850 0.07681 0.76 0.447
## LastAuthorFemale1 0.04794 0.07650 0.63 0.532
## Year1997 0.26024 0.22906 1.14 0.258
## Year1998 0.42277 0.23115 1.83 0.069 .
## Year1999 -0.14746 0.36512 -0.40 0.687
## Year2000 0.02769 0.23594 0.12 0.907
## Year2001 0.32868 0.23018 1.43 0.155
## Year2002 -0.00989 0.22476 -0.04 0.965
## Year2003 0.00306 0.17348 0.02 0.986
## Year2004 0.30153 0.18984 1.59 0.114
## Year2005 -0.50530 0.26795 -1.89 0.061 .
```

```

## Year2006      -0.03584    0.22095   -0.16    0.871
## Year2007      0.04318    0.21034    0.21    0.838
## Year2008      0.04907    0.17698    0.28    0.782
## Year2009      0.01553    0.18694    0.08    0.934
## Year2010      0.00137    0.15300    0.01    0.993
## Year2011      0.01514    0.16981    0.09    0.929
## Year2012      0.00355    0.17299    0.02    0.984
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.423
## Multiple R-squared:  0.12,   Adjusted R-squared:  0.0249
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 174 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.413  0.893  0.952  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      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 1.552 1      1.246
## Year              1.552 16      1.014

```

Residuals from first author



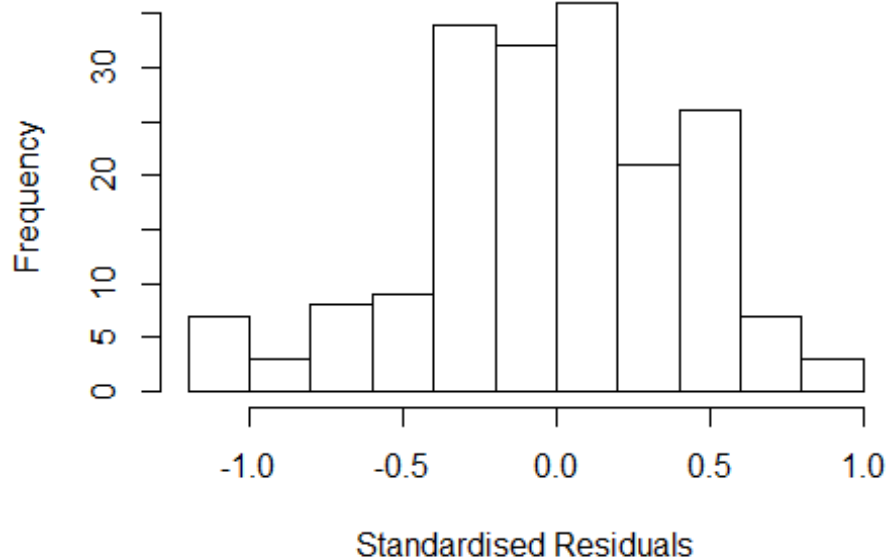
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.20644 -0.30474 0.00657 0.29491 1.00369
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.12159 0.14837 7.56 2.5e-12 ***
## FirstAuthorFemale1 0.05997 0.07604 0.79 0.431
## Year1997 0.25660 0.22864 1.12 0.263
## Year1998 0.42901 0.22525 1.90 0.059 .
## Year1999 -0.17325 0.35036 -0.49 0.622
## Year2000 0.03010 0.24083 0.12 0.901
## Year2001 0.33065 0.21976 1.50 0.134
## Year2002 -0.00883 0.22431 -0.04 0.969
## Year2003 0.00107 0.17483 0.01 0.995
## Year2004 0.31595 0.18870 1.67 0.096 .
## Year2005 -0.49726 0.26910 -1.85 0.066 .
## Year2006 -0.03929 0.22121 -0.18 0.859
```

```

## Year2007          0.04645    0.20709    0.22    0.823
## Year2008          0.04197    0.17752    0.24    0.813
## Year2009          0.00420    0.18719    0.02    0.982
## Year2010          0.00782    0.15455    0.05    0.960
## Year2011          0.02488    0.16906    0.15    0.883
## Year2012          0.00994    0.17393    0.06    0.954
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.436
## Multiple R-squared:  0.117, Adjusted R-squared:  0.0273
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 173 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.425  0.898  0.952  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      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 1.524 1          1.234
## Year            1.524 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
## -1.16788 -0.30021 -0.00153 0.30928 1.03101
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.13199 0.14302 7.91 3.2e-13 ***
## LastAuthorFemale1 0.04719 0.07615 0.62 0.536
## Year1997 0.25489 0.23281 1.09 0.275
## Year1998 0.42239 0.22186 1.90 0.059 .
## Year1999 -0.15066 0.36211 -0.42 0.678
## Year2000 0.00665 0.23263 0.03 0.977
## Year2001 0.36624 0.22387 1.64 0.104
## Year2002 -0.01243 0.22463 -0.06 0.956
## Year2003 -0.00786 0.17145 -0.05 0.963
## Year2004 0.30342 0.19072 1.59 0.114
## Year2005 -0.52219 0.27313 -1.91 0.058 .
## Year2006 -0.01130 0.21207 -0.05 0.958
```

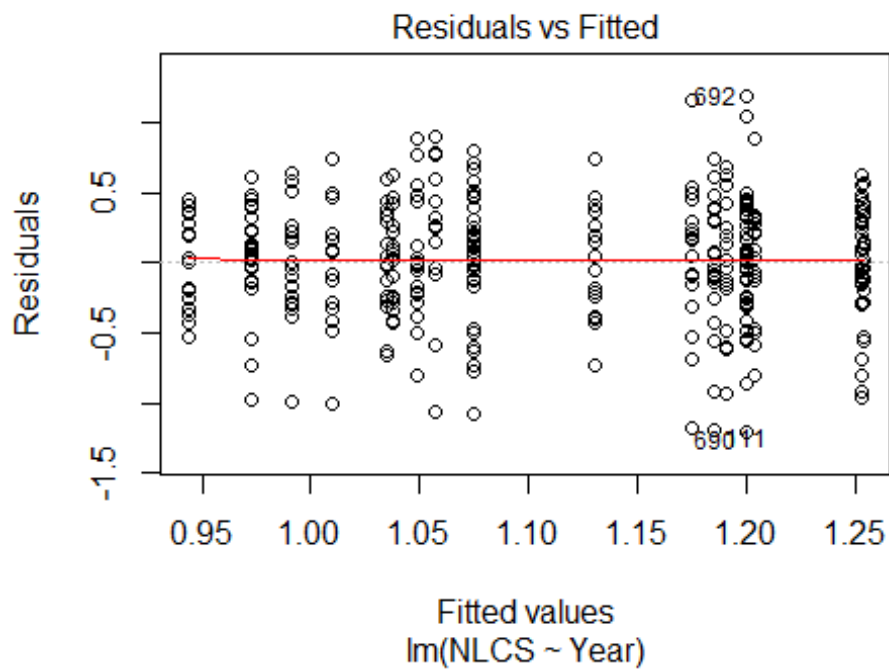
```

## Year2007      0.03011      0.20595      0.15      0.884
## Year2008      0.06127      0.17283      0.35      0.723
## Year2009      0.03323      0.18339      0.18      0.856
## Year2010     -0.00306      0.14975     -0.02      0.984
## Year2011      0.02511      0.16807      0.15      0.881
## Year2012      0.00639      0.17151      0.04      0.970
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.419
## Multiple R-squared:  0.118, Adjusted R-squared:  0.0291
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 173 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.418  0.886  0.948  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      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 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
##   33   35   23   43   29   34   22   22   19   24   27   34   26   27   51
## 2011 2012
##   55   49
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   15   18   16   24   15   16   16   17   16   18   18   25   21   18   32
## 2011 2012

```



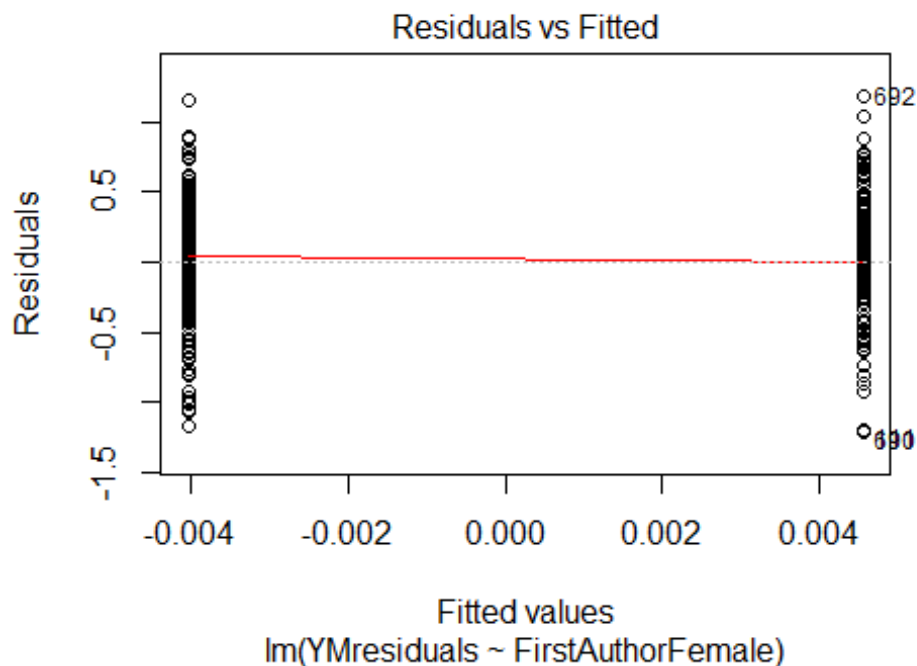
```
## 39 40
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 12 17 13 21 13 12 13 14 12 16 14 24 20 15 29
## 2011 2012
## 36 34
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 18, 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: 4.88519805159947"
## [1] "Male first author team size 2018 geometric mean: 2.7123532530899"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

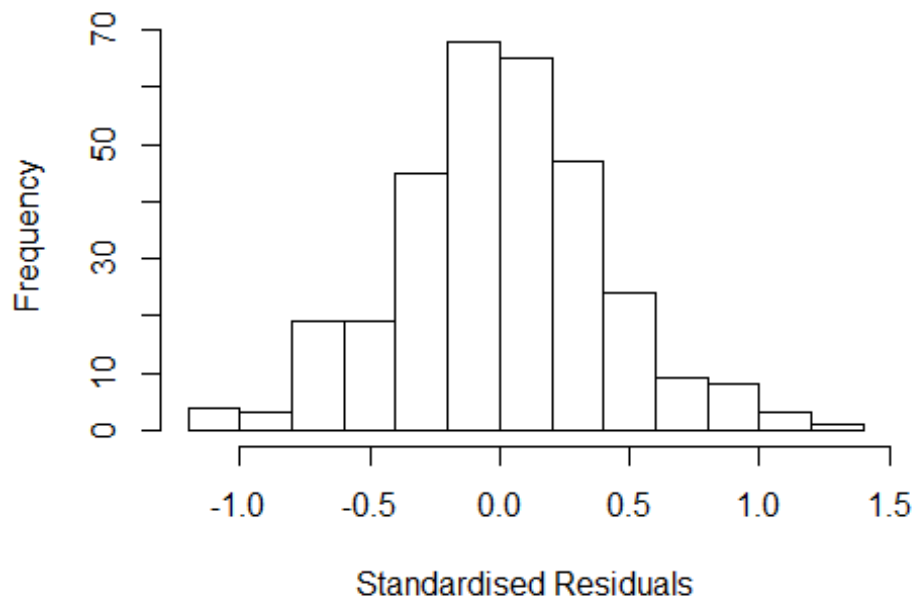
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 56, p-value = 0.05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.36599377745599"
## [1] "Male last author team size 2018 geometric mean: 3.50079251758821"

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



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 41, 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 1.385 1      1.177
## LastAuthorFemale  1.663 1      1.289
## UniqueAuthors    2.997 4      1.147
## Year             4.137 16      1.045
```

Residuals from first and last author and team size



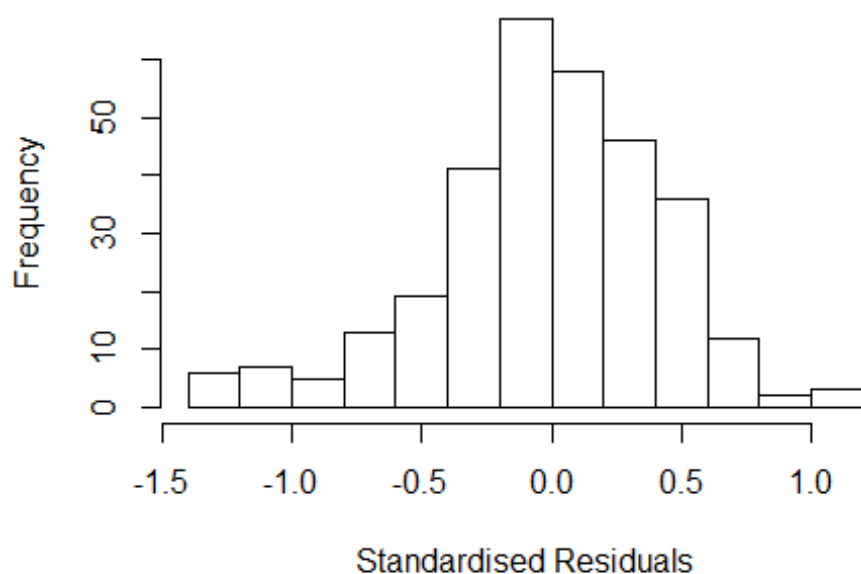
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.184736 -0.240503 -0.000722 0.253591 1.323086
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6884 0.1654 4.16 4.2e-05 ***
## FirstAuthorFemale1 -0.0699 0.0501 -1.39 0.16415
## LastAuthorFemale1 -0.0323 0.0744 -0.43 0.66473
## UniqueAuthors2 0.4454 0.1328 3.35 0.00090 ***
## UniqueAuthors3 0.4159 0.1325 3.14 0.00188 **
## UniqueAuthors4 0.4995 0.1334 3.75 0.00022 ***
## UniqueAuthors5 0.5421 0.1258 4.31 2.2e-05 ***
## Year1997 0.1644 0.1864 0.88 0.37863
## Year1998 0.0833 0.2755 0.30 0.76263
## Year1999 0.0913 0.1599 0.57 0.56849
```

```

## Year2000          0.0140      0.1593      0.09  0.93006
## Year2001         -0.0634      0.1730     -0.37  0.71408
## Year2002          0.0125      0.1807      0.07  0.94493
## Year2003         -0.1248      0.1628     -0.77  0.44406
## Year2004          0.2484      0.1599      1.55  0.12138
## Year2005         -0.0934      0.1629     -0.57  0.56705
## Year2006          0.0986      0.1794      0.55  0.58292
## Year2007         -0.0981      0.1571     -0.62  0.53272
## Year2008         -0.0844      0.1648     -0.51  0.60874
## Year2009          0.2125      0.1914      1.11  0.26791
## Year2010          0.1831      0.1565      1.17  0.24277
## Year2011          0.0577      0.1530      0.38  0.70642
## Year2012          0.1170      0.1521      0.77  0.44216
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.395
## Multiple R-squared:  0.185, Adjusted R-squared:  0.124
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 287 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.240  0.869  0.955  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      3.17e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.191 1      1.091
## LastAuthorFemale  1.481 1      1.217
## Year              1.702 16      1.017

```

Residuals from first and last author



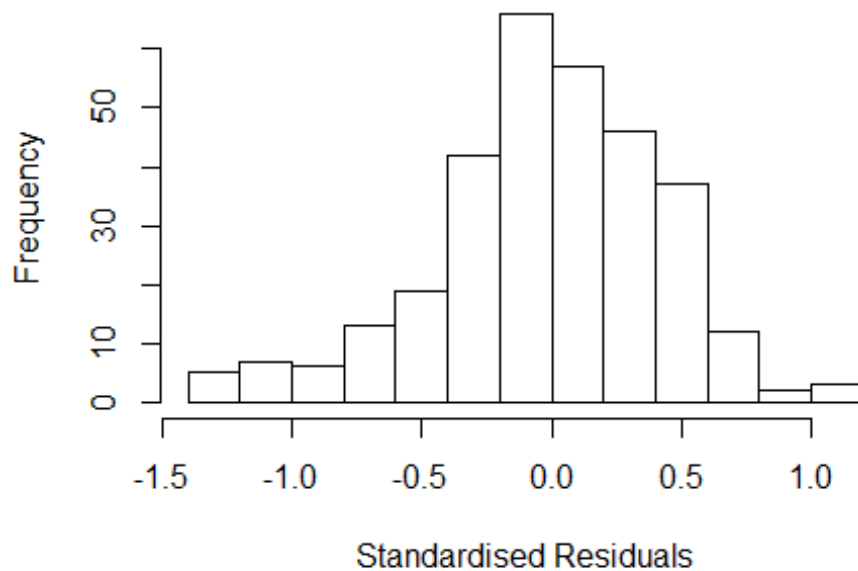
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.24204 -0.24575 -0.00704  0.28325  1.18036
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.02840    0.14591   7.05 1.3e-11 ***
## FirstAuthorFemale1 0.00285    0.04974   0.06  0.954
## LastAuthorFemale1 0.00652    0.07126   0.09  0.927
## Year1997         0.21363    0.18591   1.15  0.251
## Year1998         0.19035    0.39482   0.48  0.630
## Year1999         0.16799    0.17464   0.96  0.337
## Year2000         0.01634    0.16949   0.10  0.923
## Year2001        -0.06533    0.17377  -0.38  0.707
## Year2002         0.09265    0.18804   0.49  0.623
## Year2003        -0.03071    0.16175  -0.19  0.850
## Year2004         0.32979    0.16488   2.00  0.046 *
## Year2005        -0.01870    0.17226  -0.11  0.914
```

```

## Year2006          0.16296    0.19073    0.85    0.394
## Year2007          -0.02006    0.16186   -0.12    0.901
## Year2008          -0.00282    0.17475   -0.02    0.987
## Year2009          0.25097    0.18426    1.36    0.174
## Year2010          0.25430    0.15754    1.61    0.108
## Year2011          0.08002    0.16371    0.49    0.625
## Year2012          0.17239    0.15766    1.09    0.275
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.409
## Multiple R-squared:  0.0667, Adjusted R-squared:  0.00995
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 285 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.336  0.878  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      3.17e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##      trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.171 1      1.082
## Year              1.171 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
## -1.24198 -0.24269 -0.00921 0.28420 1.17814
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.03044 0.14228 7.24 3.8e-12 ***
## FirstAuthorFemale1 0.00271 0.04959 0.05 0.956
## Year1997 0.21154 0.18301 1.16 0.249
## Year1998 0.18643 0.38268 0.49 0.627
## Year1999 0.16604 0.17140 0.97 0.333
## Year2000 0.01430 0.16636 0.09 0.932
## Year2001 -0.06620 0.17260 -0.38 0.702
## Year2002 0.09192 0.18735 0.49 0.624
## Year2003 -0.03162 0.16042 -0.20 0.844
## Year2004 0.33007 0.16549 1.99 0.047 *
## Year2005 -0.01921 0.17245 -0.11 0.911
## Year2006 0.16098 0.18810 0.86 0.393
```

```

## Year2007          -0.02022    0.16218   -0.12    0.901
## Year2008          -0.00426    0.17364   -0.02    0.980
## Year2009           0.24950    0.18298    1.36    0.174
## Year2010           0.25345    0.15672    1.62    0.107
## Year2011           0.07948    0.16340    0.49    0.627
## Year2012           0.17272    0.15846    1.09    0.277
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.41
## Multiple R-squared:  0.0664, Adjusted R-squared:  0.013
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 285 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.338  0.879  0.953  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      3.17e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.435 1      1.198
## Year              1.435 16      1.011
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.24324 -0.24566 -0.00555  0.28342  1.18173

```



```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.03003    0.14506   7.10 9.2e-12 ***
## LastAuthorFemale1 0.00675    0.07114   0.09  0.924
## Year1997        0.21321    0.18685   1.14  0.255
## Year1998        0.19353    0.39834   0.49  0.627
## Year1999        0.16808    0.17499   0.96  0.338
## Year2000        0.01502    0.16955   0.09  0.929
## Year2001       -0.06583    0.17404  -0.38  0.706
## Year2002        0.09205    0.18820   0.49  0.625
## Year2003       -0.03157    0.16221  -0.19  0.846
## Year2004        0.32952    0.16495   2.00  0.047 *
## Year2005       -0.01947    0.17259  -0.11  0.910
## Year2006        0.16299    0.19085   0.85  0.394
## Year2007       -0.01990    0.16200  -0.12  0.902
## Year2008       -0.00278    0.17416  -0.02  0.987
## Year2009        0.25176    0.18337   1.37  0.171
## Year2010        0.25418    0.15788   1.61  0.108
## Year2011        0.08023    0.16369   0.49  0.624
## Year2012        0.17224    0.15781   1.09  0.276
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.407
## Multiple R-squared:  0.0671, Adjusted R-squared:  0.0137
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 287 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.331  0.878  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      3.17e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 315"
## [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
##    7   12    5    5    3    6    4    4   11    6   11   14   13    8   17
## 2011 2012
##   21   12
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    6    1    2    3    3    2    2    7    6    4    7    6    5   12
## 2011 2012
##   14    8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    5    1    2    2    3    2    2    5    5    4    7    5    4   12
## 2011 2012
##   14    5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.78938895318699"
## [1] "Male first author team size 2018 geometric mean: 1.25992104989487"

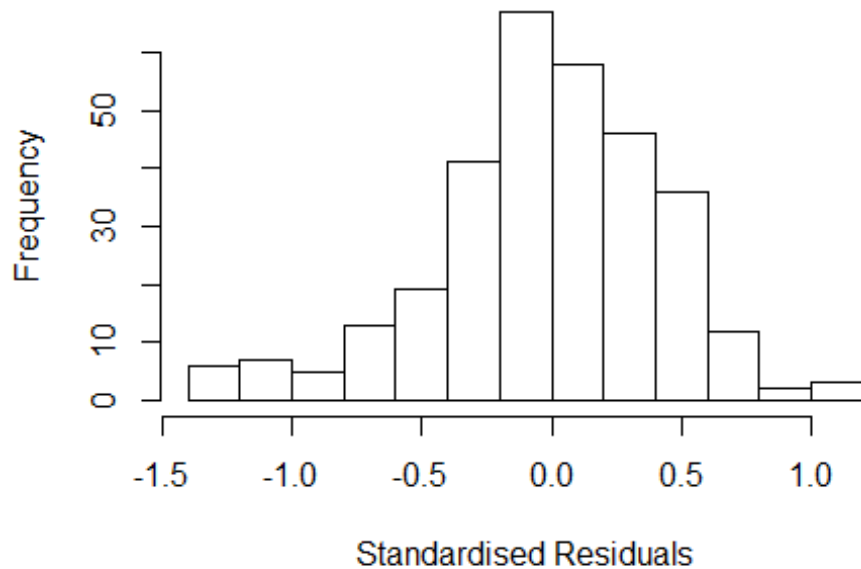
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 14, p-value = 0.05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.51984209978975"
## [1] "Male last author team size 2018 geometric mean: 3.15981830574927"

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

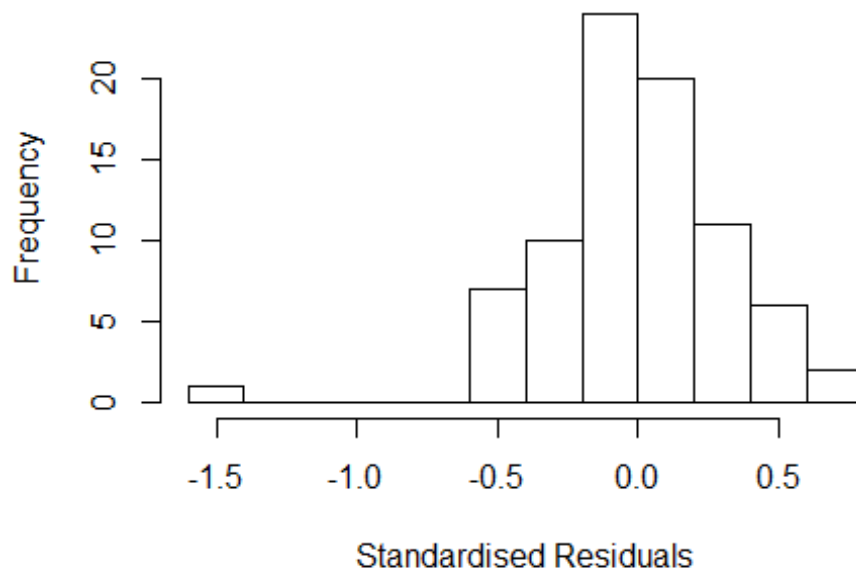
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 6, 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    8.171  1      2.858
## LastAuthorFemale    20.138  1      4.488
## UniqueAuthors    777.267  3      3.032
## Year          9077.598 16      1.329
```

Residuals from first and last author and team size



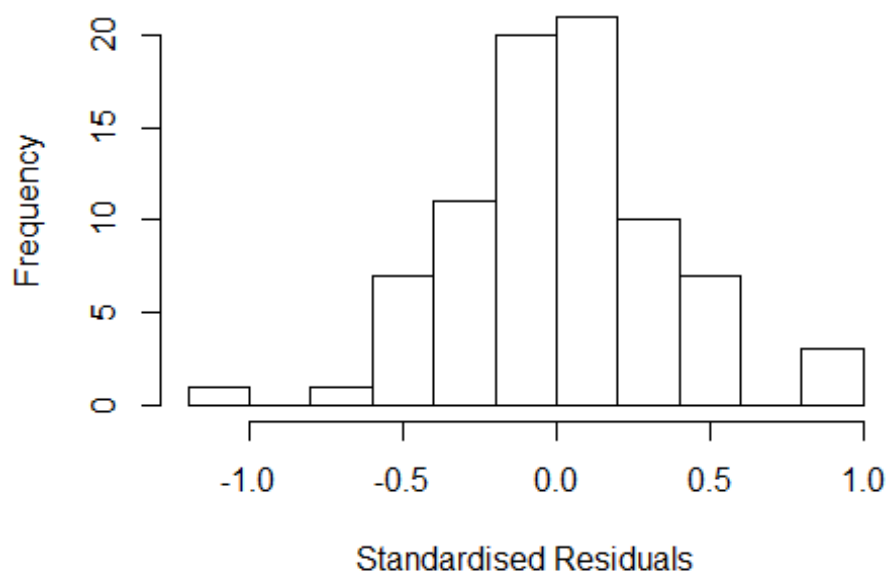
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4304 -0.1794 -0.0175 0.1758 0.6618
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.073512 0.203267 5.28 1.9e-06 ***
## FirstAuthorFemale1 -0.133737 0.085456 -1.56 0.12293
## LastAuthorFemale1 -0.000468 0.079827 -0.01 0.99535
## UniqueAuthors3 0.212937 0.121332 1.75 0.08445 .
## UniqueAuthors4 -0.042980 0.117210 -0.37 0.71516
## UniqueAuthors5 0.397475 0.119591 3.32 0.00153 **
## Year1997 0.166238 0.206324 0.81 0.42364
## Year1998 0.235936 0.204671 1.15 0.25366
## Year1999 -0.726391 0.194190 -3.74 0.00042 ***
## Year2000 -0.618253 0.233216 -2.65 0.01029 *
```

```

## Year2001      -0.258787    0.232642    -1.11    0.27049
## Year2002      -0.540512    0.204450    -2.64    0.01049 *
## Year2003      -0.242987    0.498938    -0.49    0.62806
## Year2004       0.033601    0.246769     0.14    0.89215
## Year2005       0.099648    0.247301     0.40    0.68845
## Year2006       0.004531    0.195205     0.02    0.98156
## Year2007       0.234261    0.215010     1.09    0.28035
## Year2008       0.144140    0.190199     0.76    0.45156
## Year2009      -0.187809    0.178734    -1.05    0.29765
## Year2010       0.177669    0.219060     0.81    0.42060
## Year2011      -0.040581    0.212555    -0.19    0.84924
## Year2012      -0.291968    0.207103    -1.41    0.16386
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.326
## Multiple R-squared:  0.485, Adjusted R-squared:  0.301
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 74 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0157 0.8930 0.9650 0.9190 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          1.23e-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 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 4.937 1 2.222
## LastAuthorFemale 45.121 1 6.717
## Year 170.764 16 1.174

```

Residuals from first and last author



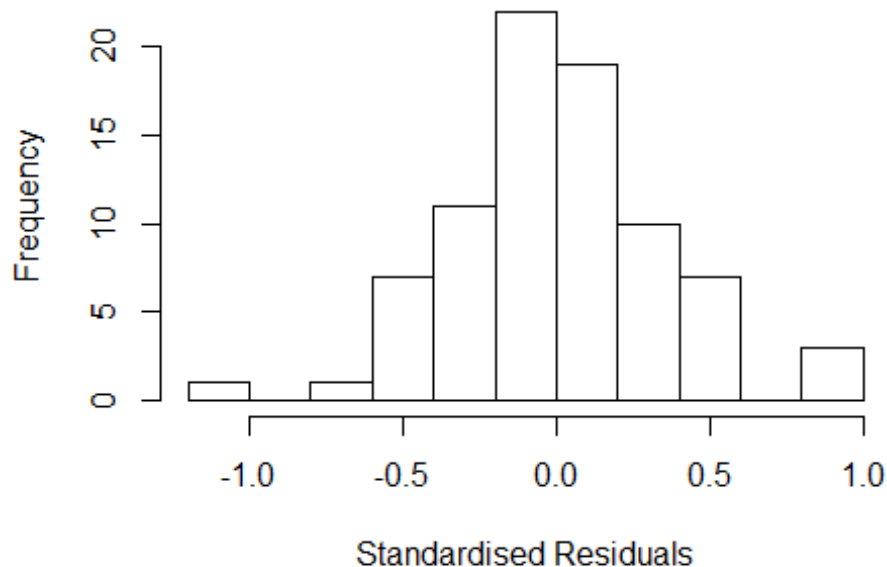
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.09659 -0.19515  0.00126  0.17612  0.86781
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.20943    0.21667    5.58 5.6e-07 ***
## FirstAuthorFemale1 -0.17540    0.08772   -2.00  0.0499 *
## LastAuthorFemale1  0.01618    0.12118    0.13  0.8942
## Year1997         0.11786    0.26237    0.45  0.6549
## Year1998         0.04039    0.24966    0.16  0.8720
## Year1999        -0.66423    0.22428   -2.96  0.0043 **
## Year2000        -0.36502    0.25982   -1.40  0.1650
## Year2001        -0.22776    0.27599   -0.83  0.4124
## Year2002        -0.67643    0.21778   -3.11  0.0029 **
## Year2003         0.01857    0.50588    0.04  0.9708
## Year2004         0.21153    0.29424    0.72  0.4749
## Year2005         0.00354    0.26996    0.01  0.9896
```

```

## Year2006          0.02991    0.23632    0.13    0.8997
## Year2007          0.15185    0.31542    0.48    0.6319
## Year2008         -0.00159    0.23160   -0.01    0.9945
## Year2009         -0.10361    0.24452   -0.42    0.6732
## Year2010          0.24876    0.24019    1.04    0.3044
## Year2011         -0.11284    0.31162   -0.36    0.7185
## Year2012         -0.24130    0.24688   -0.98    0.3322
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.328
## Multiple R-squared:  0.351, Adjusted R-squared:  0.162
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 70 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.242  0.863  0.950  0.898  0.988  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      1.23e-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"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 5.338 1      2.310
## Year              5.338 16      1.054

```

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
## -1.102157 -0.187896 -0.000696 0.164500 0.857304
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.21086 0.21327 5.68 3.7e-07 ***
## FirstAuthorFemale1 -0.17046 0.09185 -1.86 0.0682 .
## Year1997 0.11879 0.25933 0.46 0.6485
## Year1998 0.05514 0.21327 0.26 0.7968
## Year1999 -0.66813 0.22175 -3.01 0.0037 **
## Year2000 -0.35836 0.24526 -1.46 0.1489
## Year2001 -0.22889 0.26241 -0.87 0.3864
## Year2002 -0.67786 0.21440 -3.16 0.0024 **
## Year2003 0.01714 0.47546 0.04 0.9714
## Year2004 0.20724 0.28390 0.73 0.4681
## Year2005 0.00600 0.24748 0.02 0.9807
## Year2006 0.03041 0.23416 0.13 0.8971
```

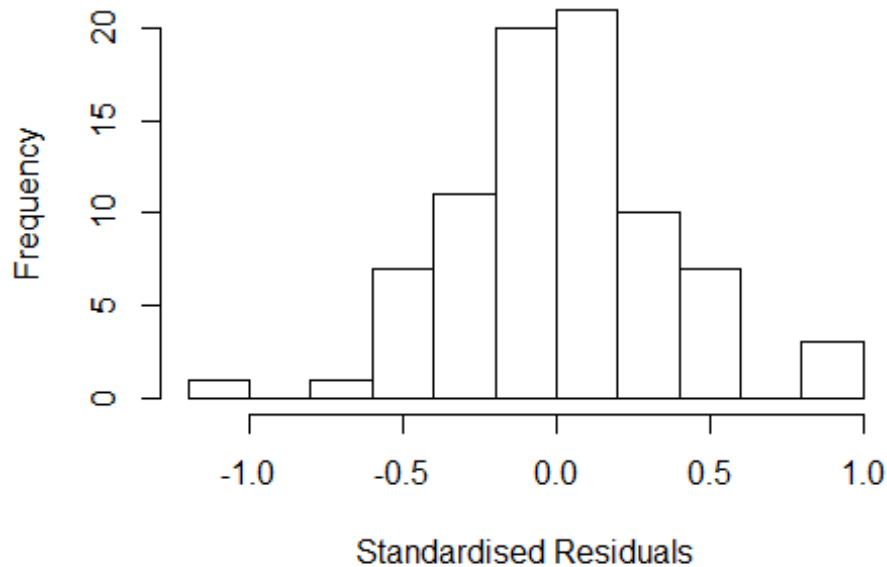


```

## Year2007          0.15927    0.30881    0.52    0.6078
## Year2008          -0.00569    0.22880   -0.02    0.9802
## Year2009          -0.10129    0.23041   -0.44    0.6617
## Year2010           0.25049    0.22752    1.10    0.2751
## Year2011          -0.10871    0.27550   -0.39    0.6945
## Year2012          -0.23567    0.23770   -0.99    0.3253
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.355
## Multiple R-squared:  0.338, Adjusted R-squared:  0.159
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 69 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.314  0.879  0.954  0.910  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      1.23e-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 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 34.94 1          5.911
## Year            34.94 16          1.117

```

Residuals from last author



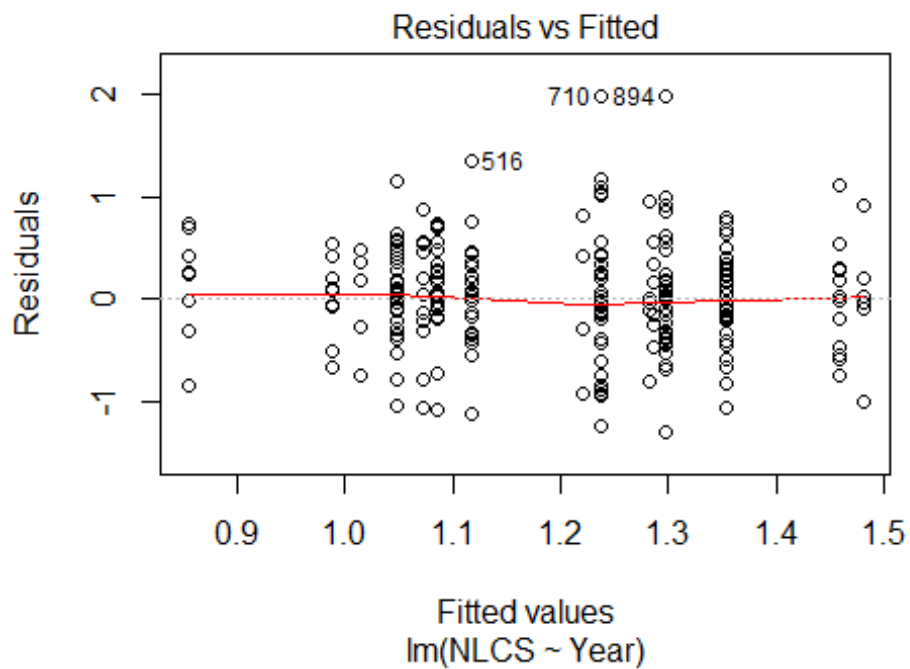
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.991 -0.206 -0.022 0.213 0.798
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.15535 0.27074 4.27 6.8e-05 ***
## LastAuthorFemale1 0.04192 0.10919 0.38 0.702
## Year1997 0.12478 0.30961 0.40 0.688
## Year1998 0.06873 0.29460 0.23 0.816
## Year1999 -0.69785 0.29733 -2.35 0.022 *
## Year2000 -0.32381 0.31018 -1.04 0.300
## Year2001 -0.35905 0.30662 -1.17 0.246
## Year2002 -0.62235 0.27163 -2.29 0.025 *
## Year2003 0.07265 0.51047 0.14 0.887
## Year2004 0.14765 0.36500 0.40 0.687
## Year2005 -0.05014 0.30819 -0.16 0.871
## Year2006 -0.00548 0.27917 -0.02 0.984
```

```

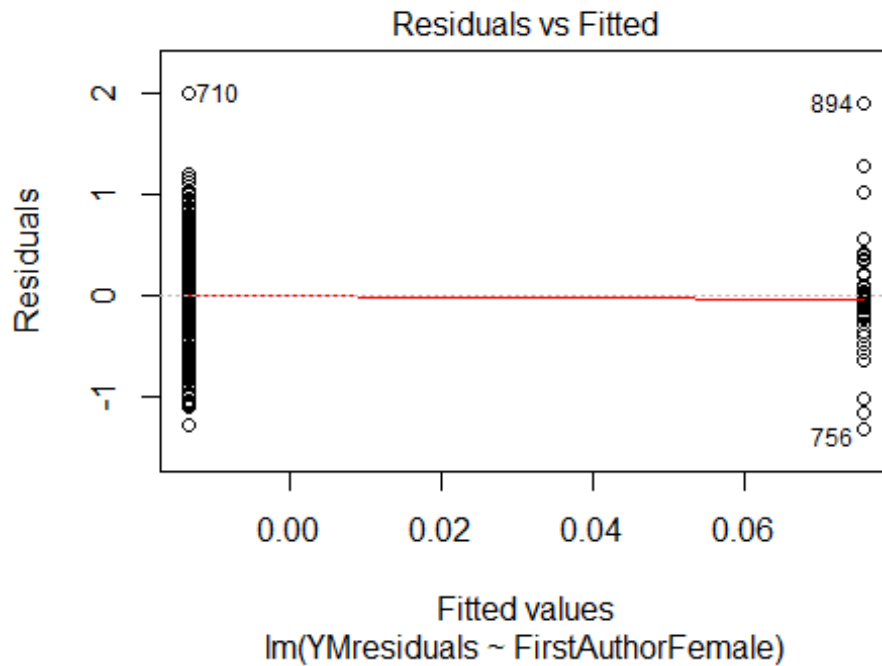
## Year2007      0.06105      0.32819      0.19      0.853
## Year2008     -0.05095      0.28680     -0.18      0.860
## Year2009     -0.19613      0.27972     -0.70      0.486
## Year2010      0.21979      0.29694      0.74      0.462
## Year2011     -0.16431      0.33629     -0.49      0.627
## Year2012     -0.30431      0.29088     -1.05      0.299
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.349
## Multiple R-squared:  0.302, Adjusted R-squared:  0.113
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 73 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.400  0.871  0.963  0.909  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      1.23e-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: 81"
## [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
##   23   27   33   19   36   23   23   29   41   38   29   46   57   65   75
## 2012
##   57
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    5    4    7    5    6   10    9   14   22   13   13   22   27   33   35
## 2012

```

```
## 35
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
## 4 3 6 5 5 9 7 14 20 11 11 20 21 30 30
## 2012
## 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 = 12, df = 15, p-value = 0.6
```

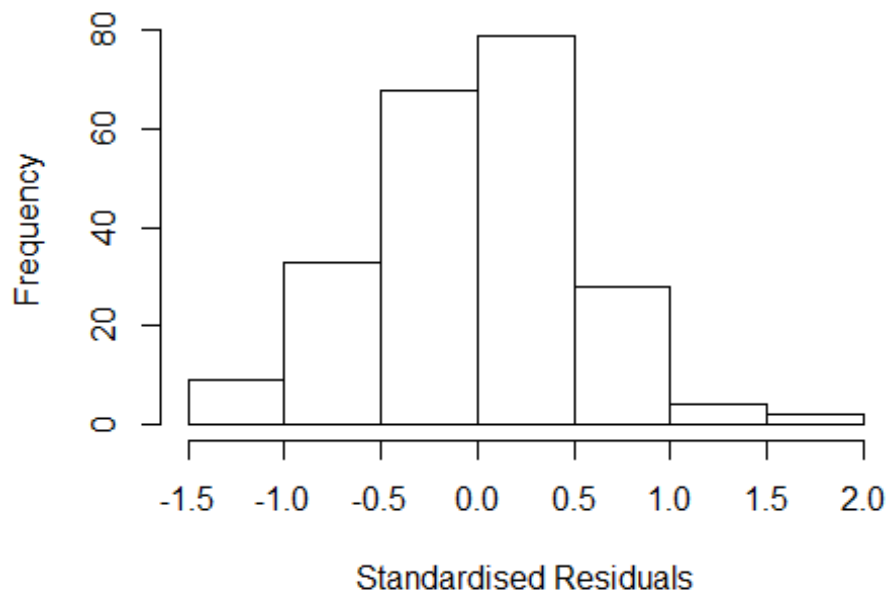


```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.62, df = 1, p-value = 0.4
```



```
## [1] "Female first author team size 2018 geometric mean: NaN"
## [1] "Male first author team size 2018 geometric mean: 2.29378153504588"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.715 1          1.310
## LastAuthorFemale  1.414 1          1.189
## UniqueAuthors    2.728 4          1.134
## Year              4.483 15         1.051
```

Residuals from first and last author and team size



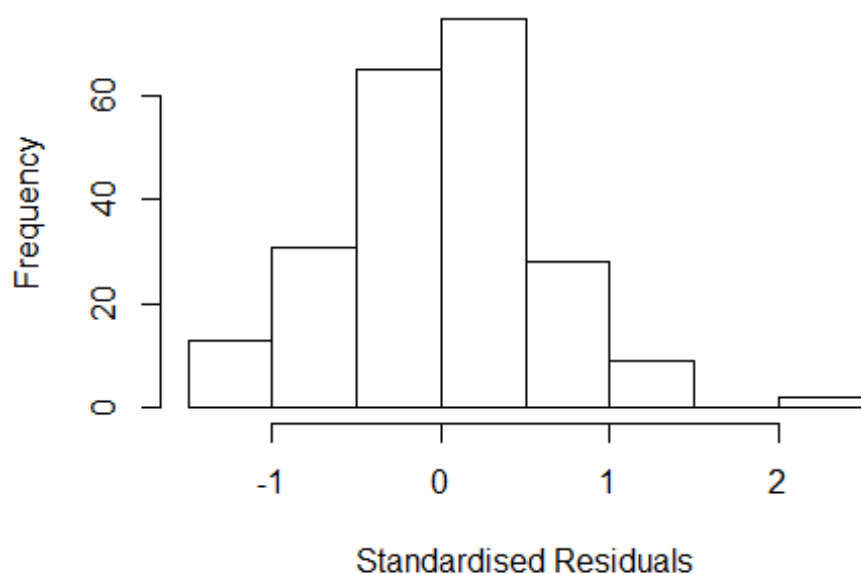
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.19517 -0.36202 0.00739 0.36410 1.78307
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2440 0.3807 3.27 0.0013 **
## FirstAuthorFemale1 0.0223 0.1313 0.17 0.8654
## LastAuthorFemale1 -0.0205 0.1637 -0.13 0.9004
## UniqueAuthors2 0.0910 0.1064 0.85 0.3938
## UniqueAuthors3 0.1764 0.1358 1.30 0.1955
## UniqueAuthors4 0.1997 0.1365 1.46 0.1449
## UniqueAuthors5 0.4714 0.1560 3.02 0.0028 **
## Year1998 0.0910 0.6403 0.14 0.8871
## Year1999 0.0053 0.4231 0.01 0.9900
## Year2000 -0.3279 0.4555 -0.72 0.4725
```

```

## Year2001      0.2392      0.4759      0.50      0.6158
## Year2002     -0.3985      0.4247     -0.94      0.3492
## Year2003     -0.3060      0.4062     -0.75      0.4521
## Year2004     -0.3728      0.4052     -0.92      0.3587
## Year2005     -0.3646      0.3961     -0.92      0.3584
## Year2006     -0.3198      0.4069     -0.79      0.4328
## Year2007      0.1215      0.4066      0.30      0.7654
## Year2008     -0.2343      0.3920     -0.60      0.5507
## Year2009     -0.2837      0.4079     -0.70      0.4875
## Year2010     -0.2428      0.4023     -0.60      0.5469
## Year2011     -0.0737      0.3906     -0.19      0.8506
## Year2012     -0.2252      0.3959     -0.57      0.5701
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.579
## Multiple R-squared:  0.11,   Adjusted R-squared:  0.0175
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 205 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.322  0.884  0.958  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      4.48e-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 1.567 1      1.252
## LastAuthorFemale  1.404 1      1.185
## Year              2.170 15      1.026

```

Residuals from first and last author



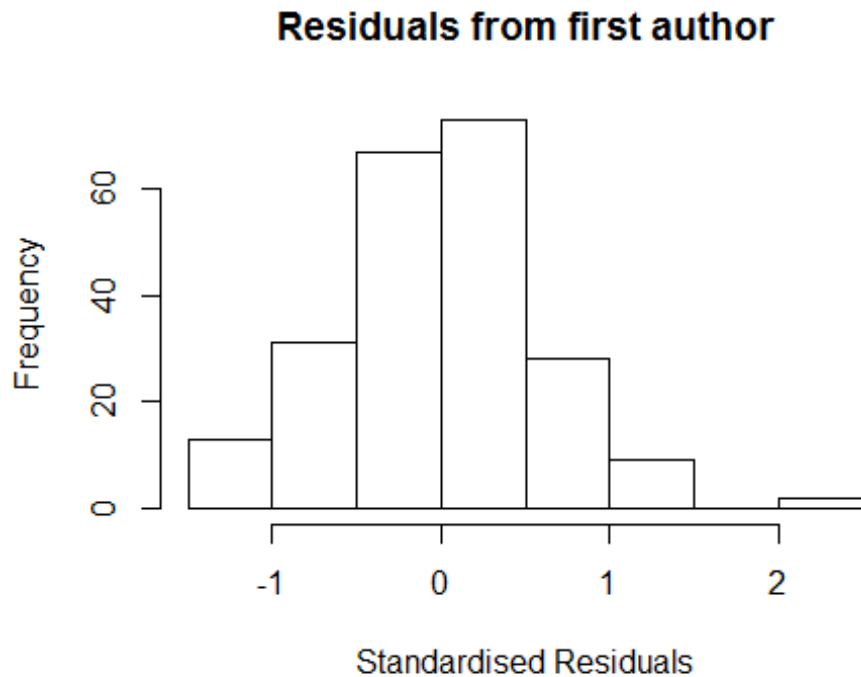
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.20097 -0.34986 0.00413 0.36517 2.05608
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.291024 0.351371 3.67 0.0003 ***
## FirstAuthorFemale1 0.042045 0.129337 0.33 0.7455
## LastAuthorFemale1 0.035522 0.169121 0.21 0.8338
## Year1998 0.132021 0.659985 0.20 0.8417
## Year1999 0.014431 0.384374 0.04 0.9701
## Year2000 -0.278047 0.418527 -0.66 0.5072
## Year2001 0.218029 0.457174 0.48 0.6339
## Year2002 -0.372472 0.410264 -0.91 0.3650
## Year2003 -0.298535 0.390434 -0.76 0.4454
## Year2004 -0.236870 0.375093 -0.63 0.5284
## Year2005 -0.268950 0.372815 -0.72 0.4715
## Year2006 -0.184168 0.396839 -0.46 0.6431
```



```

## Year2007          0.212148    0.383423    0.55    0.5807
## Year2008          -0.197787    0.374650   -0.53    0.5981
## Year2009          -0.187196    0.386703   -0.48    0.6288
## Year2010          -0.132102    0.378115   -0.35    0.7272
## Year2011           0.000663    0.364643    0.00    0.9986
## Year2012          -0.101229    0.370118   -0.27    0.7847
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.569
## Multiple R-squared:  0.0612, Adjusted R-squared:  -0.0166
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 203 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.164  0.876  0.957  0.906  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      4.48e-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 1.547 1          1.244
## Year              1.547 15          1.015

```



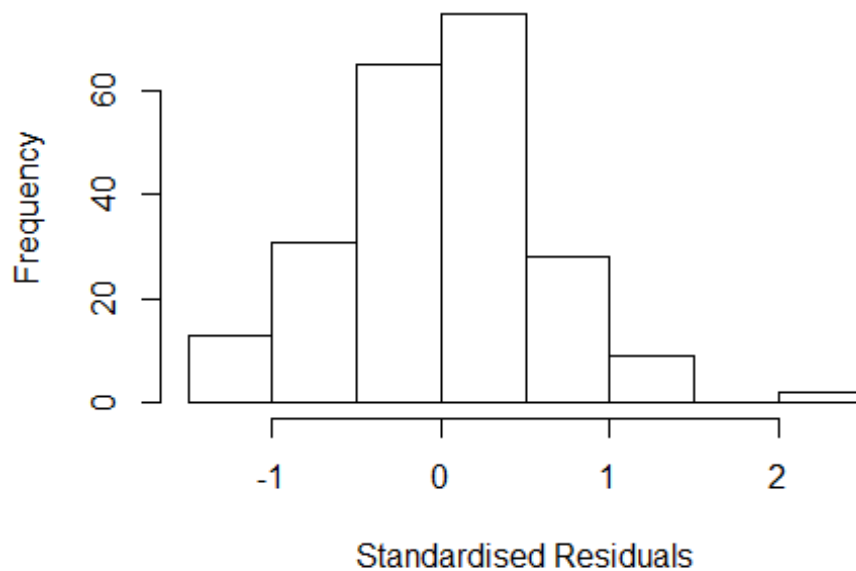
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.20587 -0.34955 0.00187 0.36530 2.05637
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2909 0.3516 3.67 0.00031 ***
## FirstAuthorFemale1 0.0472 0.1302 0.36 0.71707
## Year1998 0.1330 0.6605 0.20 0.84067
## Year1999 0.0138 0.3849 0.04 0.97141
## Year2000 -0.2800 0.4188 -0.67 0.50455
## Year2001 0.2183 0.4574 0.48 0.63373
## Year2002 -0.3722 0.4105 -0.91 0.36567
## Year2003 -0.2984 0.3907 -0.76 0.44590
## Year2004 -0.2259 0.3687 -0.61 0.54068
## Year2005 -0.2670 0.3729 -0.72 0.47483
## Year2006 -0.1839 0.3971 -0.46 0.64376
## Year2007 0.2186 0.3821 0.57 0.56791
```

```

## Year2008          -0.1948      0.3745   -0.52  0.60357
## Year2009          -0.1860      0.3865   -0.48  0.63088
## Year2010          -0.1323      0.3786   -0.35  0.72714
## Year2011           0.0026      0.3647    0.01  0.99432
## Year2012          -0.1014      0.3703   -0.27  0.78442
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.567
## Multiple R-squared:  0.0611, Adjusted R-squared:  -0.0119
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 201 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.161  0.872  0.955  0.905  0.987  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      4.48e-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 1.409 1          1.187
## Year            1.409 15          1.011

```

Residuals from last author



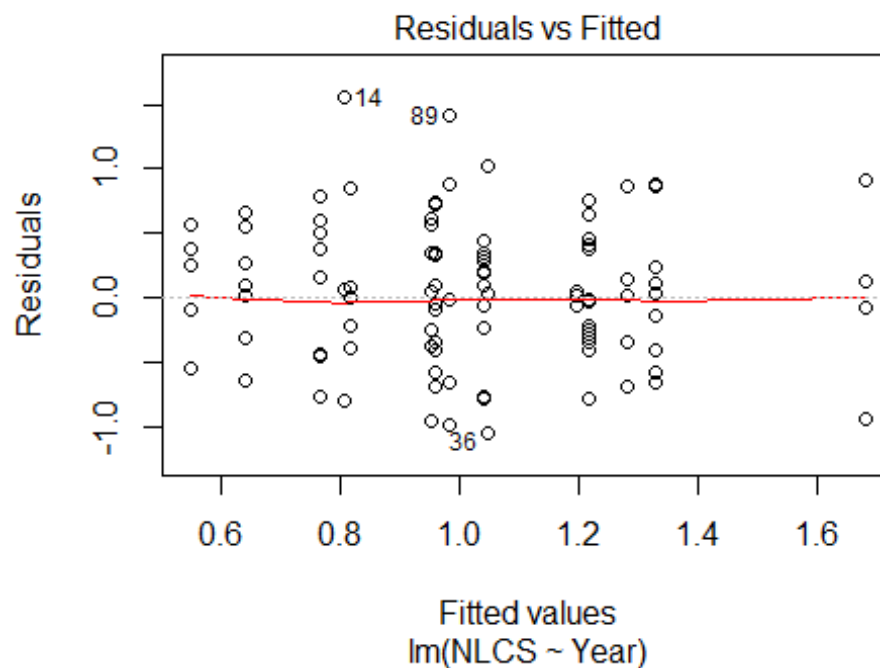
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.19179 -0.35113 0.00195 0.35938 2.05552
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2909 0.3517 3.67 0.00031 ***
## LastAuthorFemale1 0.0461 0.1720 0.27 0.78887
## Year1998 0.1333 0.6608 0.20 0.84039
## Year1999 0.0207 0.3831 0.05 0.95699
## Year2000 -0.2601 0.4140 -0.63 0.53054
## Year2001 0.2184 0.4575 0.48 0.63367
## Year2002 -0.3721 0.4106 -0.91 0.36592
## Year2003 -0.2983 0.3907 -0.76 0.44608
## Year2004 -0.2364 0.3756 -0.63 0.52971
## Year2005 -0.2667 0.3728 -0.72 0.47516
## Year2006 -0.1838 0.3972 -0.46 0.64398
## Year2007 0.2146 0.3839 0.56 0.57679
```

```

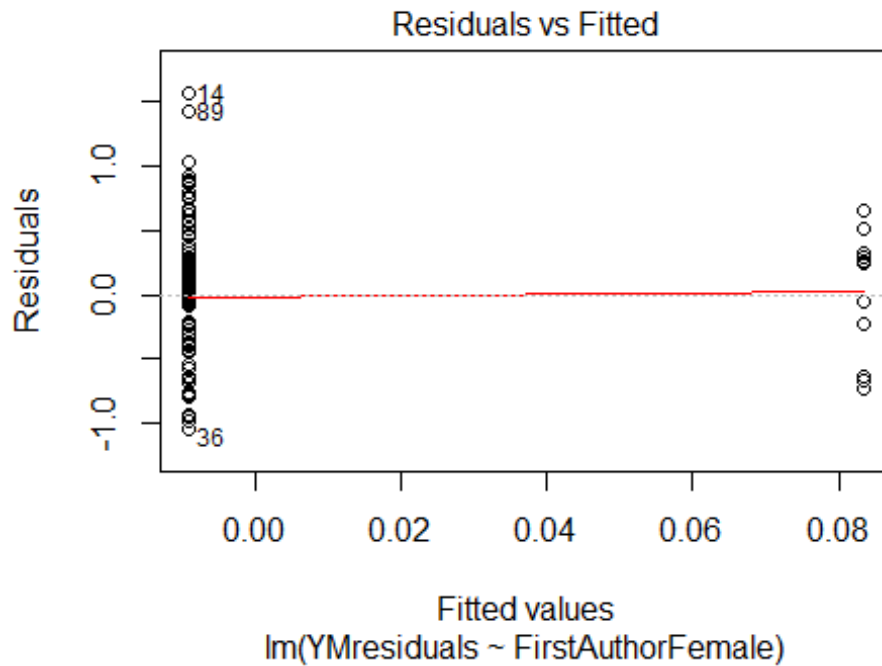
## Year2008          -0.1922      0.3750   -0.51  0.60889
## Year2009          -0.1736      0.3809   -0.46  0.64907
## Year2010          -0.1314      0.3785   -0.35  0.72883
## Year2011           0.0131      0.3620    0.04  0.97118
## Year2012          -0.0991      0.3703   -0.27  0.78930
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.567
## Multiple R-squared:  0.0608, Adjusted R-squared:  -0.0121
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 203 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.161  0.877  0.956  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      4.48e-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: 223"
## [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
##   10   8   7  17  16  12  20  10  17  21  25  30  32  12  22
## 2011 2012
##   23  17
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3   4   3   4   8   6   6   3   6  11   5   7  11   7  12
## 2011 2012
##   11   3

```

```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    4    3    4    7    6    5    2    6    4    4    4   10    6   10
## 2011 2012
##   10    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
```



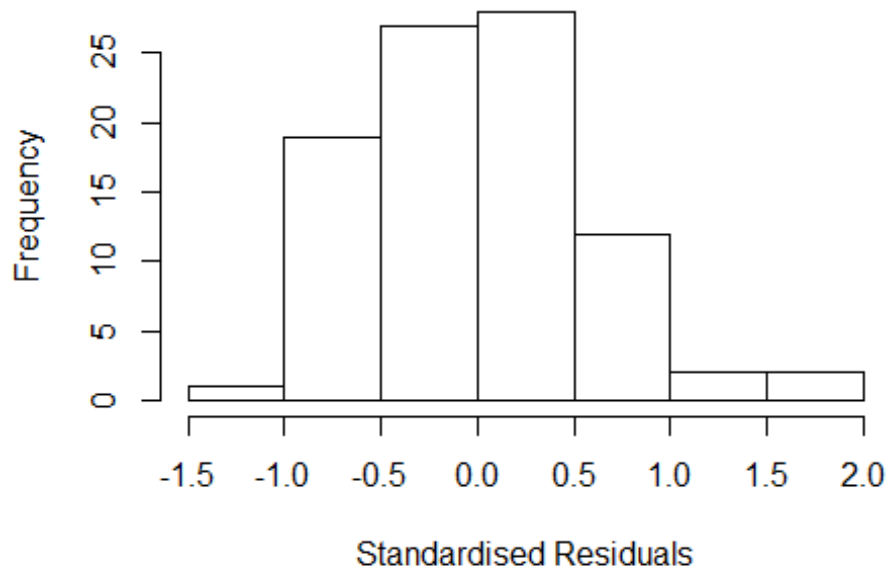
```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.19, df = 1, p-value = 0.7
```



```
## [1] "Female first author team size 2018 geometric mean: NaN"
## [1] "Male first author team size 2018 geometric mean: 1.64375182951723"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
```

| | GVIF | Df | GVIF^(1/(2*Df)) |
|----------------------|--------|----|-----------------|
| ## FirstAuthorFemale | 4.142 | 1 | 2.035 |
| ## LastAuthorFemale | 5.824 | 1 | 2.413 |
| ## UniqueAuthors | 18.335 | 4 | 1.438 |
| ## Year | 34.479 | 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
## -1.0609 -0.4091 -0.0252  0.3703  1.6575
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.21831    0.19870     6.13   5e-08 ***
## FirstAuthorFemale1 -0.07479    0.30844    -0.24   0.809
## LastAuthorFemale1  0.15308    0.52369     0.29   0.771
## UniqueAuthors2   -0.00316    0.22923    -0.01   0.989
## UniqueAuthors3    0.34520    0.23469     1.47   0.146
## UniqueAuthors4    0.12758    0.26712     0.48   0.634
## UniqueAuthors5    0.52854    0.22482     2.35   0.022 *
## Year1997         -0.66171    0.70069    -0.94   0.348
## Year1998         -0.15740    0.69600    -0.23   0.822
## Year1999          0.46970    0.41752     1.12   0.265
```

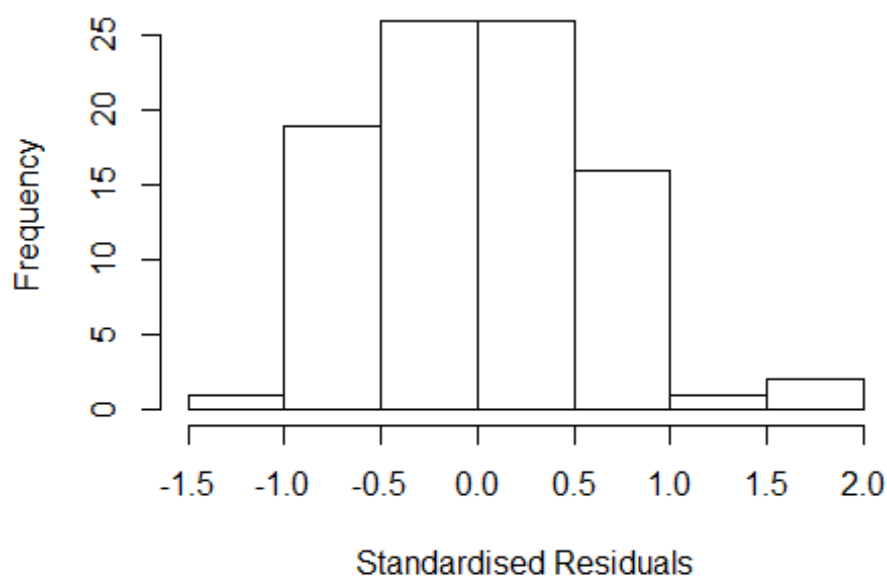


```

## Year2000      -0.59066    0.27084   -2.18    0.033 *
## Year2001      -0.34256    0.51156   -0.67    0.505
## Year2002      -0.34519    0.35007   -0.99    0.328
## Year2003      -0.02873    0.17974   -0.16    0.873
## Year2004      -0.69288    0.28409   -2.44    0.017 *
## Year2005      -0.64266    0.26302   -2.44    0.017 *
## Year2006      -0.10212    0.26624   -0.38    0.702
## Year2007      -0.45092    0.35122   -1.28    0.204
## Year2008      -0.24386    0.25853   -0.94    0.349
## Year2009      -0.52288    0.32913   -1.59    0.117
## Year2010      -0.42695    0.22676   -1.88    0.064 .
## Year2011       0.02451    0.27852    0.09    0.930
## Year2012      -0.52090    0.37825   -1.38    0.173
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.656
## Multiple R-squared:  0.248, Adjusted R-squared:  0.00422
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 84 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.503  0.923  0.962  0.932  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      1.10e-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 5.442 1      2.333
## LastAuthorFemale  6.623 1      2.573
## Year              3.010 16      1.035

```

Residuals from first and last author



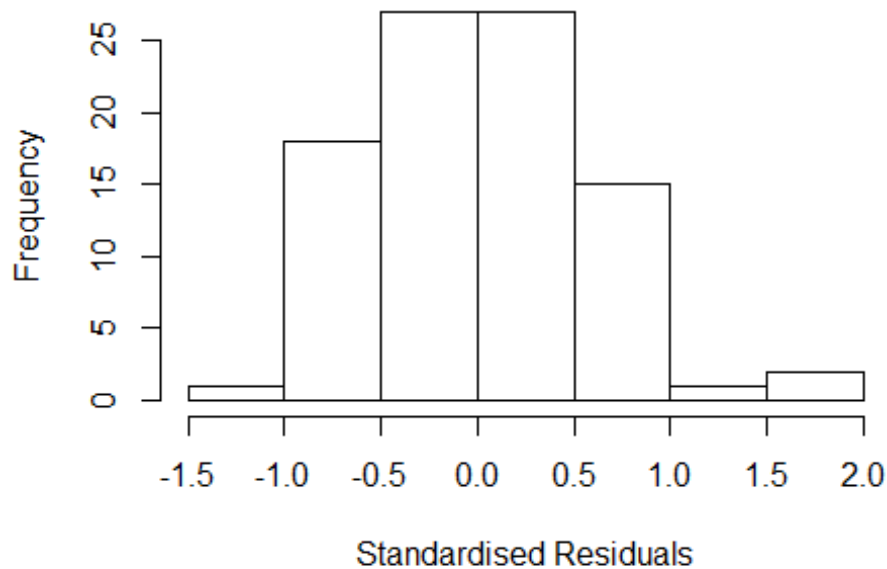
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.057 -0.426 -0.015 0.420 1.600
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.21729 0.17441 6.98 1.2e-09 ***
## FirstAuthorFemale1 0.12966 0.32358 0.40 0.6898
## LastAuthorFemale1 0.13753 0.48241 0.29 0.7764
## Year1997 -0.59127 0.58995 -1.00 0.3196
## Year1998 -0.16056 0.64830 -0.25 0.8051
## Year1999 0.46942 0.40187 1.17 0.2466
## Year2000 -0.59057 0.26789 -2.20 0.0307 *
## Year2001 -0.32494 0.48724 -0.67 0.5070
## Year2002 -0.34725 0.34510 -1.01 0.3177
## Year2003 -0.02929 0.17855 -0.16 0.8702
## Year2004 -0.75332 0.27540 -2.74 0.0078 **
## Year2005 -0.46904 0.31984 -1.47 0.1469
```

```

## Year2006          0.00598    0.35797    0.02    0.9867
## Year2007          -0.40895    0.31783   -1.29    0.2023
## Year2008          -0.03216    0.24113   -0.13    0.8943
## Year2009          -0.35184    0.27956   -1.26    0.2123
## Year2010          -0.28240    0.23478   -1.20    0.2330
## Year2011           0.09612    0.25625    0.38    0.7087
## Year2012          -0.47583    0.38922   -1.22    0.2255
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.696
## Multiple R-squared:  0.199, Adjusted R-squared:  -0.000932
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 80 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.576  0.912  0.954  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      1.10e-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 1.444 1      1.202
## Year              1.444 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
## -1.0572 -0.3956 -0.0159 0.4197 1.7633
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.21726 0.17458 6.97 1.2e-09 ***
## FirstAuthorFemale1 0.19785 0.17384 1.14 0.2588
## Year1997 -0.61657 0.59734 -1.03 0.3054
## Year1998 -0.16003 0.65491 -0.24 0.8076
## Year1999 0.46969 0.40325 1.16 0.2479
## Year2000 -0.59054 0.26850 -2.20 0.0310 *
## Year2001 -0.33000 0.49243 -0.67 0.5049
## Year2002 -0.34650 0.34655 -1.00 0.3207
## Year2003 -0.02926 0.17872 -0.16 0.8704
## Year2004 -0.73083 0.26613 -2.75 0.0076 **
## Year2005 -0.46902 0.32093 -1.46 0.1482
## Year2006 0.00493 0.35890 0.01 0.9891
```

```

## Year2007          -0.40992    0.31839   -1.29    0.2020
## Year2008          -0.02467    0.24079   -0.10    0.9187
## Year2009          -0.35150    0.28003   -1.26    0.2134
## Year2010          -0.28837    0.23283   -1.24    0.2195
## Year2011           0.10434    0.25214    0.41    0.6802
## Year2012          -0.49833    0.36823   -1.35    0.1801
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.683
## Multiple R-squared:  0.202, Adjusted R-squared:  0.0166
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 80 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.485  0.910  0.954  0.930  0.984  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      1.10e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.958 1      1.399
## Year              1.958 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
## -1.0563 -0.4412 -0.0247  0.4207  1.4974

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.21732    0.17426   6.99 1.1e-09 ***
## LastAuthorFemale1 0.23633    0.29365   0.80  0.424
## Year1997       -0.58184    0.54545  -1.07  0.290
## Year1998       -0.16100    0.64208  -0.25  0.803
## Year1999        0.46918    0.40069   1.17  0.245
## Year2000       -0.59060    0.26736  -2.21  0.030 *
## Year2001       -0.32073    0.48197  -0.67  0.508
## Year2002       -0.34789    0.34374  -1.01  0.315
## Year2003       -0.02932    0.17840  -0.16  0.870
## Year2004       -0.74363    0.27690  -2.69  0.009 **
## Year2005       -0.46907    0.31903  -1.47  0.146
## Year2006        0.00687    0.35706   0.02  0.985
## Year2007       -0.40813    0.31724  -1.29  0.202
## Year2008       -0.02846    0.24123  -0.12  0.906
## Year2009       -0.35214    0.27913  -1.26  0.211
## Year2010       -0.25397    0.23676  -1.07  0.287
## Year2011        0.11735    0.24913   0.47  0.639
## Year2012       -0.43323    0.40368  -1.07  0.287
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.708
## Multiple R-squared:  0.196, Adjusted R-squared:  0.00836
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 82 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.634  0.914  0.956  0.937  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      1.10e-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 3102"
## [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 2006 2007 2008 2009 2010 2011 2012
##    1    6    4    6    1    1    2    5    7    3    2   11    6   14    5
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2006 2007 2008 2009 2010 2011 2012
##    0    0    1    1    0    0    1    1    4    3    2    6    1    9    4
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2006 2007 2008 2009 2010 2011 2012
##    0    0    1    1    0    0    1    1    4    2    2    6    1    8    2
## [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: 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 = 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: 4"
## [1] "Male last author team size 2018 geometric mean: 2.78315768371374"

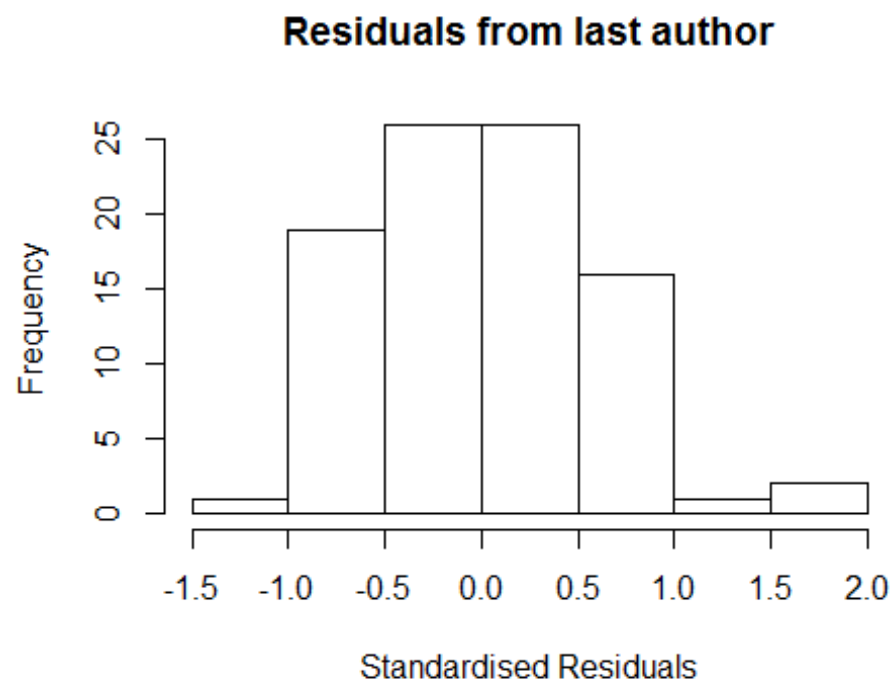
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2.5, 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"

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

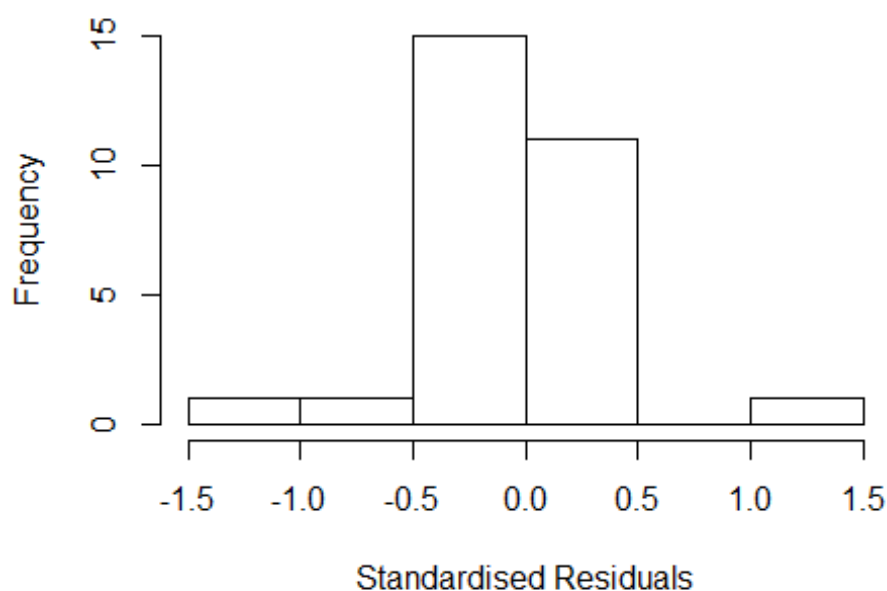
```

```
## Warning in cov2cor(v): diag(.) had 0 or NA entries; non-finite result is
## doubtful
```



| | | | |
|----------------------|------|----|-----------------|
| ## | GVIF | Df | GVIF^(1/(2*Df)) |
| ## FirstAuthorFemale | NaN | 1 | NaN |
| ## LastAuthorFemale | NaN | 1 | NaN |
| ## Year | NaN | 10 | 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
## -1.34e+00 -1.59e-01 -7.77e-16 1.58e-01 1.17e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4650 0.0000 Inf < 2e-16 ***
## FirstAuthorFemale1 0.3483 0.3028 1.15 0.26702
## LastAuthorFemale1 0.2096 0.3173 0.66 0.51818
## Year2000 -0.0560 0.0000 -Inf < 2e-16 ***
## Year2003 -0.8170 0.0000 -Inf < 2e-16 ***
## Year2004 -0.6970 0.0000 -Inf < 2e-16 ***
## Year2006 -0.4364 0.0929 -4.70 0.00024 ***
## Year2007 0.0095 0.0110 0.87 0.39901
## Year2008 -0.4280 0.1159 -3.69 0.00198 **
## Year2009 -0.1235 0.1465 -0.84 0.41140
## Year2010 -0.0790 0.0000 -Inf < 2e-16 ***
## Year2011 -0.4286 0.3422 -1.25 0.22846
```

```

## Year2012          -0.4740      0.1686   -2.81  0.01252 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.392
## Multiple R-squared:  0.364, Adjusted R-squared:  -0.113
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 20 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.218  0.934  0.971  0.891  0.985  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           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 3952 1          62.862
## Year              3952 10          1.513

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##   control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##           Min           1Q       Median           3Q           Max
## -1.36e+00 -1.59e-01  1.11e-16  1.09e-01  1.10e+00
##
## Coefficients:
##           Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   1.47e+00   1.11e-08  1.32e+08 < 2e-16 ***
## FirstAuthorFemale1 3.81e-01  3.33e-01  1.14e+00  0.26925

```

```

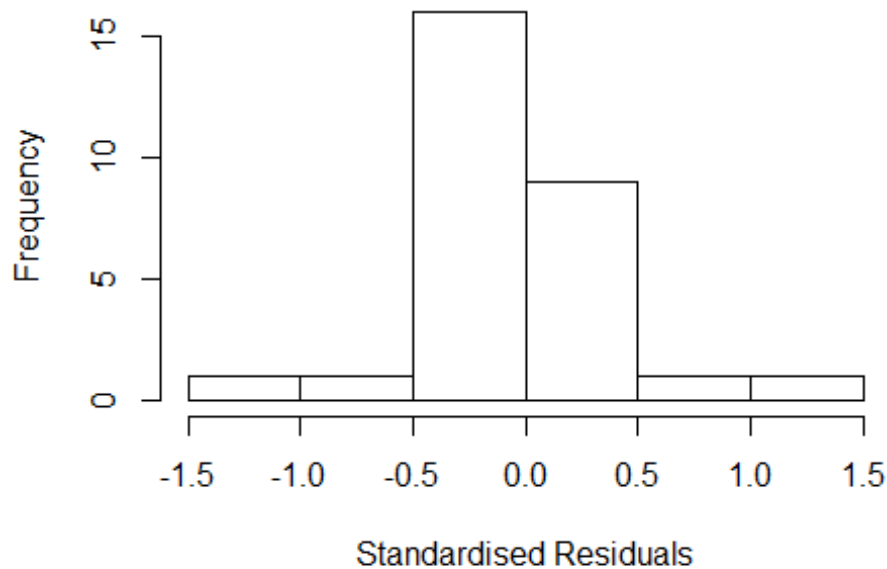
## Year2000      -5.60e-02  1.28e-08 -4.38e+06 < 2e-16 ***
## Year2003      -8.17e-01  6.40e-09 -1.28e+08 < 2e-16 ***
## Year2004      -6.97e-01  4.52e-09 -1.54e+08 < 2e-16 ***
## Year2006      -4.36e-01  9.32e-02 -4.68e+00 0.00022 ***
## Year2007       9.50e-03  1.10e-02  8.70e-01 0.39828
## Year2008      -4.28e-01  1.16e-01 -3.68e+00 0.00187 **
## Year2009      -1.02e-01  1.30e-01 -7.90e-01 0.44171
## Year2010      -7.90e-02  3.21e-09 -2.46e+07 < 2e-16 ***
## Year2011      -3.56e-01  3.33e-01 -1.07e+00 0.30013
## Year2012      -4.74e-01  1.70e-01 -2.79e+00 0.01263 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.368
## Multiple R-squared:  0.374, Adjusted R-squared:  -0.0308
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
##  11 weights are ~ = 1. The remaining 18 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.142  0.933   0.965   0.865   0.981   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      3.45e-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 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q      Median        3Q       Max
## -1.31e+00 -1.59e-01 -1.11e-16  1.59e-01  1.03e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.4650     0.0000   Inf <2e-16 ***
## LastAuthorFemale1  0.2118     0.3977   0.53  0.6013
## Year2000          -0.0560     0.0000  -Inf <2e-16 ***
## Year2003          -0.8170     0.0000  -Inf <2e-16 ***
## Year2004          -0.6970     0.0000  -Inf <2e-16 ***
## Year2006          -0.4368     0.0925  -4.72  0.0002 ***
## Year2007           0.0095     0.0110   0.87  0.3982
## Year2008          -0.4280     0.1154  -3.71  0.0017 **
```

```

## Year2009          -0.1510      0.1748   -0.86   0.3997
## Year2010          -0.0790      0.0000   -Inf    <2e-16 ***
## Year2011          -0.2897      0.3204   -0.90   0.3786
## Year2012          -0.4740      0.1669   -2.84   0.0113 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.427
## Multiple R-squared:  0.28,   Adjusted R-squared:  -0.186
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 20 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.323  0.933  0.967  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      3.45e-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 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
##    7    7    2    5    5    3    5    6    9    7    7    9    9   12   11
## 2011 2012
##   12    8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    2    0    1    0    0    2    3    3    2    6    8    5    6    6
## 2011 2012
##    4    2
##

```

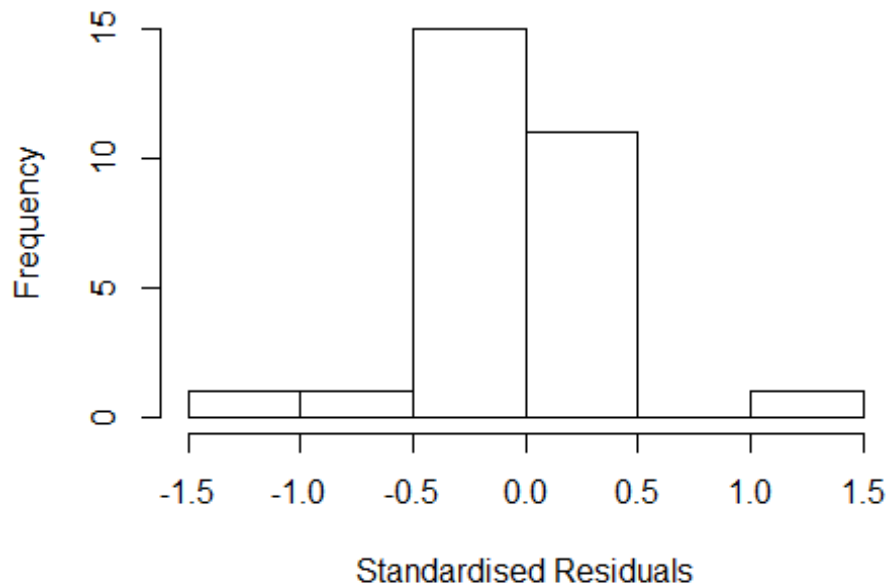
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    2    0    1    0    0    0    3    3    2    6    8    4    5    5
## 2011 2012
##    4    2
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.73205080756888"
## [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 = 3.5, p-value = 1
## 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: 2.21336383940064"

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

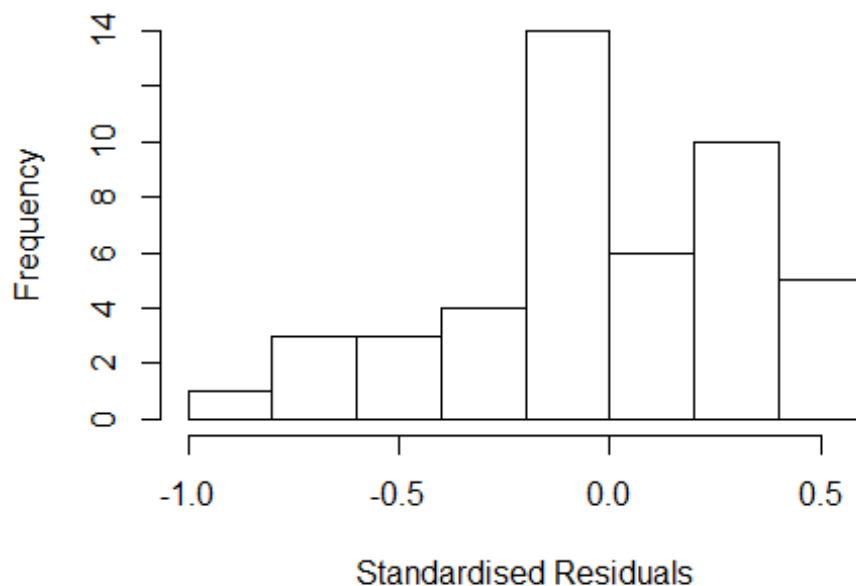
Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
```

```
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3.5, 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 -4.252e+14  1          NaN
## LastAuthorFemale  -9.365e+14  1          NaN
## UniqueAuthors    -1.998e+31  4          NaN
## Year              4.074e+29 12          17.13
```

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
## -8.29e-01 -1.85e-01 -2.01e-16  2.49e-01  5.87e-01
```

```

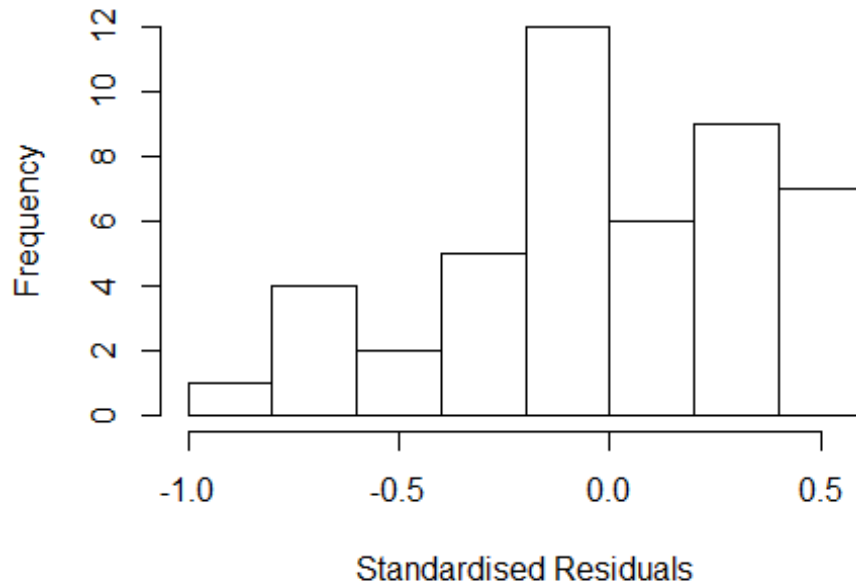
##
## Coefficients:
##               Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.07e+00   1.45e-08  7.42e+07 < 2e-16 ***
## FirstAuthorFemale1 1.34e-01   2.66e-01   5.00e-01  0.61812
## LastAuthorFemale1 -5.08e-01   4.24e-01  -1.20e+00  0.24219
## UniqueAuthors2    1.91e-01   1.75e-01   1.09e+00  0.28448
## UniqueAuthors3    7.39e-02   3.59e-01   2.10e-01  0.83839
## UniqueAuthors4    7.18e-01   3.91e-01   1.84e+00  0.07696 .
## UniqueAuthors5   -1.98e-01   3.90e-01  -5.10e-01  0.61643
## Year1997          7.59e-03   2.05e-01   4.00e-02  0.97069
## Year1999         -2.01e-01   2.99e-01  -6.70e-01  0.50800
## Year2003         -1.01e-01   3.37e-01  -3.00e-01  0.76637
## Year2004          2.42e-01   1.09e-01   2.23e+00  0.03450 *
## Year2005         -5.76e-01   2.02e-01  -2.84e+00  0.00842 **
## Year2006         -3.20e-01   2.56e-01  -1.25e+00  0.22119
## Year2007         -4.17e-01   2.51e-01  -1.66e+00  0.10808
## Year2008         -4.72e-01   2.83e-01  -1.67e+00  0.10652
## Year2009         -4.86e-01   1.08e-01  -4.50e+00  0.00012 ***
## Year2010         -1.01e-01   1.99e-01  -5.10e-01  0.61657
## Year2011         -1.14e-01   2.72e-01  -4.20e-01  0.67781
## Year2012         -4.86e-02   2.46e-01  -2.00e-01  0.84466
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.477
## Multiple R-squared:  0.346, Adjusted R-squared:  -0.0894
## 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.
##   0.744  0.913  0.969   0.942   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 2: First author gender, Last author gender, Year as

```



```
factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -1.032e+15  1      NaN
## LastAuthorFemale  -2.545e+15  1      NaN
## Year              -4.949e+15 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
## -8.47e-01 -2.03e-01  7.63e-16  2.73e-01  5.43e-01
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.07e+00   3.29e-08  3.26e+07 < 2e-16 ***
## FirstAuthorFemale1 -6.25e-02   1.65e-01 -3.80e-01  0.70780
## LastAuthorFemale1 -2.84e-01   2.59e-01 -1.09e+00  0.28253
## Year1997         -5.96e-03   2.06e-01 -3.00e-02  0.97708
## Year1999         -2.28e-01   2.82e-01 -8.10e-01  0.42454
## Year2003         -3.69e-02   3.40e-01 -1.10e-01  0.91420
```

```

## Year2004          3.94e-01  9.77e-02  4.04e+00  0.00033 ***
## Year2005          -4.80e-01  2.67e-01 -1.80e+00  0.08171 .
## Year2006          -2.28e-01  2.43e-01 -9.40e-01  0.35674
## Year2007          -2.97e-01  1.99e-01 -1.49e+00  0.14562
## Year2008          -3.95e-01  2.24e-01 -1.76e+00  0.08764 .
## Year2009          -3.76e-01  1.14e-01 -3.30e+00  0.00241 **
## Year2010           1.12e-01  1.03e-01  1.09e+00  0.28575
## Year2011           4.33e-02  2.52e-01  1.70e-01  0.86471
## Year2012           8.67e-02  1.42e-01  6.10e-01  0.54455
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.475
## Multiple R-squared:  0.291, Adjusted R-squared:  -0.0286
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 3 weights are ~= 1. The remaining 43 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.731  0.916  0.974   0.942   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 3: First author gender, Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.501e+14  1          1.225e+07
## Year              1.501e+14 12          3.897e+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"

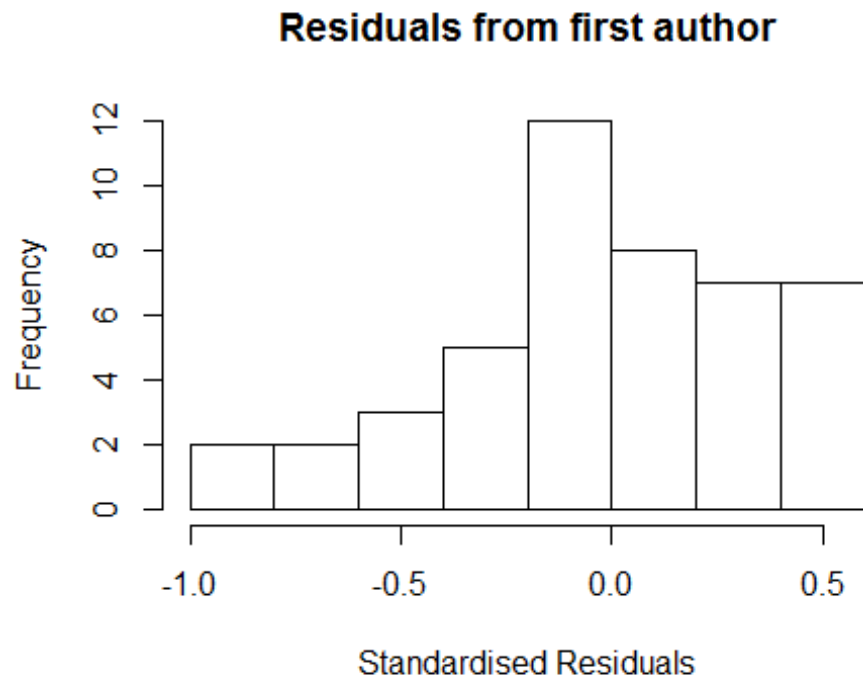
```

```

## Residuals:
##      Min      1Q    Median      3Q      Max
## -9.45e-01 -2.06e-01 -7.77e-16  2.84e-01  5.56e-01
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.07e+00   6.39e-09  1.68e+08  <2e-16 ***
## FirstAuthorFemale1 -1.69e-01   2.07e-01 -8.20e-01   0.420
## Year1997        -9.45e-02   2.48e-01 -3.80e-01   0.706
## Year1999        -4.05e-01   2.07e-01 -1.96e+00   0.059 .
## Year2003        -1.06e-01   3.06e-01 -3.40e-01   0.733
## Year2004         4.28e-01   1.33e-01  3.21e+00   0.003 **
## Year2005        -4.80e-01   2.70e-01 -1.78e+00   0.085 .
## Year2006        -2.24e-01   2.52e-01 -8.90e-01   0.381
## Year2007        -2.95e-01   2.05e-01 -1.44e+00   0.159
## Year2008        -4.41e-01   2.05e-01 -2.15e+00   0.039 *
## Year2009        -3.61e-01   1.36e-01 -2.65e+00   0.012 *
## Year2010         7.70e-02   7.36e-02  1.05e+00   0.303
## Year2011         3.86e-02   3.19e-01  1.20e-01   0.905
## Year2012         1.40e-01   1.28e-01  1.09e+00   0.283
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.449
## Multiple R-squared:  0.278, Adjusted R-squared:  -0.0154
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~ = 1. The remaining 41 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.638  0.902  0.961  0.930  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      2.17e-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"

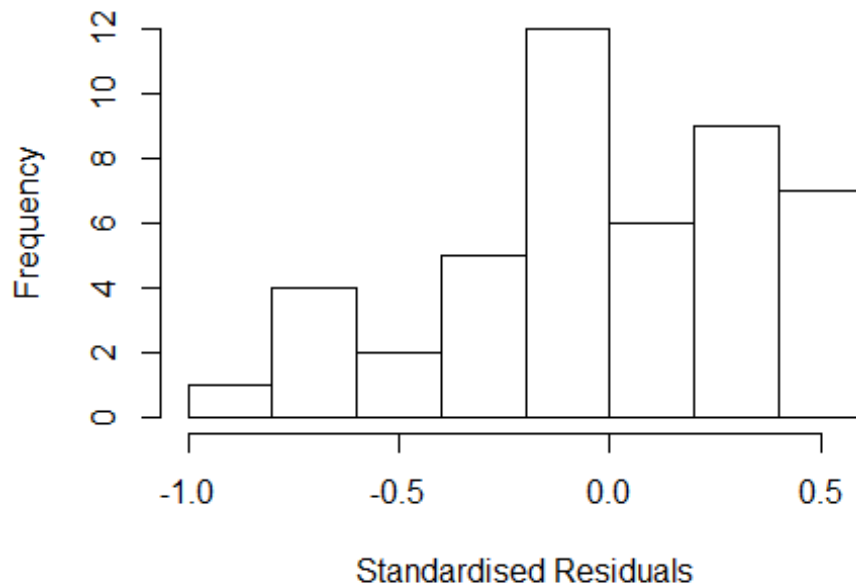
```

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



```
##               GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.325e+14  1      1.151e+07
## Year             1.325e+14 12      3.876e+00
```

Residuals from last author



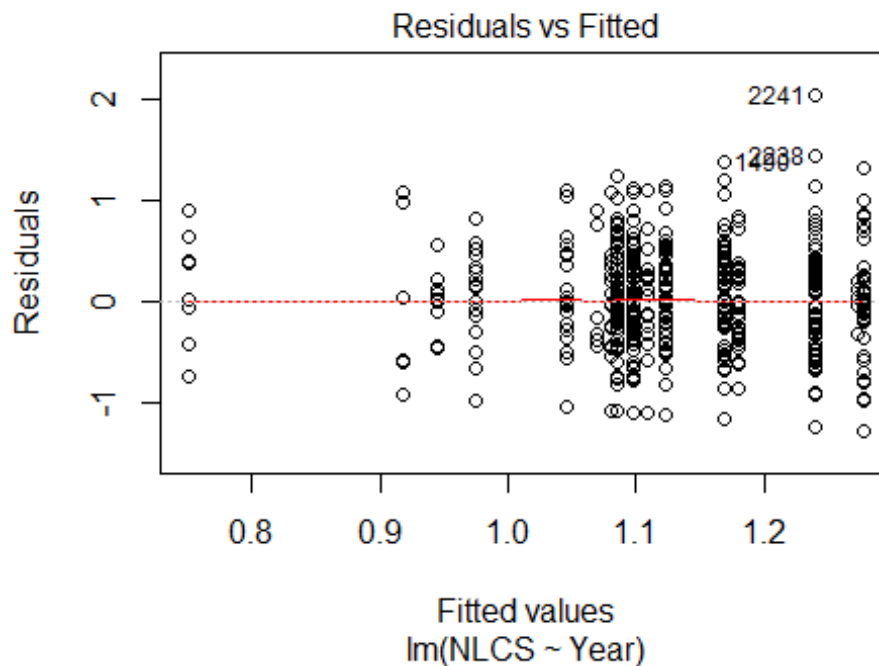
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -8.48e-01 -2.09e-01 5.00e-16 2.67e-01 5.44e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0750 0.0000 Inf < 2e-16 ***
## LastAuthorFemale1 -0.3249 0.2734 -1.19 0.2434
## Year1997 -0.0165 0.2093 -0.08 0.9375
## Year1999 -0.2491 0.2734 -0.91 0.3690
## Year2003 -0.0448 0.3343 -0.13 0.8942
## Year2004 0.3741 0.0734 5.10 1.5e-05 ***
## Year2005 -0.4800 0.2672 -1.80 0.0819 .
## Year2006 -0.2272 0.2423 -0.94 0.3553
## Year2007 -0.2967 0.1986 -1.49 0.1450
## Year2008 -0.4000 0.2204 -1.81 0.0789 .
## Year2009 -0.3869 0.1125 -3.44 0.0016 **
## Year2010 0.1076 0.1102 0.98 0.3364
```

```

## Year2011          0.0416      0.2589      0.16      0.8735
## Year2012          0.0555      0.1389      0.40      0.6921
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.471
## Multiple R-squared:  0.29,   Adjusted R-squared:  0.00197
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 3 weights are ~= 1. The remaining 43 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.727  0.917  0.970  0.941  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          2.17e-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 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
##   54   36   46    4   52   75   47   61   77   92   78  125  110  119  143
## 2011 2012
##  157  123
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    9    5    6    0    7   11   16   18   21   30   24   52   41   59   67
## 2011 2012
##   80   71
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    9    5    4    0    5    9   15   15   19   20   22   43   35   55   55

```

```
## 2011 2012
## 68 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 = 16, df = 15, p-value = 0.4
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 5, df = 1, p-value = 0.02

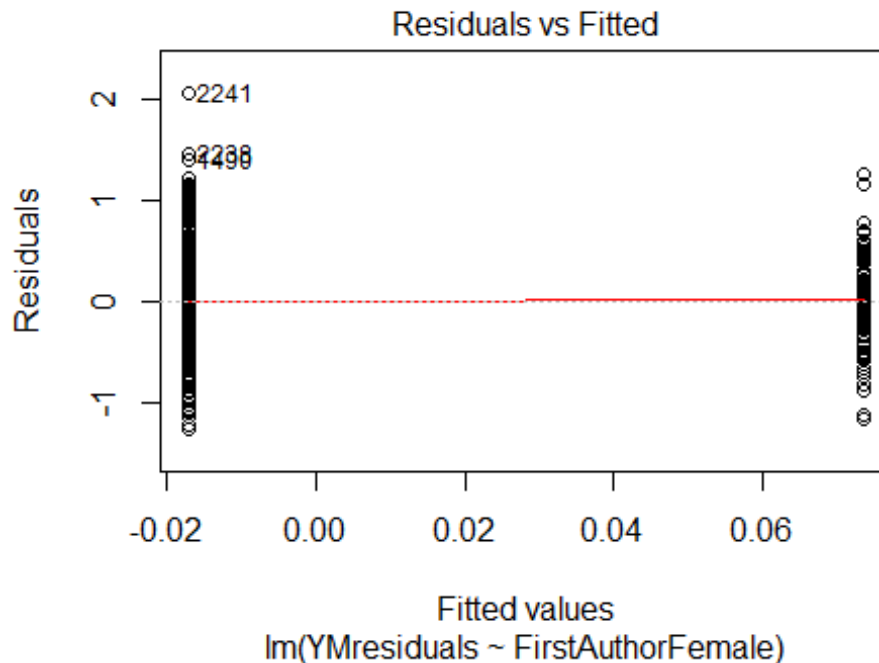
## [1] "Female first author team size 2018 geometric mean: 5.01586747266111"
## [1] "Male first author team size 2018 geometric mean: 3.3492273859029"

## Warning in wilcox.test.default(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 = 0.02
```

```
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.94202163544084"
## [1] "Male last author team size 2018 geometric mean: 3.44638722186446"

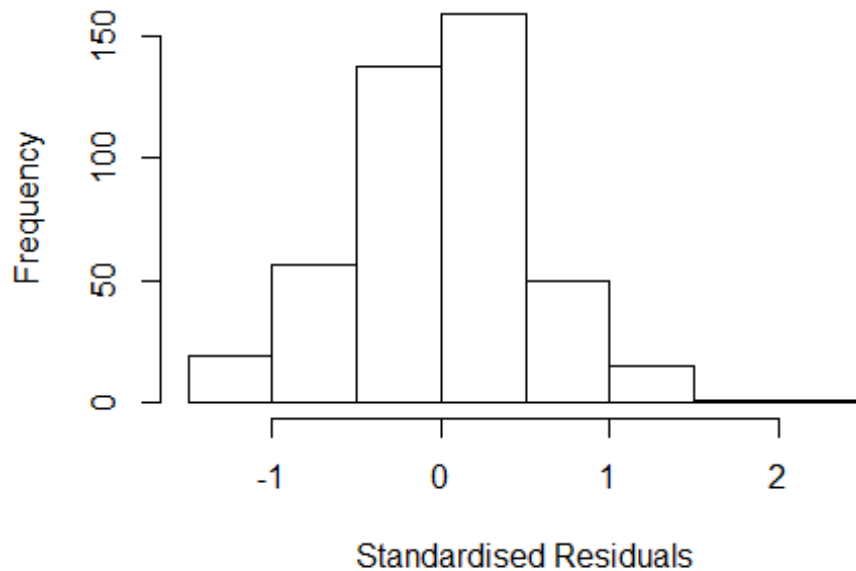
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 230, p-value = 0.06
## 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 | 1.373 | 1 | 1.172 |
| LastAuthorFemale | 1.238 | 1 | 1.112 |
| UniqueAuthors | 1.992 | 4 | 1.090 |
| Year | 2.592 | 15 | 1.032 |

Residuals from first and last author and team size



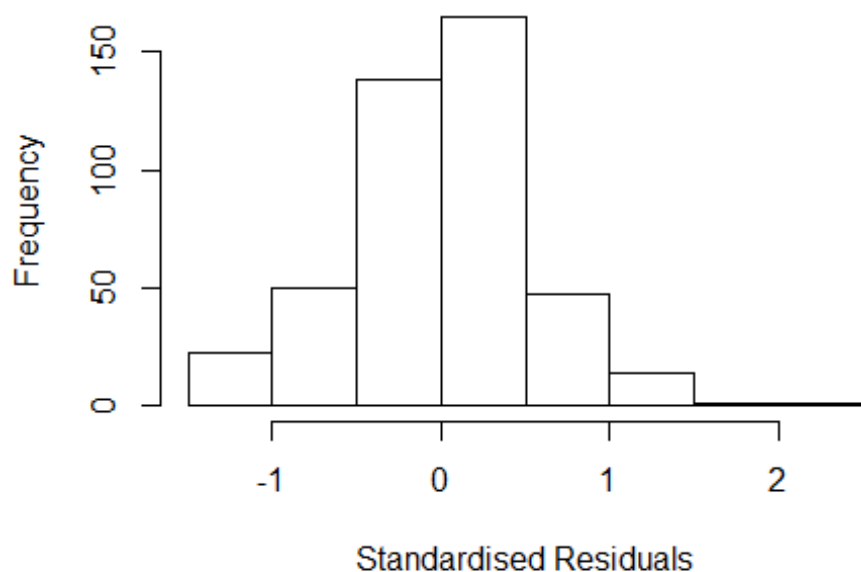
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2014 -0.3072 0.0245 0.3359 2.0756
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7488 0.1455 5.15 4.1e-07 ***
## FirstAuthorFemale1 0.1037 0.0612 1.69 0.0909 .
## LastAuthorFemale1 -0.0198 0.0794 -0.25 0.8026
## UniqueAuthors2 0.1388 0.1117 1.24 0.2148
## UniqueAuthors3 0.1797 0.1172 1.53 0.1261
## UniqueAuthors4 0.2384 0.1146 2.08 0.0381 *
## UniqueAuthors5 0.2333 0.1156 2.02 0.0442 *
## Year1997 0.2382 0.1390 1.71 0.0873 .
## Year1998 0.1448 1.2672 0.11 0.9091
## Year2000 0.2031 0.3447 0.59 0.5560
```

```

## Year2001          -0.1733      0.2739   -0.63    0.5273
## Year2002           0.2058      0.1681     1.22    0.2214
## Year2003           0.0625      0.1701     0.37    0.7133
## Year2004           0.1764      0.1665     1.06    0.2900
## Year2005           0.3577      0.1492     2.40    0.0169 *
## Year2006           0.0810      0.1614     0.50    0.6159
## Year2007           0.1676      0.1374     1.22    0.2230
## Year2008           0.3676      0.1369     2.68    0.0076 **
## Year2009           0.2194      0.1345     1.63    0.1035
## Year2010           0.1941      0.1389     1.40    0.1631
## Year2011           0.1608      0.1310     1.23    0.2203
## Year2012           0.2142      0.1407     1.52    0.1286
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.493
## Multiple R-squared:  0.0558, Adjusted R-squared:  0.00815
## Convergence in 34 IRWLS iterations
##
## Robustness weights:
## 42 weights are ~= 1. The remaining 396 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0366 0.8510 0.9550 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          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 1.233 1          1.111
## LastAuthorFemale  1.170 1          1.082
## Year              1.415 15          1.012

```

Residuals from first and last author



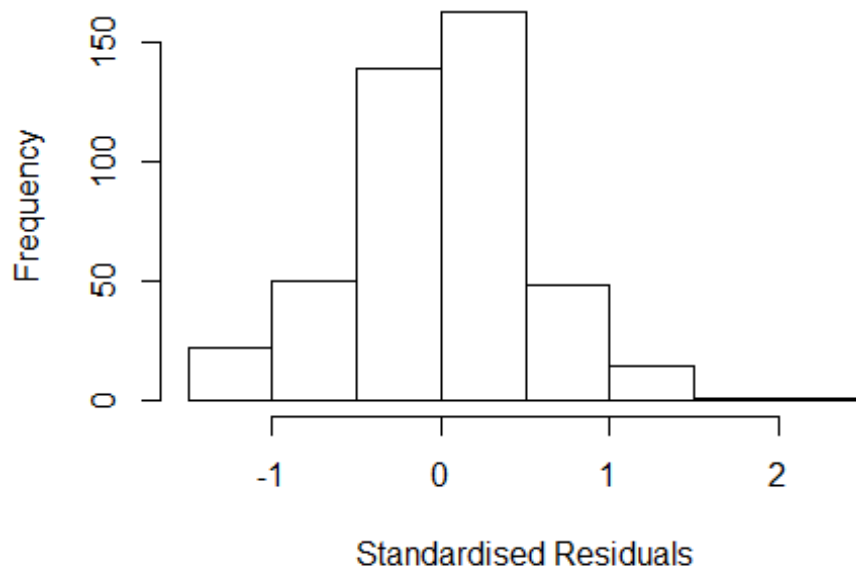
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1535 -0.3137 0.0259 0.3301 2.1263
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9015 0.1080 8.34 1e-15 ***
## FirstAuthorFemale1 0.1267 0.0596 2.12 0.0342 *
## LastAuthorFemale1 -0.0174 0.0794 -0.22 0.8269
## Year1997 0.2735 0.1366 2.00 0.0458 *
## Year1998 0.1967 1.4871 0.13 0.8948
## Year2000 0.2054 0.3442 0.60 0.5510
## Year2001 -0.1279 0.2636 -0.49 0.6277
## Year2002 0.2055 0.1640 1.25 0.2108
## Year2003 0.0340 0.1630 0.21 0.8349
## Year2004 0.1816 0.1619 1.12 0.2626
## Year2005 0.3213 0.1481 2.17 0.0306 *
## Year2006 0.0706 0.1565 0.45 0.6524
```

```

## Year2007          0.1915      0.1317      1.45      0.1468
## Year2008          0.3758      0.1319      2.85      0.0046 **
## Year2009          0.2520      0.1275      1.98      0.0487 *
## Year2010          0.2132      0.1333      1.60      0.1105
## Year2011          0.1883      0.1251      1.51      0.1330
## Year2012          0.2492      0.1345      1.85      0.0647 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.492
## Multiple R-squared:  0.0396, Adjusted R-squared:  0.000726
## Convergence in 40 IRWLS iterations
##
## Robustness weights:
## 36 weights are ~= 1. The remaining 402 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0218 0.8520 0.9500 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      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 1.213 1      1.101
## Year              1.213 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
## -1.1524 -0.3189 0.0262 0.3320 2.1299
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9021 0.1079 8.36 9.3e-16 ***
## FirstAuthorFemale1 0.1249 0.0591 2.11 0.0352 *
## Year1997 0.2744 0.1364 2.01 0.0449 *
## Year1998 0.1973 1.5155 0.13 0.8965
## Year2000 0.2009 0.3449 0.58 0.5604
## Year2001 -0.1281 0.2637 -0.49 0.6275
## Year2002 0.2053 0.1639 1.25 0.2111
## Year2003 0.0304 0.1619 0.19 0.8513
## Year2004 0.1770 0.1585 1.12 0.2646
## Year2005 0.3179 0.1463 2.17 0.0303 *
## Year2006 0.0687 0.1562 0.44 0.6602
## Year2007 0.1905 0.1317 1.45 0.1489
```

```

## Year2008          0.3738      0.1318      2.84      0.0048 **
## Year2009          0.2504      0.1273      1.97      0.0499 *
## Year2010          0.2106      0.1317      1.60      0.1107
## Year2011          0.1854      0.1249      1.48      0.1383
## Year2012          0.2451      0.1336      1.83      0.0672 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.49
## Multiple R-squared:  0.0396, Adjusted R-squared:  0.00306
## Convergence in 41 IRWLS iterations
##
## Robustness weights:
## 36 weights are ~= 1. The remaining 402 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0193 0.8520 0.9490 0.8930 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      2.28e-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 1.157 1          1.075
## Year          1.157 15          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
## -1.1766 -0.3168  0.0212  0.3210  2.1080
##

```

```

## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.94400    0.10174   9.28  <2e-16 ***
## LastAuthorFemale1 0.00218    0.08135   0.03  0.9787
## Year1997        0.33136    0.13035   2.54  0.0114 *
## Year1998        0.15049    1.39952   0.11  0.9144
## Year2000        0.24390    0.33574   0.73  0.4680
## Year2001       -0.14424    0.24662  -0.58  0.5590
## Year2002        0.18132    0.15733   1.15  0.2498
## Year2003        0.01581    0.16294   0.10  0.9228
## Year2004        0.16686    0.16117   1.04  0.3011
## Year2005        0.30169    0.14603   2.07  0.0394 *
## Year2006        0.06306    0.15549   0.41  0.6853
## Year2007        0.16647    0.12987   1.28  0.2006
## Year2008        0.35815    0.12754   2.81  0.0052 **
## Year2009        0.23258    0.12411   1.87  0.0616 .
## Year2010        0.18595    0.12785   1.45  0.1466
## Year2011        0.16780    0.12216   1.37  0.1703
## Year2012        0.22504    0.13146   1.71  0.0876 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.498
## Multiple R-squared:  0.0311, Adjusted R-squared:  -0.00571
## Convergence in 37 IRWLS iterations
##
## Robustness weights:
## 41 weights are ~= 1. The remaining 397 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0334 0.8530 0.9490 0.8940 0.9860 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      2.28e-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] "#####"
```

```

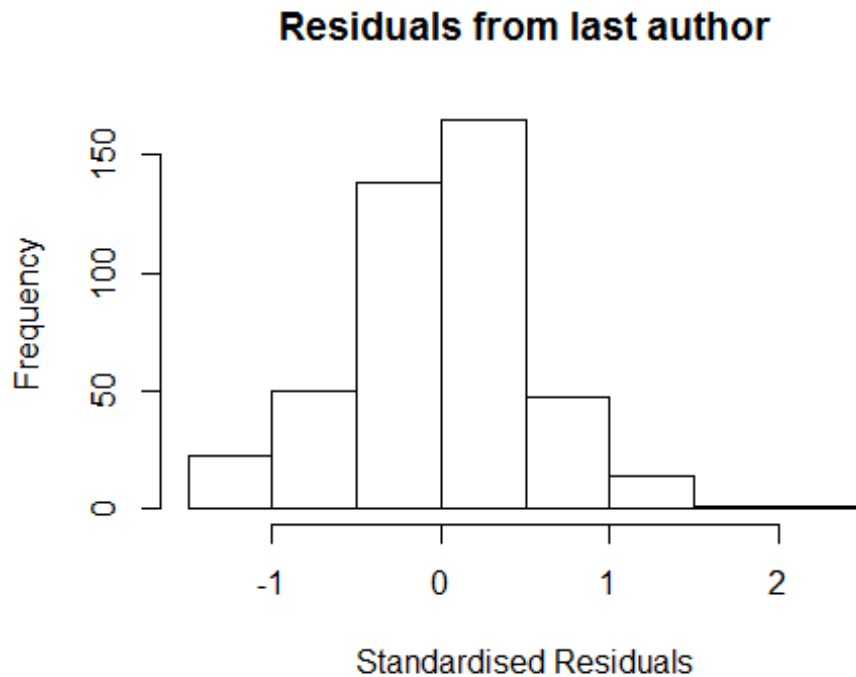
## [1] "Analysis of AJSC 3105"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    6   11   18   11    4   10   11   12   18   21   24   17   25   20
## 2011 2012
##   15   20
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    5    5    2    3    3    4    7    7    9   12   14    3   13   10
## 2011 2012
##   11   16
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    5    4    1    1    2    4    6    6    8   11   12    1   11    8
## 2011 2012
##   11   13
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.28750365903445"
## [1] "Male first author team size 2018 geometric mean: 3.16978638492223"

## Warning in wilcox.test.default(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"
## [1] "Male last author team size 2018 geometric mean: 3.40447488031232"

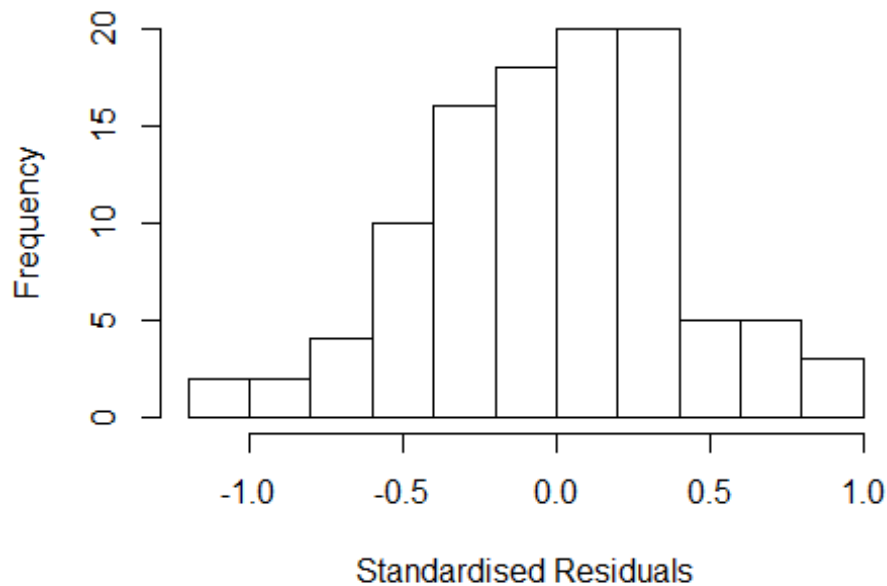
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```

```
##
## 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"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.213e+14 1      1.101e+07
## LastAuthorFemale  1.799e+14 1      1.341e+07
## UniqueAuthors    4.025e+16 4      1.190e+02
## Year              4.768e+29 16      8.462e+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
## -1.12156 -0.28529 0.00579 0.23927 0.91553
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 7.92e-01 7.21e-08 1.10e+07 < 2e-16 ***
## FirstAuthorFemale1 -1.93e-02 1.18e-01 -1.60e-01 0.871
## LastAuthorFemale1 2.06e-01 1.45e-01 1.42e+00 0.160
## UniqueAuthors2 4.26e-02 1.53e-01 2.80e-01 0.782
## UniqueAuthors3 2.97e-01 2.34e-01 1.27e+00 0.209
## UniqueAuthors4 1.08e-01 1.71e-01 6.30e-01 0.530
## UniqueAuthors5 2.11e-01 1.56e-01 1.35e+00 0.180
## Year1997 3.53e-01 1.50e-01 2.36e+00 0.021 *
## Year1998 3.31e-01 4.48e-01 7.40e-01 0.461
## Year1999 -2.33e-01 1.93e-01 -1.21e+00 0.231
```

```

## Year2000      -1.47e-01   7.45e-08 -1.97e+06 < 2e-16 ***
## Year2001      -1.41e-02   2.36e-01 -6.00e-02   0.952
## Year2002       2.55e-01   2.27e-01  1.13e+00   0.263
## Year2003       5.38e-01   2.28e-01  2.36e+00   0.020 *
## Year2004       2.37e-01   1.28e-01  1.86e+00   0.067 .
## Year2005       3.40e-01   1.82e-01  1.87e+00   0.065 .
## Year2006       4.16e-01   1.79e-01  2.33e+00   0.023 *
## Year2007       4.37e-01   2.05e-01  2.13e+00   0.036 *
## Year2008       8.97e-01   1.71e-01  5.26e+00  1.1e-06 ***
## Year2009       2.91e-01   2.02e-01  1.44e+00   0.153
## Year2010       3.58e-01   1.51e-01  2.38e+00   0.020 *
## Year2011       3.01e-01   2.29e-01  1.31e+00   0.193
## Year2012       2.14e-01   1.91e-01  1.12e+00   0.265
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.429
## Multiple R-squared:  0.205, Adjusted R-squared:  -0.00773
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 91 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.475  0.887  0.959  0.912  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      9.52e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.997e+15  1      4.469e+07
## LastAuthorFemale  2.818e+15  1      5.308e+07
## Year              1.286e+16 16      3.187e+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)

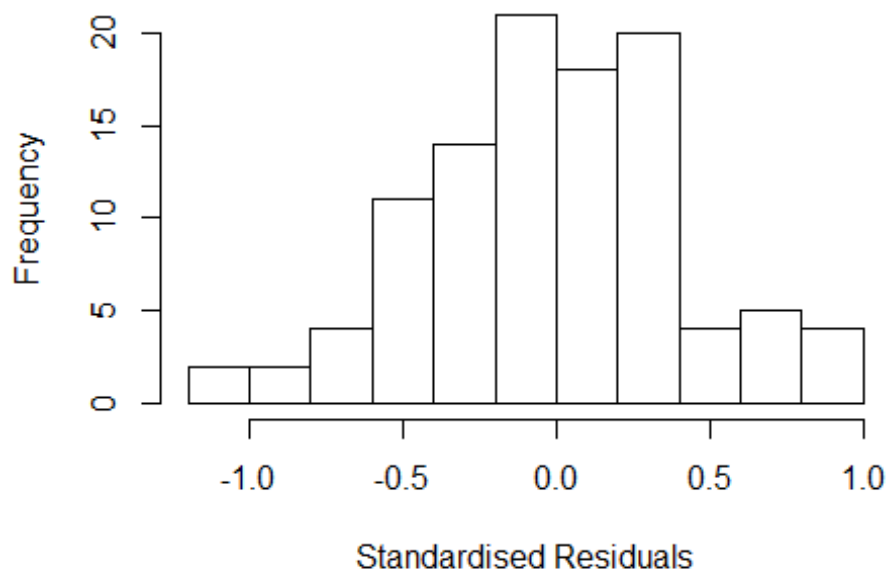
```

```
##
## 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
## -1.05e+00 -2.55e-01 -8.33e-17  2.40e-01  9.46e-01
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    7.92e-01   5.73e-08  1.38e+07 < 2e-16 ***
## FirstAuthorFemale1 2.34e-02   1.14e-01  2.10e-01  0.83760
## LastAuthorFemale1 2.66e-01   1.35e-01  1.96e+00  0.05269 .
## Year1997        4.61e-01   1.14e-01  4.03e+00  0.00012 ***
## Year1998        3.43e-01   3.41e-01  1.00e+00  0.31794
## Year1999       -3.35e-01   1.70e-01 -1.97e+00  0.05251 .
## Year2000       -1.47e-01   5.38e-08 -2.73e+06 < 2e-16 ***
## Year2001       -1.42e-02   2.27e-01 -6.00e-02  0.95014
## Year2002        3.75e-01   2.28e-01  1.64e+00  0.10403
## Year2003        7.07e-01   1.74e-01  4.06e+00  0.00011 ***
## Year2004        2.95e-01   7.92e-02  3.72e+00  0.00035 ***
## Year2005        4.63e-01   1.84e-01  2.52e+00  0.01361 *
## Year2006        5.14e-01   1.51e-01  3.40e+00  0.00103 **
## Year2007        5.54e-01   1.68e-01  3.30e+00  0.00142 **
## Year2008        1.00e+00   5.90e-08  1.70e+07 < 2e-16 ***
## Year2009        4.05e-01   1.43e-01  2.83e+00  0.00585 **
## Year2010        4.85e-01   8.69e-02  5.59e+00  2.7e-07 ***
## Year2011        4.72e-01   1.88e-01  2.50e+00  0.01419 *
## Year2012        3.77e-01   1.38e-01  2.73e+00  0.00774 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.455
## Multiple R-squared:  0.162, Adjusted R-squared:  -0.0129
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 95 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.576  0.901  0.964  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      9.52e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
```

```
##           500           50           2           1           1000           200
##   trace.lev      mts   compute.rd
##           0           1000           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 5.665e+13 1 7.527e+06
## Year 5.665e+13 16 2.690e+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:
```

```

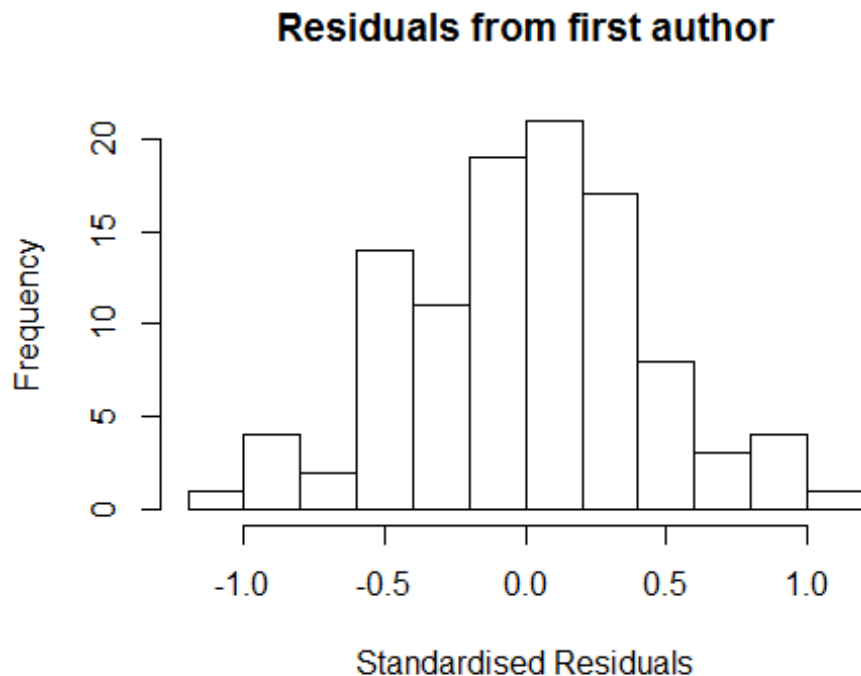
##      Min      1Q      Median      3Q      Max
## -1.07306 -0.29389  0.00116  0.26474  1.01194
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    7.92e-01   0.00e+00      Inf < 2e-16 ***
## FirstAuthorFemale1 3.68e-02   1.17e-01   3.10e-01  0.75497
## Year1997        4.51e-01   1.16e-01   3.90e+00  0.00019 ***
## Year1998        3.40e-01   3.30e-01   1.03e+00  0.30592
## Year1999       -8.28e-02   1.17e-01  -7.00e-01  0.48294
## Year2000       -1.47e-01   2.33e-08  -6.30e+06 < 2e-16 ***
## Year2001       -2.09e-02   2.30e-01  -9.00e-02  0.92791
## Year2002        3.76e-01   2.23e-01   1.69e+00  0.09538 .
## Year2003        7.57e-01   1.59e-01   4.76e+00  7.7e-06 ***
## Year2004        3.82e-01   6.05e-02   6.31e+00  1.1e-08 ***
## Year2005        4.59e-01   1.81e-01   2.53e+00  0.01316 *
## Year2006        5.68e-01   1.84e-01   3.08e+00  0.00276 **
## Year2007        5.75e-01   1.63e-01   3.52e+00  0.00069 ***
## Year2008        1.00e+00   2.02e-08   4.98e+07 < 2e-16 ***
## Year2009        4.86e-01   1.59e-01   3.06e+00  0.00294 **
## Year2010        5.10e-01   9.69e-02   5.26e+00  1.0e-06 ***
## Year2011        5.34e-01   1.68e-01   3.17e+00  0.00209 **
## Year2012        4.01e-01   1.39e-01   2.89e+00  0.00483 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.49
## Multiple R-squared:  0.123, Adjusted R-squared:  -0.048
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~ = 1. The remaining 90 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.611  0.896  0.959   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      9.52e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500            50          2            1            1000      200
##      trace.lev      mts      compute.rd
##      0              1000          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
## -1.05e+00 -2.66e-01  8.33e-16  2.58e-01  9.48e-01
##
## Coefficients:
```

```

##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.7920     0.0000    Inf < 2e-16 ***
## LastAuthorFemale1 0.2677     0.1348     1.99 0.05016 .
## Year1997          0.4801     0.0717     6.69 2.1e-09 ***
## Year1998          0.3431     0.3428     1.00 0.31966
## Year1999         -0.3137     0.1348    -2.33 0.02226 *
## Year2000         -0.1470     0.0000   -Inf < 2e-16 ***
## Year2001         -0.0025     0.2093    -0.01 0.99050
## Year2002          0.3749     0.2289     1.64 0.10514
## Year2003          0.7099     0.1727     4.11 8.9e-05 ***
## Year2004          0.2980     0.0769     3.88 0.00020 ***
## Year2005          0.4653     0.1872     2.49 0.01482 *
## Year2006          0.5218     0.1402     3.72 0.00035 ***
## Year2007          0.5675     0.1388     4.09 9.6e-05 ***
## Year2008          1.0050     0.0000    Inf < 2e-16 ***
## Year2009          0.4088     0.1436     2.85 0.00551 **
## Year2010          0.4965     0.0719     6.91 7.7e-10 ***
## Year2011          0.4702     0.1892     2.48 0.01488 *
## Year2012          0.3817     0.1349     2.83 0.00577 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.451
## Multiple R-squared:  0.163, Adjusted R-squared:  -0.000659
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 95 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.565  0.904   0.966   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      9.52e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 105"
## [1] ""
## [1] ""
## [1] "#####"

```



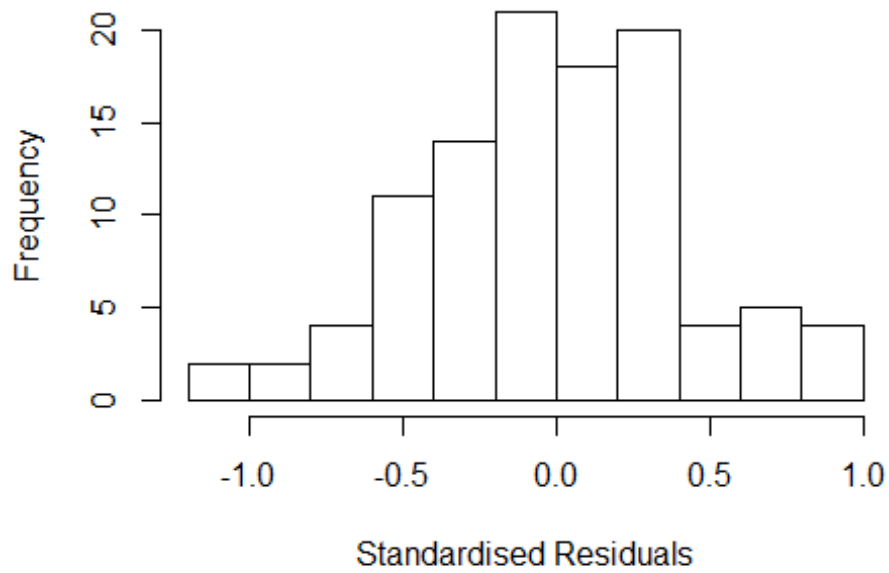
```

## [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
##    7   10   19   22   21   18   25   10   20   20   12   13   17   14   11
## 2011 2012
##    10    9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    4    7    7   11    8   13    9   15    8   10    7    9   12    9
## 2011 2012
##    7    7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    4    7    6   11    7   12    6   12    5    6    4    8   12    9
## 2011 2012
##    7    7
## [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: 1.73205080756888"
## [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

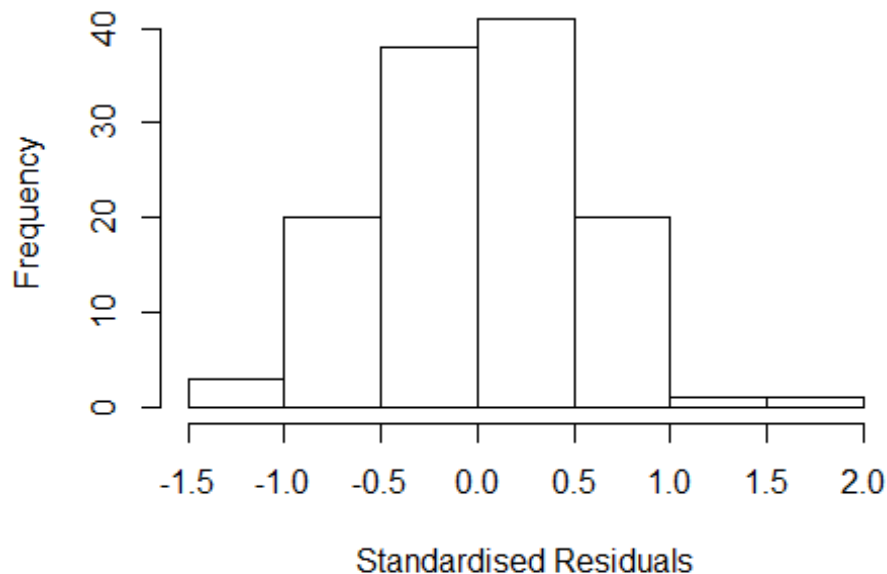
```

Residuals from last author



| ## | | GVIF | Df | GVIF^(1/(2*Df)) |
|----|-------------------|--------|----|-----------------|
| ## | FirstAuthorFemale | 4.088 | 1 | 2.022 |
| ## | LastAuthorFemale | 4.645 | 1 | 2.155 |
| ## | UniqueAuthors | 63.266 | 4 | 1.679 |
| ## | Year | 28.952 | 16 | 1.111 |

Residuals from first and last author and team size



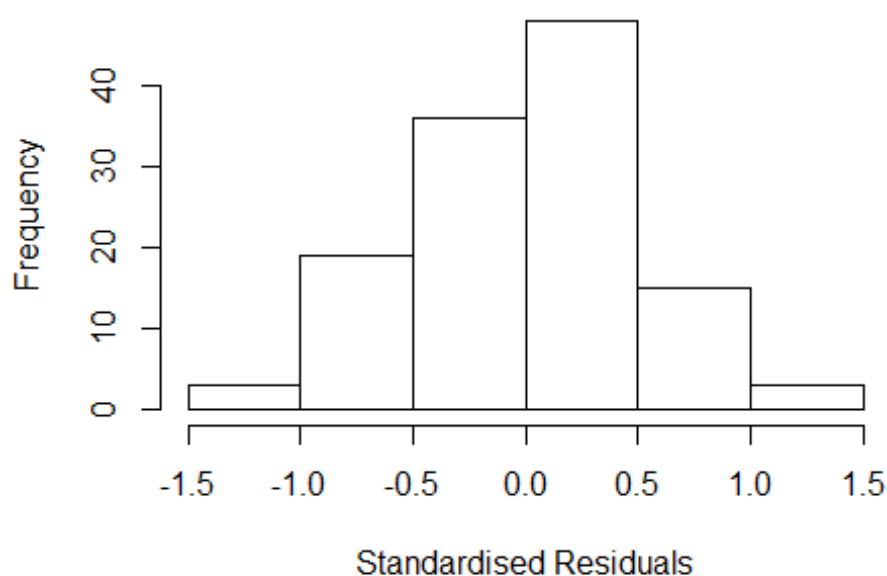
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.04656 -0.28605 0.00633 0.34241 1.53013
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9510 0.0000 Inf <2e-16 ***
## FirstAuthorFemale1 -0.5259 0.3559 -1.48 0.143
## LastAuthorFemale1 0.7633 0.4648 1.64 0.104
## UniqueAuthors2 0.2731 0.1235 2.21 0.029 *
## UniqueAuthors3 0.5789 0.2244 2.58 0.011 *
## UniqueAuthors4 -0.5324 0.5733 -0.93 0.355
## UniqueAuthors5 0.2890 0.3969 0.73 0.468
## Year1997 0.4308 0.1892 2.28 0.025 *
## Year1998 -0.0711 0.2577 -0.28 0.783
## Year1999 0.2684 0.3941 0.68 0.497
```

```

## Year2000          -0.2585      0.1626    -1.59      0.115
## Year2001          -0.0871      0.4277    -0.20      0.839
## Year2002           0.0682      0.1335      0.51      0.611
## Year2003           0.2031      0.1254      1.62      0.108
## Year2004          -0.1775      0.2446    -0.73      0.470
## Year2005          -0.1204      0.2513    -0.48      0.633
## Year2006          -0.0819      0.1567    -0.52      0.602
## Year2007          -0.5103      0.2376    -2.15      0.034 *
## Year2008          -0.0761      0.1255    -0.61      0.546
## Year2009          -0.1286      0.1281    -1.00      0.318
## Year2010          -0.0296      0.1629    -0.18      0.856
## Year2011           0.2230      0.3190      0.70      0.486
## Year2012          -0.1180      0.1192    -0.99      0.324
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.504
## Multiple R-squared:  0.186, Adjusted R-squared:  0.00872
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 109 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.337  0.876  0.949  0.909  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      8.06e-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 1.823 1      1.350
## LastAuthorFemale  1.574 1      1.255
## Year              1.939 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
## -1.1337 -0.4102 0.0754 0.3626 1.3214
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 9.51e-01 8.95e-08 1.06e+07 <2e-16 ***
## FirstAuthorFemale1 -2.04e-01 2.43e-01 -8.40e-01 0.4045
## LastAuthorFemale1 2.75e-01 2.46e-01 1.12e+00 0.2657
## Year1997 5.89e-01 1.93e-01 3.06e+00 0.0028 **
## Year1998 -2.32e-02 2.50e-01 -9.00e-02 0.9264
## Year1999 3.32e-01 3.46e-01 9.60e-01 0.3395
## Year2000 -1.88e-01 1.53e-01 -1.23e+00 0.2199
## Year2001 1.22e-01 5.00e-01 2.40e-01 0.8085
## Year2002 1.36e-01 1.39e-01 9.80e-01 0.3285
## Year2003 3.71e-01 1.13e-01 3.28e+00 0.0014 **
## Year2004 -1.56e-01 2.12e-01 -7.30e-01 0.4654
## Year2005 1.32e-01 1.95e-01 6.80e-01 0.4988
```

```

## Year2006          8.17e-03   1.94e-01   4.00e-02   0.9665
## Year2007         -3.59e-01   2.17e-01  -1.66e+00   0.1006
## Year2008          8.85e-02   1.42e-01   6.30e-01   0.5332
## Year2009          2.79e-02   1.32e-01   2.10e-01   0.8330
## Year2010          4.26e-02   1.23e-01   3.50e-01   0.7294
## Year2011          1.83e-01   3.77e-01   4.90e-01   0.6285
## Year2012          7.11e-02   1.00e-01   7.10e-01   0.4790
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.543
## Multiple R-squared:  0.119, Adjusted R-squared:  -0.0323
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~ = 1. The remaining 118 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.533  0.902  0.951  0.922  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      8.06e-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 1.424 1      1.194
## Year              1.424 16      1.011
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max

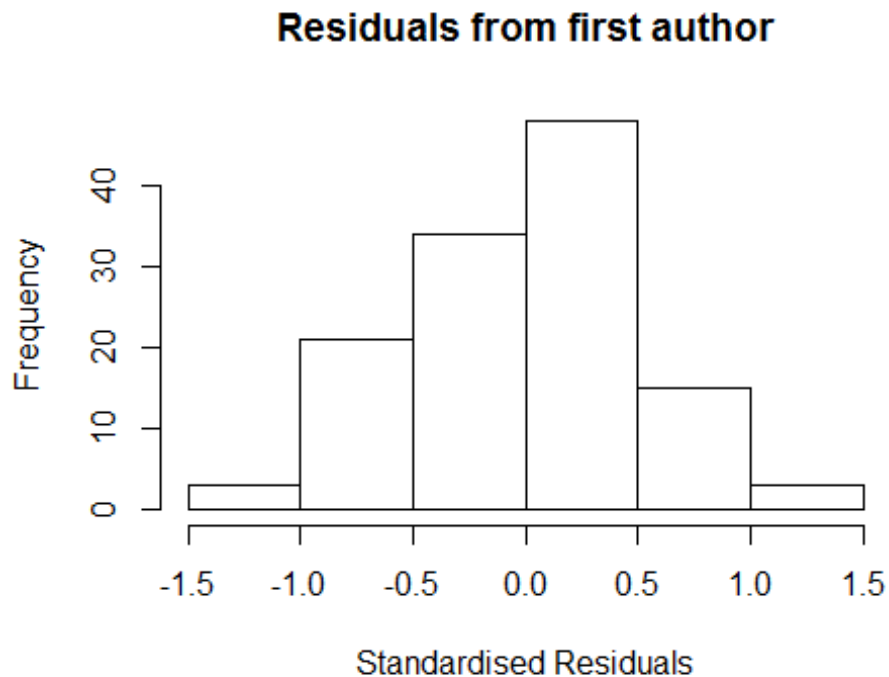
```

```

## -1.1345 -0.4315 0.0791 0.3497 1.3222
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    9.51e-01   6.87e-08  1.38e+07  <2e-16 ***
## FirstAuthorFemale1 -6.50e-02   2.14e-01  -3.00e-01   0.7620
## Year1997        5.89e-01   1.93e-01   3.05e+00   0.0028 **
## Year1998       -2.37e-02   2.51e-01  -9.00e-02   0.9247
## Year1999        3.01e-01   3.56e-01   8.50e-01   0.3989
## Year2000       -1.88e-01   1.53e-01  -1.23e+00   0.2212
## Year2001        1.21e-01   5.07e-01   2.40e-01   0.8121
## Year2002        1.37e-01   1.39e-01   9.90e-01   0.3266
## Year2003        3.88e-01   1.29e-01   3.00e+00   0.0034 **
## Year2004       -1.11e-01   2.16e-01  -5.10e-01   0.6082
## Year2005        1.33e-01   1.95e-01   6.80e-01   0.4976
## Year2006        8.17e-03   1.95e-01   4.00e-02   0.9666
## Year2007       -3.59e-01   2.17e-01  -1.65e+00   0.1012
## Year2008        8.86e-02   1.42e-01   6.20e-01   0.5335
## Year2009        4.08e-02   1.29e-01   3.20e-01   0.7534
## Year2010        5.95e-02   1.21e-01   4.90e-01   0.6250
## Year2011        1.83e-01   3.79e-01   4.80e-01   0.6297
## Year2012        5.70e-02   9.05e-02   6.30e-01   0.5305
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.538
## Multiple R-squared: 0.112, Adjusted R-squared: -0.0306
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 117 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.526  0.893  0.947   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          8.06e-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"

```

```
## 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 1.222 1          1.105
## Year            1.222 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
## -1.133 -0.427  0.051  0.368  1.320
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.95100    0.00000      Inf <2e-16 ***
## LastAuthorFemale1 0.15077    0.24089     0.63  0.5327
## Year1997        0.58910    0.19213     3.07  0.0028 **
## Year1998       -0.02232    0.24957    -0.09  0.9289
## Year1999        0.28709    0.34876     0.82  0.4122
```



```

## Year2000      -0.18876      0.15226      -1.24      0.2178
## Year2001      0.12289      0.49078      0.25      0.8028
## Year2002      0.13558      0.13900      0.98      0.3316
## Year2003      0.36036      0.10692      3.37      0.0010 **
## Year2004     -0.15892      0.21428     -0.74      0.4599
## Year2005      0.13166      0.19481      0.68      0.5006
## Year2006      0.00819      0.19373      0.04      0.9664
## Year2007     -0.35938      0.21622     -1.66      0.0995 .
## Year2008      0.08842      0.14128      0.63      0.5328
## Year2009      0.02072      0.13229      0.16      0.8758
## Year2010      0.03260      0.12094      0.27      0.7880
## Year2011      0.18162      0.37232      0.49      0.6267
## Year2012      0.04991      0.08856      0.56      0.5743
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.551
## Multiple R-squared:  0.114, Adjusted R-squared:  -0.0283
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 116 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.545  0.903   0.949   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      8.06e-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: 124"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 3107"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010

```

```

##      8      8    18      3    22    43    37    30      1    47      6    70    63    71    79
## 2011 2012
##    93    63
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      4      2      5      0      7      9    19    16      0    20      1    35    33    33    44
## 2011 2012
##    55    43
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      1      5      0      5      8    17    14      0    17      1    24    30    30    39
## 2011 2012
##    54    40
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.5298688016454"
## [1] "Male first author team size 2018 geometric mean: 3.5968310806903"

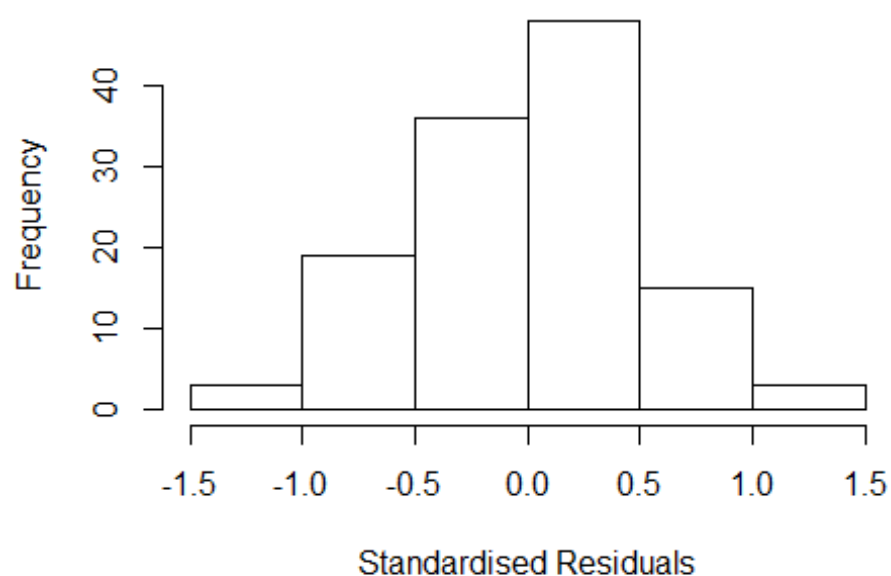
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## 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: 4.81746241999498"
## [1] "Male last author team size 2018 geometric mean: 3.55797182780347"

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

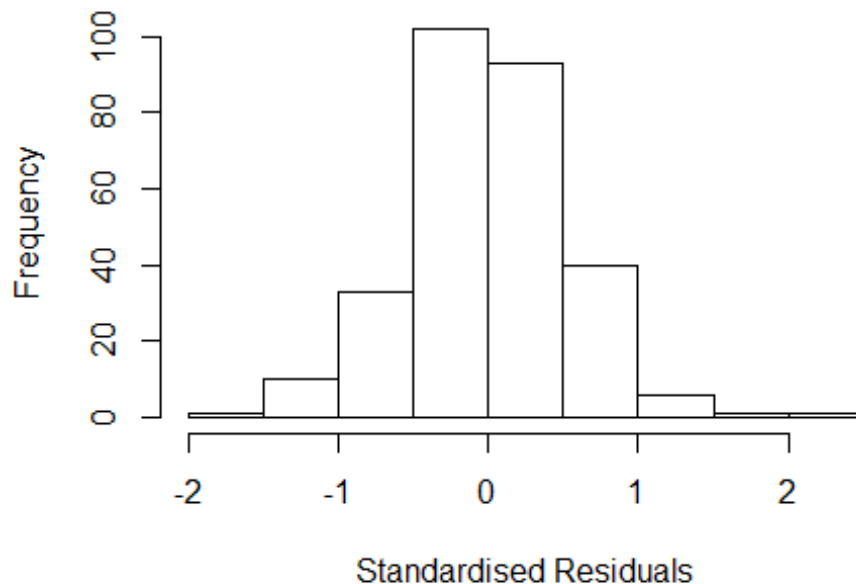
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 130, p-value = 0.3
## 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 1.742e+00 1          1.320
## LastAuthorFemale  1.460e+00 1          1.208
## UniqueAuthors    4.180e+13 4          50.425
## Year              5.200e+13 14          3.089
```

Residuals from first and last author and team size



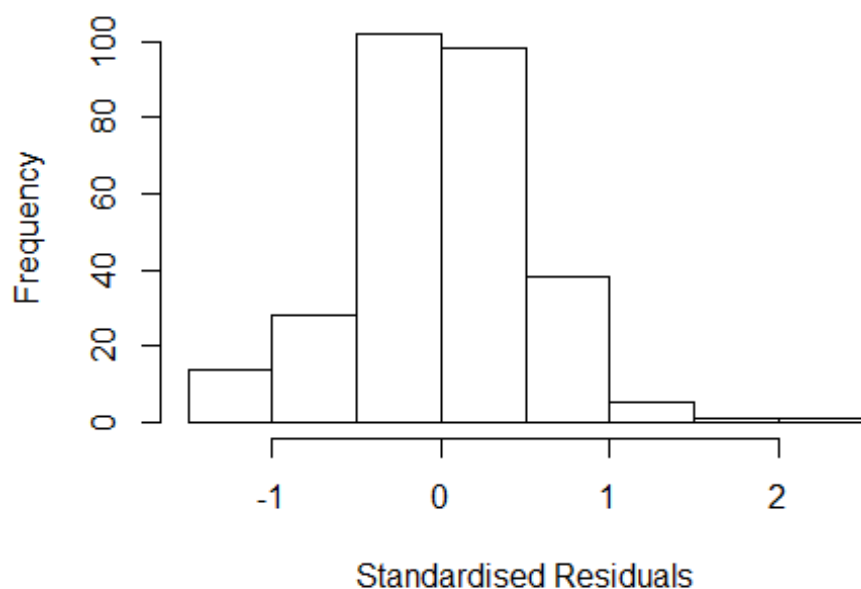
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.51714 -0.30828 -0.00705 0.32030 2.22131
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8490 0.0727 11.69 < 2e-16 ***
## FirstAuthorFemale1 -0.0706 0.0952 -0.74 0.45914
## LastAuthorFemale1 -0.1379 0.1439 -0.96 0.33885
## UniqueAuthors2 0.4273 0.1739 2.46 0.01462 *
## UniqueAuthors3 0.6165 0.1760 3.50 0.00054 ***
## UniqueAuthors4 0.4866 0.1848 2.63 0.00895 **
## UniqueAuthors5 0.4554 0.1700 2.68 0.00787 **
## Year1997 0.4250 0.0727 5.85 1.4e-08 ***
## Year1998 -0.1309 0.6946 -0.19 0.85066
## Year2000 0.3507 0.3547 0.99 0.32370
```

```

## Year2001          -0.1082      0.2038   -0.53   0.59601
## Year2002           0.0516      0.1965    0.26   0.79300
## Year2003           0.1177      0.2175    0.54   0.58876
## Year2005           0.1634      0.2053    0.80   0.42688
## Year2006           0.0116      0.1849    0.06   0.94991
## Year2007          -0.1494      0.1999   -0.75   0.45564
## Year2008          -0.1068      0.1944   -0.55   0.58320
## Year2009          -0.3216      0.1995   -1.61   0.10818
## Year2010          -0.0962      0.1885   -0.51   0.61013
## Year2011          -0.0398      0.1840   -0.22   0.82890
## Year2012          -0.2799      0.2062   -1.36   0.17576
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.51
## Multiple R-squared:  0.125, Adjusted R-squared:  0.0595
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 261 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0188 0.8640 0.9590 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          3.48e-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 1.450 1 1.204
## LastAuthorFemale 1.209 1 1.099
## Year 1.729 14 1.020

```

Residuals from first and last author



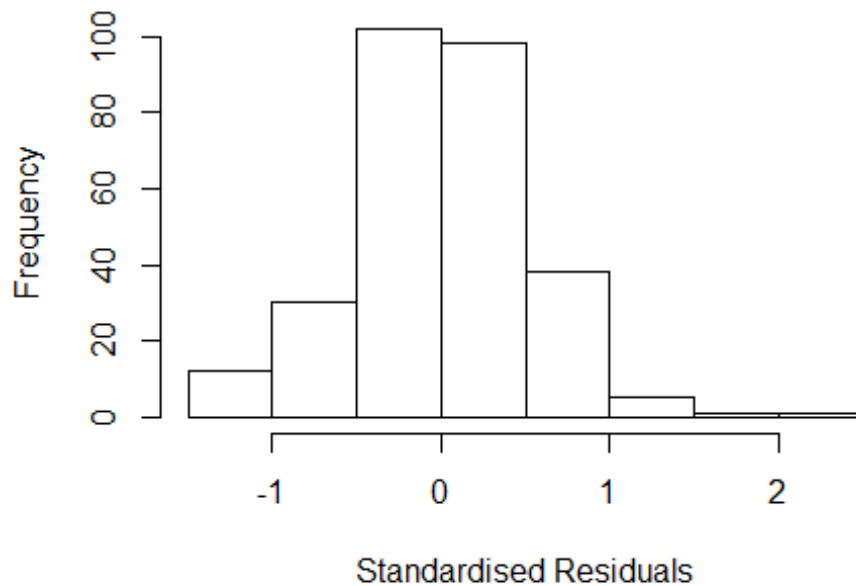
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.23e+00 -3.30e-01 -5.55e-16 3.78e-01 2.24e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8490 0.0726 11.69 < 2e-16 ***
## FirstAuthorFemale1 -0.0347 0.0895 -0.39 0.69869
## LastAuthorFemale1 -0.0673 0.1435 -0.47 0.63969
## Year1997 0.4250 0.0726 5.85 1.4e-08 ***
## Year1998 0.3359 1.1626 0.29 0.77290
## Year2000 0.7749 0.3118 2.48 0.01357 *
## Year2001 0.3498 0.1301 2.69 0.00763 **
## Year2002 0.3665 0.1688 2.17 0.03085 *
## Year2003 0.4968 0.1760 2.82 0.00511 **
## Year2005 0.6099 0.1374 4.44 1.3e-05 ***
## Year2006 0.4670 0.0726 6.43 5.8e-10 ***
## Year2007 0.2902 0.1365 2.13 0.03440 *
```

```

## Year2008          0.4025      0.1073      3.75  0.00022 ***
## Year2009          0.1529      0.1264      1.21  0.22756
## Year2010          0.3799      0.1198      3.17  0.00170 **
## Year2011          0.4263      0.1050      4.06  6.4e-05 ***
## Year2012          0.1925      0.1183      1.63  0.10481
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.522
## Multiple R-squared:  0.0641, Adjusted R-squared:  0.0086
## Convergence in 36 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 266 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0273 0.8670 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      3.48e-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 1.447 1          1.203
## Year              1.447 14          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
## -1.23e+00 -3.24e-01 -2.22e-16 3.72e-01 2.24e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8490 0.0726 11.69 < 2e-16 ***
## FirstAuthorFemale1 -0.0339 0.0891 -0.38 0.70403
## Year1997 0.4250 0.0726 5.85 1.4e-08 ***
## Year1998 0.3371 1.1723 0.29 0.77389
## Year2000 0.7750 0.3121 2.48 0.01364 *
## Year2001 0.3496 0.1300 2.69 0.00762 **
## Year2002 0.3569 0.1646 2.17 0.03101 *
## Year2003 0.4866 0.1714 2.84 0.00488 **
## Year2005 0.6092 0.1365 4.46 1.2e-05 ***
## Year2006 0.4670 0.0726 6.43 5.7e-10 ***
## Year2007 0.2868 0.1370 2.09 0.03730 *
## Year2008 0.3995 0.1087 3.67 0.00029 ***
```



```

## Year2009          0.1448      0.1242      1.17  0.24482
## Year2010          0.3774      0.1196      3.16  0.00178 **
## Year2011          0.4207      0.1040      4.05  6.8e-05 ***
## Year2012          0.1886      0.1183      1.60  0.11187
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.521
## Multiple R-squared:  0.0636, Adjusted R-squared:  0.0117
## Convergence in 36 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 266 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0253 0.8650 0.9550 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      3.48e-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 1.21 1          1.100
## Year            1.21 14          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
## -1.22e+00 -3.38e-01  1.05e-15  3.76e-01  2.24e+00
##
## Coefficients:

```

```

##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.8490      0.0726  11.69 < 2e-16 ***
## LastAuthorFemale1 -0.0664      0.1430   -0.46  0.64269
## Year1997          0.4250      0.0726   5.85  1.4e-08 ***
## Year1998          0.3388      1.1843   0.29  0.77506
## Year2000          0.7740      0.3184   2.43  0.01570 *
## Year2001          0.3430      0.1295   2.65  0.00857 **
## Year2002          0.3669      0.1688   2.17  0.03062 *
## Year2003          0.4918      0.1740   2.83  0.00505 **
## Year2005          0.6062      0.1377   4.40  1.5e-05 ***
## Year2006          0.4670      0.0726   6.43  5.8e-10 ***
## Year2007          0.2854      0.1356   2.11  0.03620 *
## Year2008          0.3999      0.1072   3.73  0.00023 ***
## Year2009          0.1444      0.1219   1.18  0.23711
## Year2010          0.3719      0.1182   3.15  0.00184 **
## Year2011          0.4212      0.1027   4.10  5.4e-05 ***
## Year2012          0.1909      0.1182   1.61  0.10757
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.52
## Multiple R-squared:  0.0641, Adjusted R-squared:  0.0123
## Convergence in 37 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 264 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0245 0.8680 0.9530 0.9020 0.9860 0.9990
## Algorithmic parameters:
##           tuning.chi           bb           tuning.psi           refine.tol
##           1.55e+00           5.00e-01           4.69e+00           1.00e-07
##           rel.tol           solve.tol           eps.outlier           eps.x
##           1.00e-07           1.00e-07           3.48e-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: 287"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 3108"
## [1] "#####"

```

```

## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    5    2    5    3   10    7    2    7    4    6    2   10   13   12
## 2011 2012
##    9    7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    2    2    2    3    2    5    1    3    3    1    1    3    8    3
## 2011 2012
##    4    5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    2    1    2    1    1    2    0    1    3    0    1    1    8    3
## 2011 2012
##    4    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"
## [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: 35"
## [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
##    7   10    8    5   11   10    6    7   15    7   10    8   16   16    9
## 2011 2012
##   22   14
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    2    2    3    0    1    2    2    6    4    4    6    7   10    4
## 2011 2012
##   10    5
##

```

```

## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      3      2      2      3      0      1      2      2      6      4      3      4      7      9      3
## 2011 2012
##      9      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: 1.5874010519682"
## [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 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

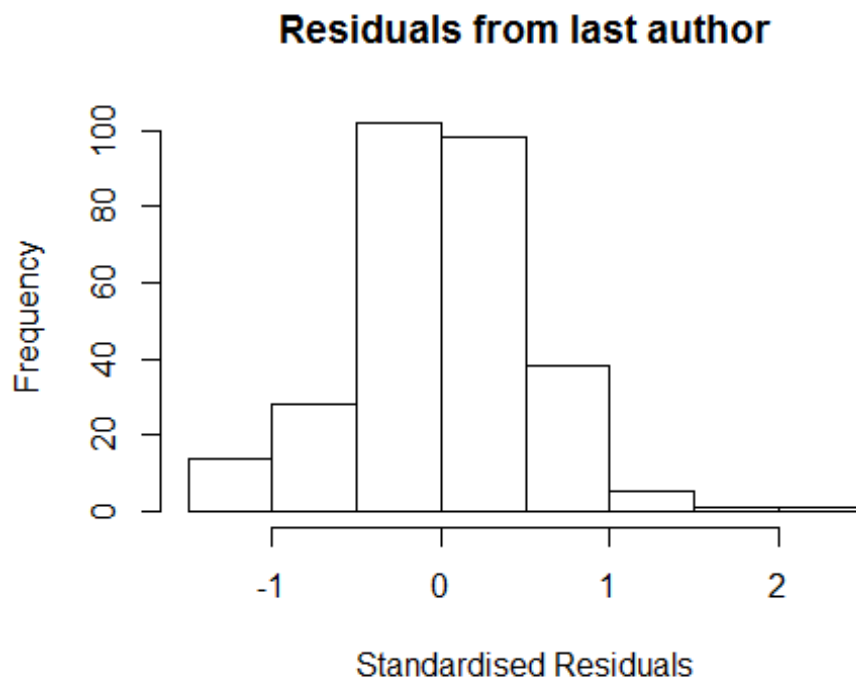
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 64"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 3110"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      7      7     12     10      7      9      7     11     10     21     11     16     21     16     24
## 2011 2012
##     19     13
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      0      1      0      1      0      1      4      2      3      7      5     12      8     10     12
## 2011 2012
##      9      5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      0      1      0      1      0      1      4      2      3      3      3      9      8      8     12
## 2011 2012
##      8      4
## [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: 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 = 6, p-value = 0.1
## 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: 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 = 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 4.453e+13 1 6.673e+06
```

```

## LastAuthorFemale  2.032e+00  1      1.426e+00
## UniqueAuthors    1.367e+15  4      7.798e+01
## Year              9.233e+15 13      4.112e+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
## -5.40e-01 -1.87e-01  1.83e-15  1.59e-01  1.12e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   1.51e+00   1.76e-01   8.61e+00 3.2e-11 ***
## FirstAuthorFemale1 1.23e-01   1.26e-01   9.70e-01  0.3347
## LastAuthorFemale1 1.62e-01   9.77e-02   1.66e+00  0.1035
## UniqueAuthors2   -2.72e-01   1.76e-01  -1.55e+00  0.1279
## UniqueAuthors3   -2.24e-01   1.40e-01  -1.60e+00  0.1171
## UniqueAuthors4   -2.30e-01   1.38e-01  -1.67e+00  0.1023
## UniqueAuthors5   -2.32e-01   1.34e-01  -1.74e+00  0.0890 .
## Year1999          4.27e-01   1.92e-01   2.22e+00  0.0312 *
## Year2001          -9.90e-02   2.57e-08  -3.85e+06 < 2e-16 ***
## Year2002          -7.05e-01   2.16e-01  -3.26e+00  0.0021 **
## Year2003          -4.84e-01   2.03e-01  -2.38e+00  0.0213 *
## Year2004          -5.87e-01   1.91e-01  -3.07e+00  0.0036 **
## Year2005           1.61e-01   1.06e-01   1.52e+00  0.1360
## Year2006          -1.37e-01   1.75e-01  -7.80e-01  0.4381
## Year2007          -1.33e-01   2.51e-01  -5.30e-01  0.5992
## Year2008           5.74e-02   1.52e-01   3.80e-01  0.7076
## Year2009          -3.45e-01   1.38e-01  -2.50e+00  0.0161 *
## Year2010          -1.96e-01   1.38e-01  -1.42e+00  0.1614
## Year2011          -4.06e-01   1.58e-01  -2.57e+00  0.0133 *
## Year2012          -4.71e-01   2.58e-01  -1.82e+00  0.0749 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.336
## Multiple R-squared:  0.446, Adjusted R-squared:  0.221
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~1. The remaining 56 ones are summarized as

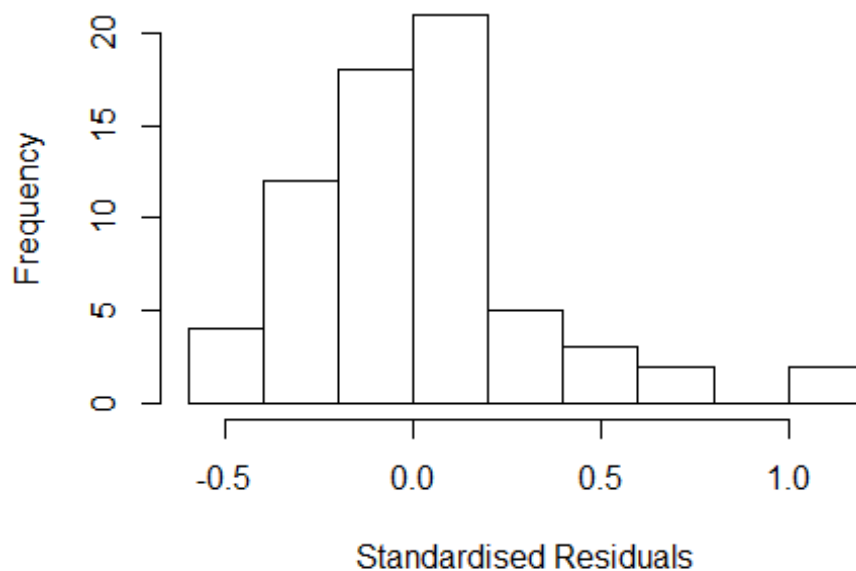
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##    0.245  0.900  0.968  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          1.49e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##      trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      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 cov2cor(v): 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 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
## -0.55610 -0.17337  0.00919  0.14479  1.12049
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2400     0.0000      Inf < 2e-16 ***
## FirstAuthorFemale1  0.1023     0.1245     0.82  0.41500
## LastAuthorFemale1  0.1473     0.0918     1.60  0.11469
## Year1999          0.4957     0.1245     3.98  0.00022 ***
## Year2001         -0.0990     0.0000    -Inf < 2e-16 ***
## Year2002         -0.6600     0.1541    -4.28  8.2e-05 ***
## Year2003         -0.4663     0.2059    -2.26  0.02784 *
## Year2004         -0.5445     0.1704    -3.19  0.00240 **
## Year2005          0.1866     0.0773     2.41  0.01942 *
## Year2006         -0.0868     0.1458    -0.60  0.55447
## Year2007         -0.0439     0.1500    -0.29  0.77092
## Year2008          0.1343     0.1053     1.27  0.20815
## Year2009         -0.3012     0.0850    -3.54  0.00086 ***
## Year2010         -0.1388     0.1109    -1.25  0.21618
## Year2011         -0.3698     0.1292    -2.86  0.00609 **
## Year2012         -0.4369     0.2213    -1.97  0.05380 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.322
## Multiple R-squared:  0.437, Adjusted R-squared:  0.272
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~ = 1. The remaining 57 ones are summarized as
##      Min. 1st Qu. Median Mean 3rd Qu. Max.
##      0.200  0.889  0.957  0.901  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

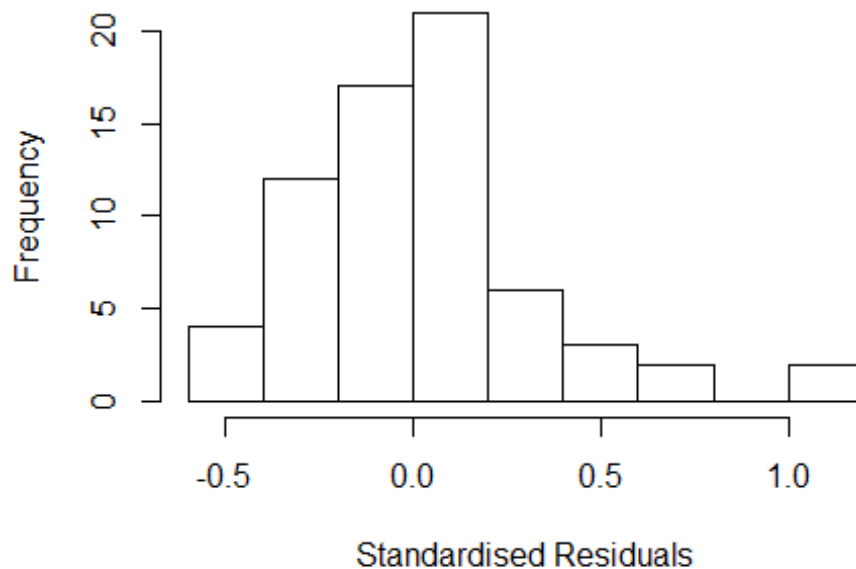
```



```
##          rel.tol          solve.tol          eps.outlier          eps.x
##          1.00e-07          1.00e-07          1.49e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
## nResample    max.it    best.r.s    k.fast.s    k.max maxit.scale
##          500        50        2        1        1000        200
## trace.lev    mts    compute.rd
##          0        1000        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 1.738e+14 1 1.318e+07
## Year 1.738e+14 13 3.529e+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:
```

```

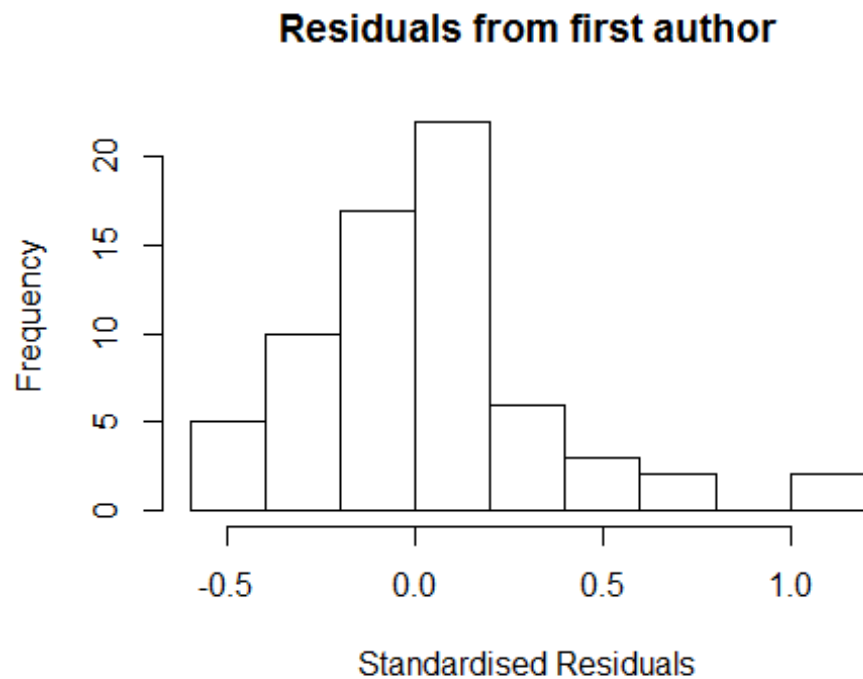
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.5745 -0.1697  0.0186  0.1399  1.0945
##
## Coefficients:
##              Estimate Std. Error   t value Pr(>|t|)
## (Intercept)    1.24e+00   0.00e+00      Inf < 2e-16 ***
## FirstAuthorFemale1 1.11e-01   1.28e-01   8.70e-01   0.3896
## Year1999        4.87e-01   1.28e-01   3.79e+00   0.0004 ***
## Year2001       -9.90e-02   1.27e-08  -7.81e+06 < 2e-16 ***
## Year2002       -6.62e-01   1.52e-01  -4.35e+00   6.4e-05 ***
## Year2003       -3.97e-01   1.39e-01  -2.86e+00   0.0061 **
## Year2004       -5.50e-01   1.73e-01  -3.18e+00   0.0025 **
## Year2005        2.32e-01   5.24e-02   4.42e+00   5.1e-05 ***
## Year2006       -9.59e-02   1.49e-01  -6.40e-01   0.5232
## Year2007       -2.55e-02   1.62e-01  -1.60e-01   0.8752
## Year2008        1.30e-01   1.06e-01   1.22e+00   0.2262
## Year2009       -2.82e-01   8.53e-02  -3.30e+00   0.0017 **
## Year2010       -1.22e-01   1.09e-01  -1.12e+00   0.2687
## Year2011       -3.50e-01   1.38e-01  -2.54e+00   0.0143 *
## Year2012       -4.38e-01   2.21e-01  -1.98e+00   0.0524 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.319
## Multiple R-squared:  0.426, Adjusted R-squared:  0.272
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 58 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.216  0.883  0.971  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      1.49e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats

```

```
##                                "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"

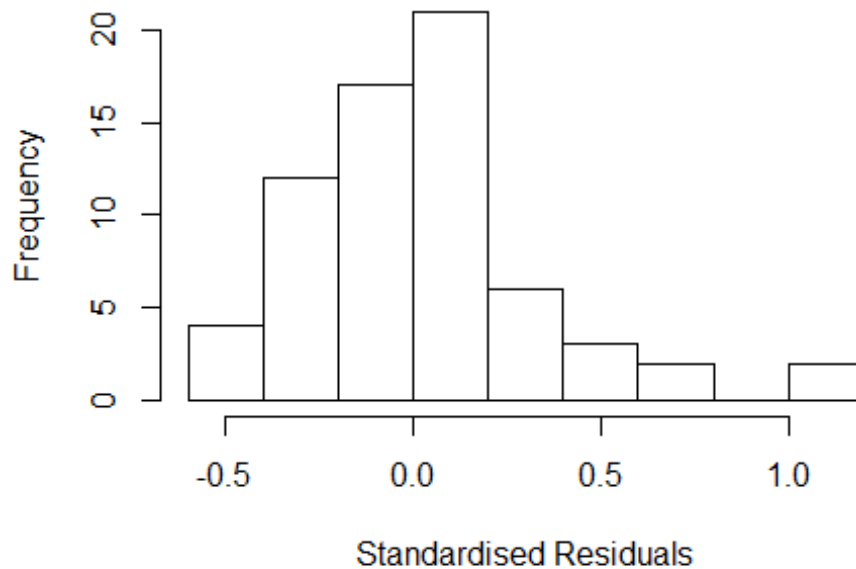
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

## 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



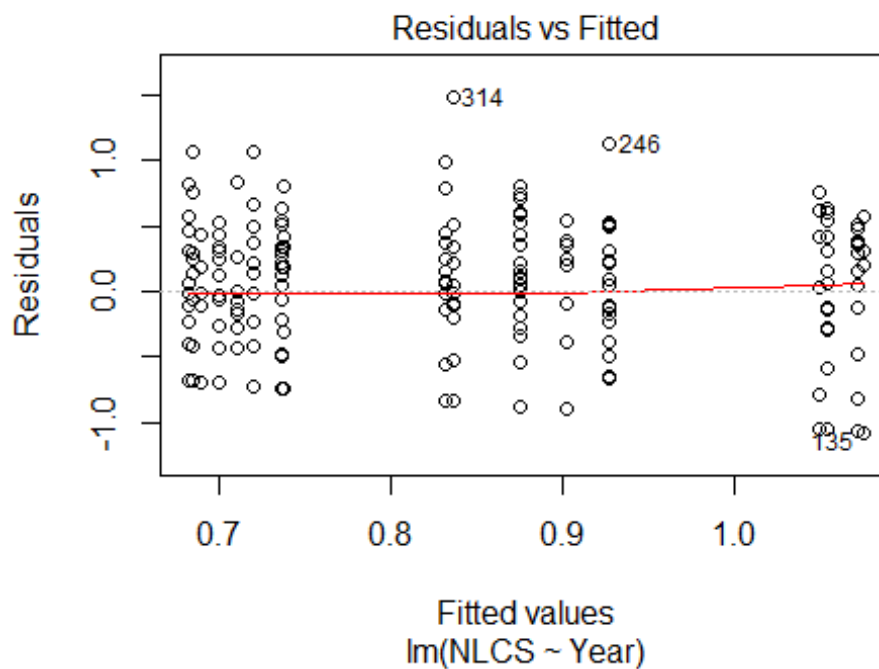
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.563 -0.171 0.000 0.131 1.226
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.24e+00 6.18e-09 2.01e+08 < 2e-16 ***
## LastAuthorFemale1 1.61e-01 9.96e-02 1.62e+00 0.11186
## Year1999 5.98e-01 0.00e+00 Inf < 2e-16 ***
## Year2001 -9.90e-02 0.00e+00 -Inf < 2e-16 ***
## Year2002 -6.52e-01 1.86e-01 -3.51e+00 0.00092 ***
## Year2003 -4.22e-01 2.68e-01 -1.57e+00 0.12174
## Year2004 -4.79e-01 1.43e-01 -3.36e+00 0.00146 **
## Year2005 2.15e-01 5.31e-02 4.06e+00 0.00017 ***
## Year2006 1.57e-02 7.50e-02 2.10e-01 0.83546
## Year2007 -3.70e-02 1.55e-01 -2.40e-01 0.81225
## Year2008 1.75e-01 8.55e-02 2.05e+00 0.04564 *
## Year2009 -2.89e-01 8.60e-02 -3.36e+00 0.00149 **
```

```

## Year2010          -1.42e-01   1.14e-01 -1.24e+00   0.21886
## Year2011          -3.48e-01   1.20e-01 -2.90e+00   0.00540 **
## Year2012          -4.49e-01   2.20e-01 -2.04e+00   0.04664 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.306
## Multiple R-squared:  0.448, Adjusted R-squared:  0.299
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 59 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.072  0.883   0.963   0.894   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      1.49e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 67"
## [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
##   18   9   22   7   13   19   11   5   10   11   19   18   22   13   17
## 2011 2012
##   14   35
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   15   6   17   7   8   16   10   4   6   9   17   14   20   13   14
## 2011 2012
##   13   33
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010

```

```
## 14 6 14 6 7 16 8 4 5 9 14 14 18 13 14
## 2011 2012
## 11 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 = 12, df = 16, p-value = 0.8
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.003, df = 1, p-value = 1

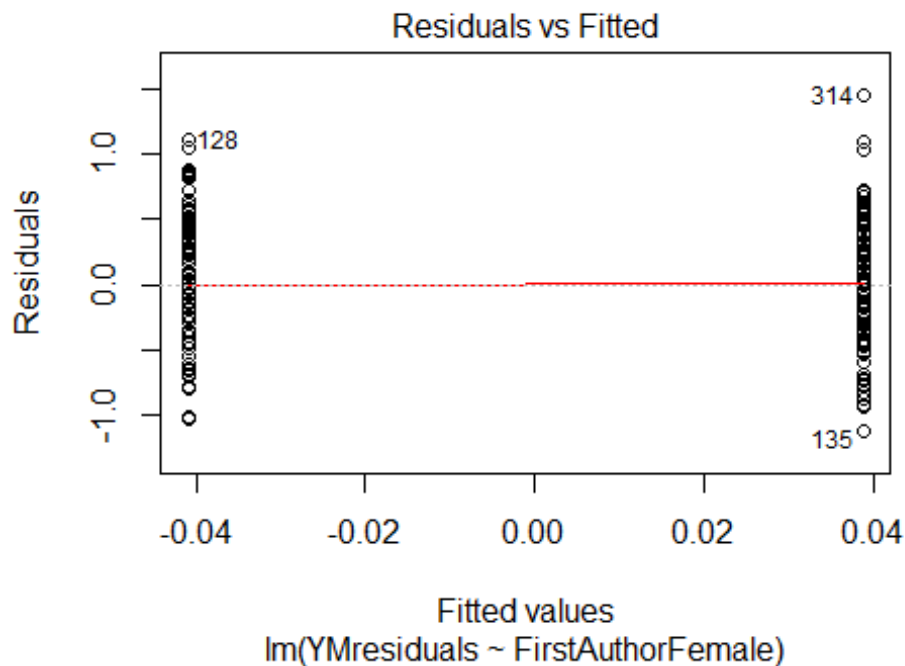
## [1] "Female first author team size 2018 geometric mean: 2.40475854740257"
## [1] "Male first author team size 2018 geometric mean: 2.75945932292243"

## Warning in wilcox.test.default(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: 2.28942848510666"
## [1] "Male last author team size 2018 geometric mean: 2.65310460741205"

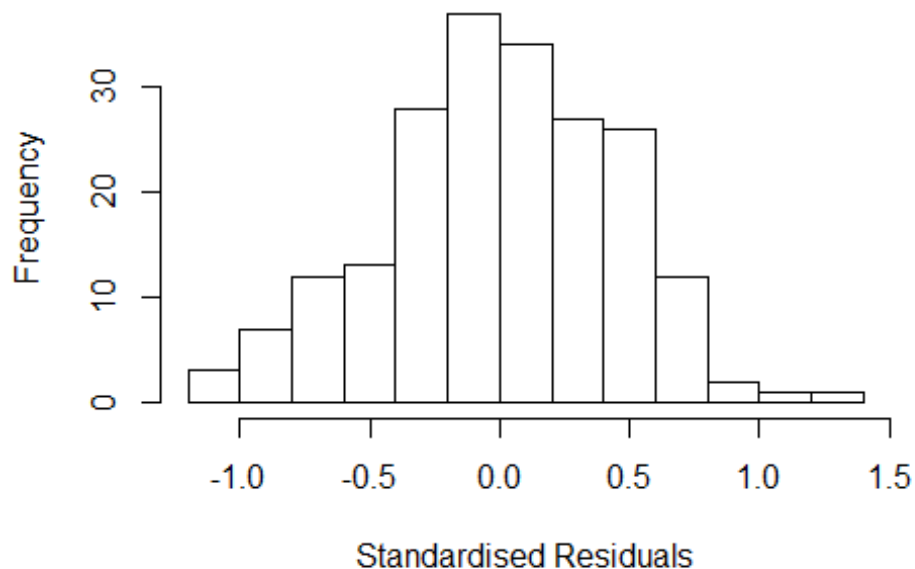
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 28, p-value = 0.9
## 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 | 1.701 | 1 | 1.304 |
| LastAuthorFemale | 1.822 | 1 | 1.350 |
| UniqueAuthors | 4.983 | 4 | 1.222 |
| Year | 9.794 | 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
## -1.13940 -0.31082  0.00771  0.31491  1.29954
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.66326    0.18090   3.67 0.00032 ***
## FirstAuthorFemale1 -0.09411    0.07470  -1.26 0.20938
## LastAuthorFemale1  0.03514    0.07656   0.46 0.64679
## UniqueAuthors2     0.25608    0.09377   2.73 0.00694 **
## UniqueAuthors3     0.05252    0.10009   0.52 0.60042
## UniqueAuthors4     0.40261    0.11622   3.46 0.00066 ***
## UniqueAuthors5     0.24088    0.17707   1.36 0.17540
## Year1997         -0.03340    0.20849  -0.16 0.87292
## Year1998         -0.04238    0.20788  -0.20 0.83869
## Year1999          0.03015    0.29153   0.10 0.91774
```

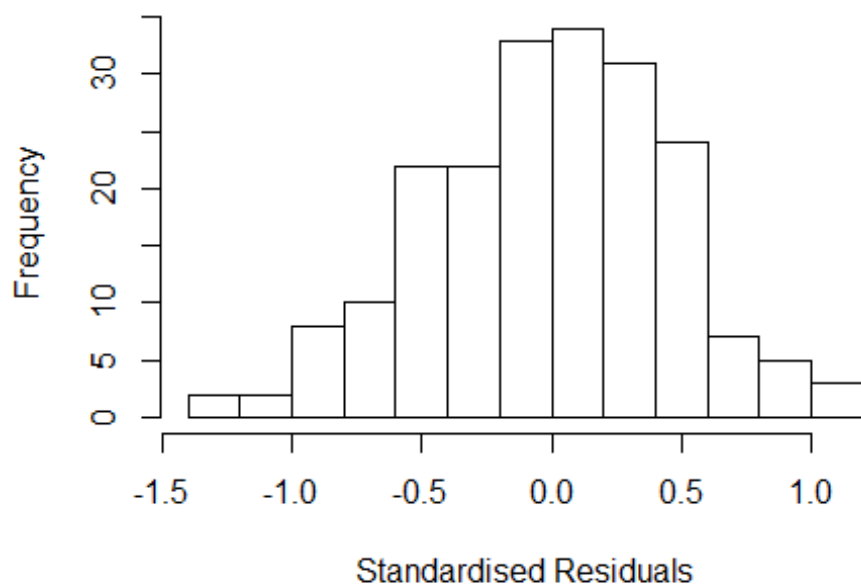


```

## Year2000      -0.03741    0.22502   -0.17  0.86813
## Year2001      -0.03497    0.21014   -0.17  0.86800
## Year2002      -0.18069    0.24170   -0.75  0.45568
## Year2003       0.38140    0.37790    1.01  0.31421
## Year2004       0.53511    0.36375    1.47  0.14302
## Year2005       0.13226    0.25071    0.53  0.59846
## Year2006       0.00147    0.22160    0.01  0.99472
## Year2007       0.07459    0.25512    0.29  0.77033
## Year2008       0.09120    0.21133    0.43  0.66657
## Year2009       0.21049    0.24633    0.85  0.39396
## Year2010       0.21503    0.22870    0.94  0.34835
## Year2011      -0.12980    0.22613   -0.57  0.56669
## Year2012       0.06645    0.20879    0.32  0.75066
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.476
## Multiple R-squared:  0.194, Adjusted R-squared:  0.0959
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 181 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.436  0.881  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      4.93e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.466 1      1.211
## LastAuthorFemale  1.538 1      1.240
## Year              2.061 16      1.023

```

Residuals from first and last author



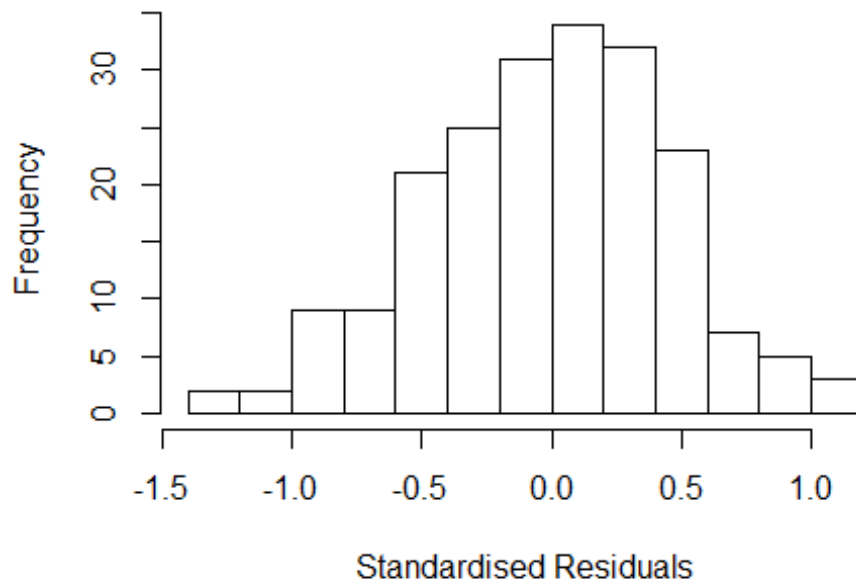
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3024 -0.3285 0.0265 0.3486 1.1035
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7710 0.1769 4.36 2.2e-05 ***
## FirstAuthorFemale1 -0.1152 0.0780 -1.48 0.142
## LastAuthorFemale1 0.0233 0.0772 0.30 0.763
## Year1997 -0.0437 0.2214 -0.20 0.844
## Year1998 -0.0529 0.2155 -0.25 0.806
## Year1999 0.1554 0.2960 0.53 0.600
## Year2000 -0.0260 0.2244 -0.12 0.908
## Year2001 -0.0280 0.2026 -0.14 0.890
## Year2002 -0.1813 0.2622 -0.69 0.490
## Year2003 0.4552 0.3865 1.18 0.240
## Year2004 0.6234 0.3240 1.92 0.056 .
## Year2005 0.1881 0.2462 0.76 0.446
```

```

## Year2006          0.0201      0.2170      0.09      0.926
## Year2007          0.0755      0.2547      0.30      0.767
## Year2008          0.1795      0.2107      0.85      0.395
## Year2009          0.3981      0.2241      1.78      0.077 .
## Year2010          0.3470      0.2168      1.60      0.111
## Year2011         -0.0658      0.2121     -0.31      0.757
## Year2012          0.1660      0.1970      0.84      0.401
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.497
## Multiple R-squared:  0.123, Adjusted R-squared:  0.0376
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 189 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.472  0.895   0.952   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      4.93e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.488 1      1.220
## Year              1.488 16      1.012

```

Residuals from first author



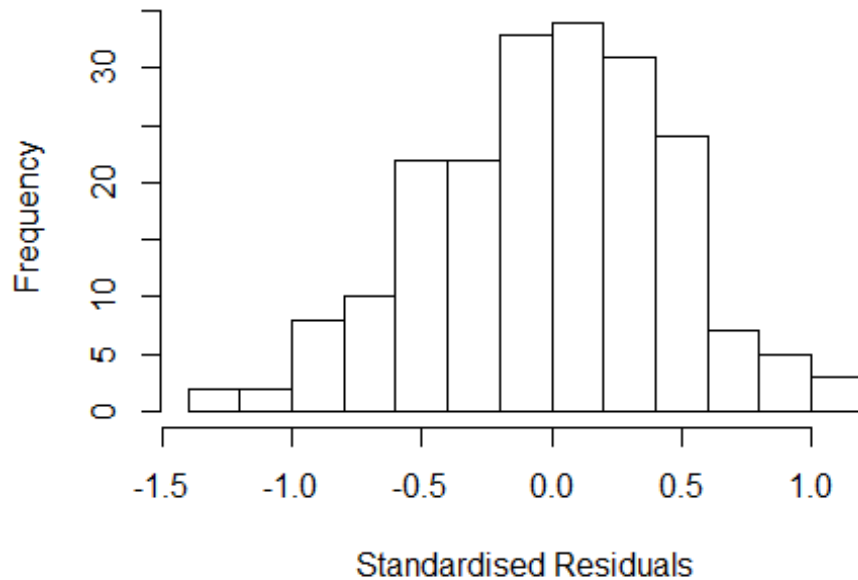
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2858 -0.3293 0.0264 0.3429 1.0958
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7746 0.1758 4.41 1.8e-05 ***
## FirstAuthorFemale1 -0.1057 0.0785 -1.35 0.180
## Year1997 -0.0398 0.2225 -0.18 0.858
## Year1998 -0.0536 0.2158 -0.25 0.804
## Year1999 0.1523 0.2940 0.52 0.605
## Year2000 -0.0199 0.2263 -0.09 0.930
## Year2001 -0.0292 0.2028 -0.14 0.886
## Year2002 -0.1775 0.2591 -0.69 0.494
## Year2003 0.4637 0.3796 1.22 0.223
## Year2004 0.6170 0.3230 1.91 0.058 .
## Year2005 0.1902 0.2423 0.78 0.434
## Year2006 0.0270 0.2119 0.13 0.899
```

```

## Year2007          0.0783      0.2554      0.31      0.760
## Year2008          0.1836      0.2104      0.87      0.384
## Year2009          0.3969      0.2202      1.80      0.073 .
## Year2010          0.3465      0.2158      1.61      0.110
## Year2011         -0.0680      0.2109     -0.32      0.748
## Year2012          0.1663      0.1957      0.85      0.397
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.499
## Multiple R-squared:  0.123, Adjusted R-squared:  0.042
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 187 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.487  0.894  0.952  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      4.93e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.537 1      1.240
## Year            1.537 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
## -1.3263 -0.3420 0.0392 0.3327 1.1412
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7456 0.1790 4.17 4.8e-05 ***
## LastAuthorFemale1 -0.0247 0.0792 -0.31 0.755
## Year1997 -0.0302 0.2285 -0.13 0.895
## Year1998 -0.0849 0.2128 -0.40 0.690
## Year1999 0.1451 0.2901 0.50 0.618
## Year2000 -0.0459 0.2323 -0.20 0.843
## Year2001 -0.0334 0.2065 -0.16 0.872
## Year2002 -0.1907 0.2780 -0.69 0.494
## Year2003 0.4771 0.3702 1.29 0.199
## Year2004 0.6054 0.3359 1.80 0.073 .
## Year2005 0.1912 0.2372 0.81 0.421
## Year2006 0.0101 0.2157 0.05 0.963
```

```

## Year2007          0.1031      0.2482      0.42      0.678
## Year2008          0.1672      0.2161      0.77      0.440
## Year2009          0.3817      0.2237      1.71      0.090 .
## Year2010          0.3367      0.2196      1.53      0.127
## Year2011         -0.0622      0.2163     -0.29      0.774
## Year2012          0.1381      0.1962      0.70      0.482
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.5
## Multiple R-squared:  0.114, Adjusted R-squared:  0.0327
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 188 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.462  0.897  0.953  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      4.93e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 203"
## [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]"
##
## 1999 2001 2007 2008 2009 2010 2011
##    1    1    2    3    1    4    1
##
## 1999 2001 2007 2008 2009 2010 2011
##    1    1    2    3    0    4    1
##
## 1999 2001 2007 2008 2009 2010 2011
##    1    1    2    3    0    4    0

```

```

## [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"
##
## 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: 2.82842712474619"
## [1] "Male last author team size 2018 geometric mean: 1"
##
## Wilcoxon rank sum test
##
## 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"
## [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 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
## 16 16 20 15 18 17 20 10 16 15 18 21 32 40 42
## 2011 2012
## 38 38
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 13 13 17 12 11 10 12 8 15 8 16 15 25 33 32
## 2011 2012
## 30 33
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 12 12 16 11 10 10 12 7 14 7 13 12 23 32 28
## 2011 2012
## 28 28
## [1] "Heteroscedasticity checks, confirming that there are problems with

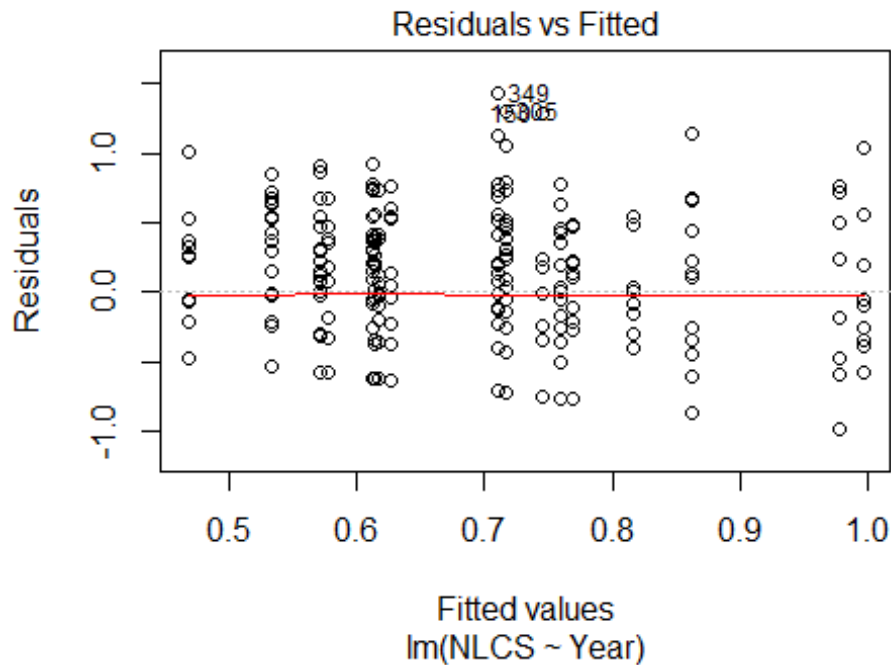
```



```

these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 15, df = 16, p-value = 0.5

```



```

##
## 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: 3.10450802591602"
## [1] "Male first author team size 2018 geometric mean: 1.87762316947124"

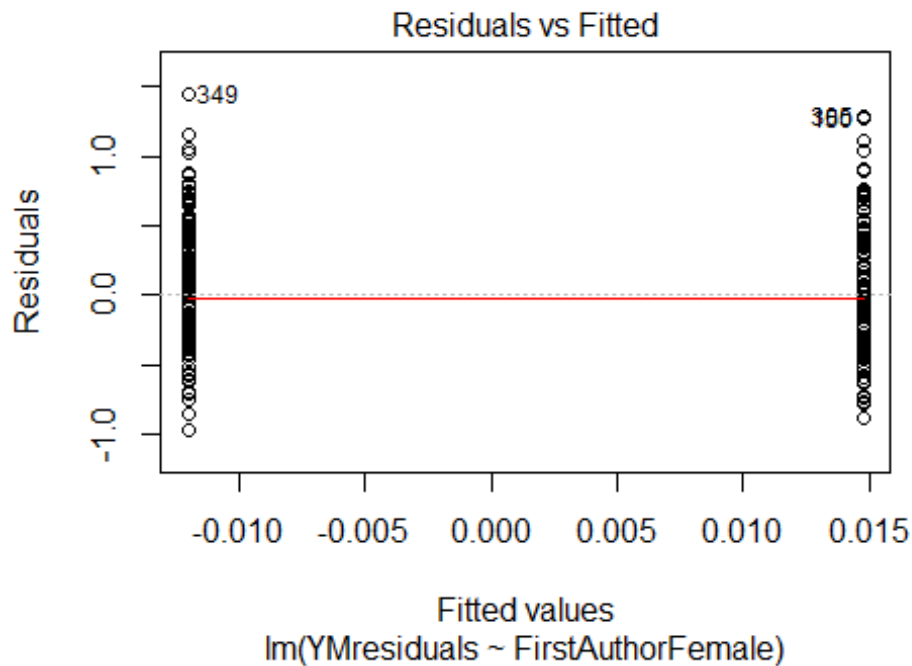
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110, p-value = 0.05
## alternative hypothesis: true location shift is not equal to 0
##

```

```
## [1] "Female last author team size 2018 geometric mean: 2.53104758846168"
## [1] "Male last author team size 2018 geometric mean: 2.35655299057984"

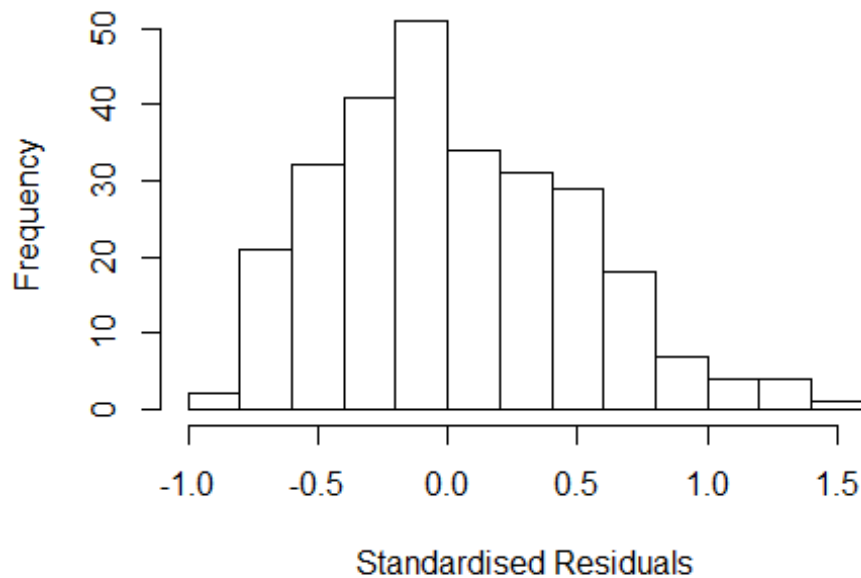
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 79, p-value = 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 | 1.468 | 1 | 1.211 |
| LastAuthorFemale | 1.282 | 1 | 1.132 |
| UniqueAuthors | 2.529 | 4 | 1.123 |
| Year | 3.388 | 16 | 1.039 |

Residuals from first and last author and team size



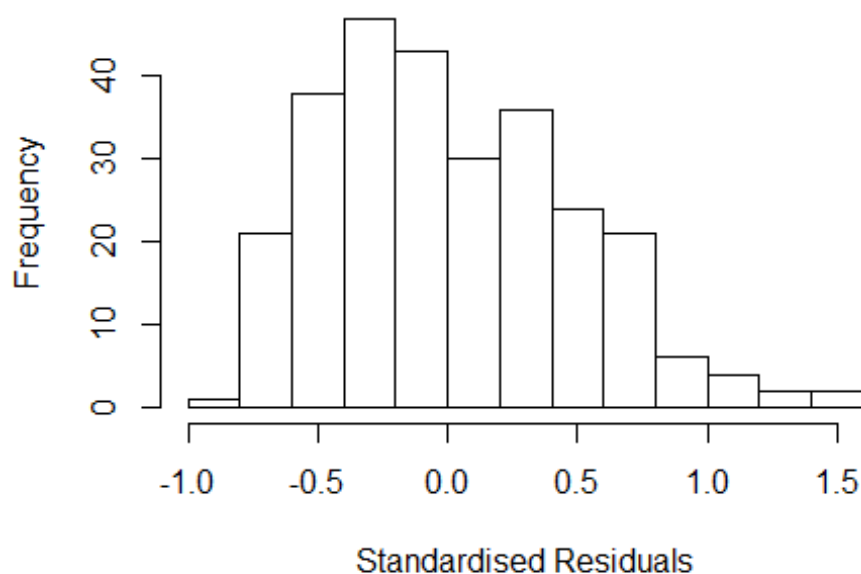
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8599 -0.3322 -0.0585 0.3516 1.4147
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 5.93e-01 1.26e-01 4.72 4e-06 ***
## FirstAuthorFemale1 2.09e-02 6.84e-02 0.31 0.760
## LastAuthorFemale1 2.67e-02 7.68e-02 0.35 0.728
## UniqueAuthors2 3.35e-02 9.26e-02 0.36 0.718
## UniqueAuthors3 -1.30e-01 9.04e-02 -1.44 0.151
## UniqueAuthors4 -2.14e-02 1.25e-01 -0.17 0.864
## UniqueAuthors5 -1.25e-01 1.24e-01 -1.01 0.313
## Year1997 2.13e-01 2.28e-01 0.94 0.351
## Year1998 -1.17e-01 1.55e-01 -0.76 0.451
## Year1999 2.45e-01 1.47e-01 1.67 0.096 .
```

```

## Year2000          1.76e-01    1.49e-01    1.18    0.239
## Year2001          3.88e-01    1.87e-01    2.08    0.039 *
## Year2002          7.56e-05    1.74e-01    0.00    1.000
## Year2003          3.91e-02    2.14e-01    0.18    0.856
## Year2004          7.26e-03    1.84e-01    0.04    0.968
## Year2005          5.24e-01    2.73e-01    1.92    0.056 .
## Year2006          2.48e-01    1.71e-01    1.45    0.148
## Year2007          8.51e-02    1.47e-01    0.58    0.564
## Year2008         -2.40e-02    1.55e-01   -0.16    0.877
## Year2009          1.21e-01    1.49e-01    0.81    0.416
## Year2010          9.03e-02    1.55e-01    0.58    0.561
## Year2011         -7.42e-02    1.57e-01   -0.47    0.637
## Year2012          9.40e-03    1.43e-01    0.07    0.948
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.484
## Multiple R-squared:  0.0972, Adjusted R-squared:  0.0184
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 259 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.372  0.875   0.945   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      3.64e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.364 1      1.168
## LastAuthorFemale  1.214 1      1.102
## Year              1.492 16      1.013

```

Residuals from first and last author



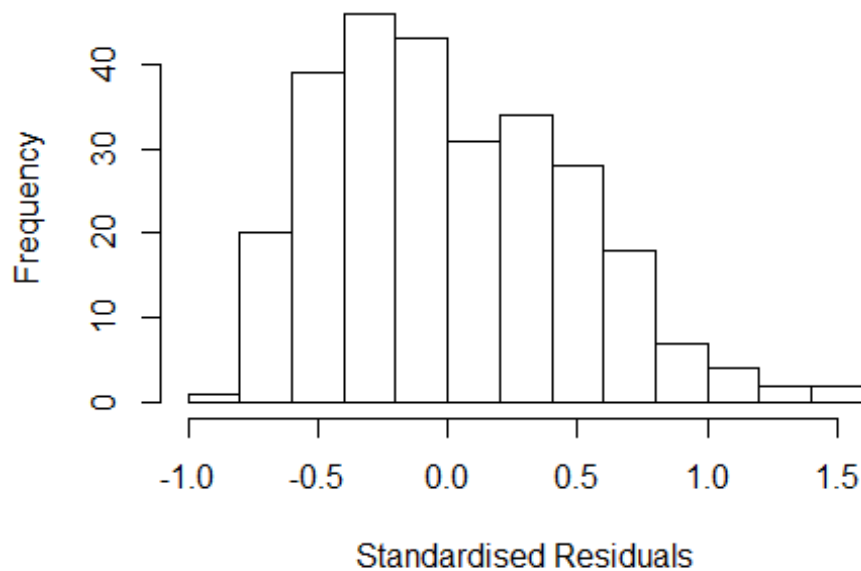
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8239 -0.3470 -0.0274 0.3564 1.4480
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.58233 0.11384 5.12 6.1e-07 ***
## FirstAuthorFemale1 0.00877 0.06765 0.13 0.897
## LastAuthorFemale1 0.03173 0.07737 0.41 0.682
## Year1997 0.19079 0.22137 0.86 0.390
## Year1998 -0.13597 0.15915 -0.85 0.394
## Year1999 0.24745 0.15259 1.62 0.106
## Year2000 0.13399 0.15294 0.88 0.382
## Year2001 0.35930 0.18282 1.97 0.050 .
## Year2002 -0.01506 0.16787 -0.09 0.929
## Year2003 0.02135 0.21957 0.10 0.923
## Year2004 0.01631 0.16908 0.10 0.923
## Year2005 0.52861 0.24895 2.12 0.035 *
```

```

## Year2006          0.20988    0.16758    1.25    0.212
## Year2007          0.05511    0.14878    0.37    0.711
## Year2008         -0.05610    0.15290   -0.37    0.714
## Year2009          0.09870    0.14867    0.66    0.507
## Year2010          0.07889    0.15968    0.49    0.622
## Year2011         -0.11349    0.15373   -0.74    0.461
## Year2012         -0.01312    0.14102   -0.09    0.926
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.517
## Multiple R-squared:  0.0779, Adjusted R-squared:  0.0131
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 246 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.413  0.887  0.948  0.918  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      3.64e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##      trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.268 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
## -0.8029 -0.3503 -0.0332 0.3516 1.4773
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5850 0.1150 5.09 7e-07 ***
## FirstAuthorFemale1 0.0154 0.0653 0.24 0.814
## Year1997 0.1862 0.2223 0.84 0.403
## Year1998 -0.1348 0.1591 -0.85 0.398
## Year1999 0.2532 0.1513 1.67 0.095 .
## Year2000 0.1354 0.1528 0.89 0.376
## Year2001 0.3636 0.1818 2.00 0.047 *
## Year2002 -0.0204 0.1684 -0.12 0.904
## Year2003 0.0203 0.2216 0.09 0.927
## Year2004 0.0167 0.1706 0.10 0.922
## Year2005 0.5308 0.2533 2.10 0.037 *
## Year2006 0.2180 0.1666 1.31 0.192
```

```

## Year2007          0.0562      0.1497      0.38      0.708
## Year2008         -0.0537      0.1547     -0.35      0.729
## Year2009          0.1000      0.1492      0.67      0.503
## Year2010          0.0788      0.1601      0.49      0.623
## Year2011         -0.1060      0.1549     -0.68      0.495
## Year2012         -0.0073      0.1433     -0.05      0.959
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.515
## Multiple R-squared:  0.0776, Adjusted R-squared:  0.0165
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 249 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.391  0.886  0.948  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      3.64e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.138 1      1.067
## Year              1.138 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
## -0.828 -0.348 -0.030  0.353  1.441

```



```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.5857    0.1095    5.35 2e-07 ***
## LastAuthorFemale1 0.0341    0.0749    0.46 0.649
## Year1997        0.1899    0.2212    0.86 0.391
## Year1998       -0.1369    0.1585   -0.86 0.388
## Year1999        0.2482    0.1524    1.63 0.105
## Year2000        0.1358    0.1530    0.89 0.376
## Year2001        0.3603    0.1821    1.98 0.049 *
## Year2002       -0.0117    0.1649   -0.07 0.943
## Year2003        0.0235    0.2219    0.11 0.916
## Year2004        0.0182    0.1680    0.11 0.914
## Year2005        0.5308    0.2492    2.13 0.034 *
## Year2006        0.2077    0.1662    1.25 0.213
## Year2007        0.0542    0.1484    0.37 0.715
## Year2008       -0.0554    0.1530   -0.36 0.717
## Year2009        0.0978    0.1480    0.66 0.509
## Year2010        0.0806    0.1583    0.51 0.611
## Year2011       -0.1154    0.1529   -0.75 0.451
## Year2012       -0.0128    0.1410   -0.09 0.928
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.519
## Multiple R-squared:  0.0778, Adjusted R-squared:  0.0168
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 246 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.420  0.887   0.948   0.918   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      3.64e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 275"
## [1] ""

```

```

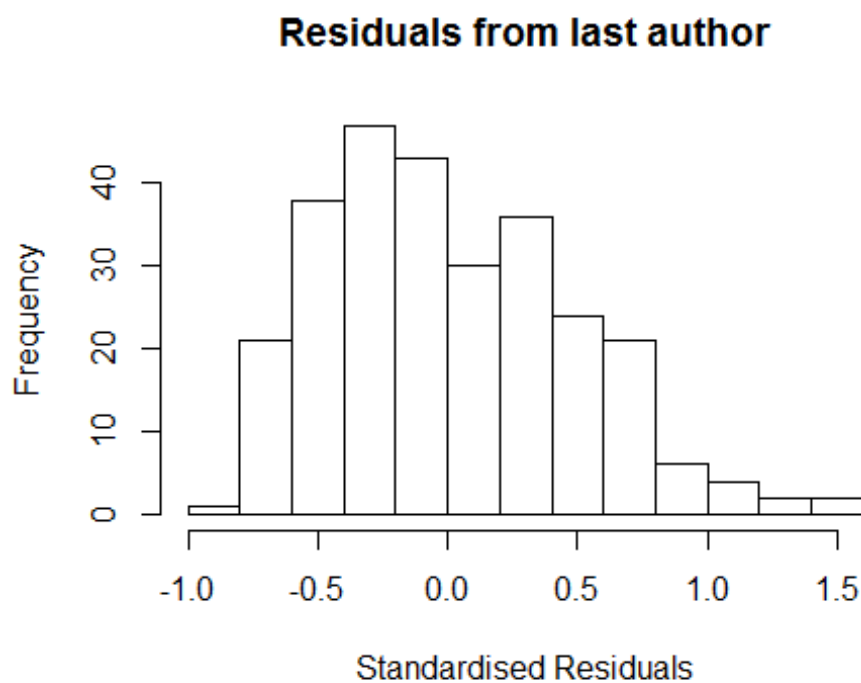
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 3203"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      8      8      5      9     16      7     11      3     12      9      8     15     15     19     21
## 2011 2012
##     23     27
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      6      8      3      5     10      5     11      1     11      7      8     13     13     18     17
## 2011 2012
##     18     25
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      5      7      2      5     10      5      9      1     11      6      7     13     12     17     17
## 2011 2012
##     17     22
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.06672651519506"
## [1] "Male first author team size 2018 geometric mean: 2.2063983750727"

## Warning in wilcox.test.default(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 = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.64562093718978"
## [1] "Male last author team size 2018 geometric mean: 2.58471889411659"

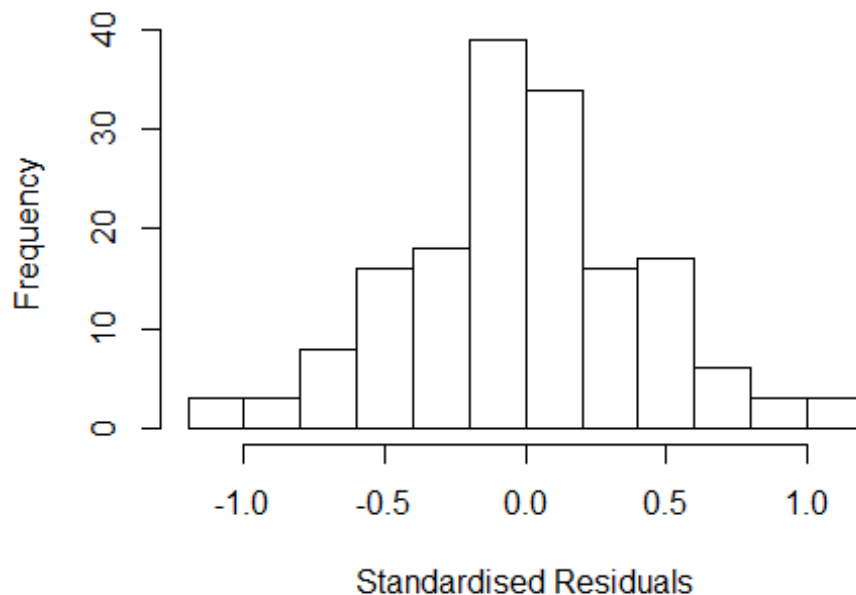
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110, p-value = 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  4.449  1      2.109
## LastAuthorFemale   4.082  1      2.020
## UniqueAuthors     68.390  4      1.696
## Year              139.464 16      1.167
```

Residuals from first and last author and team size



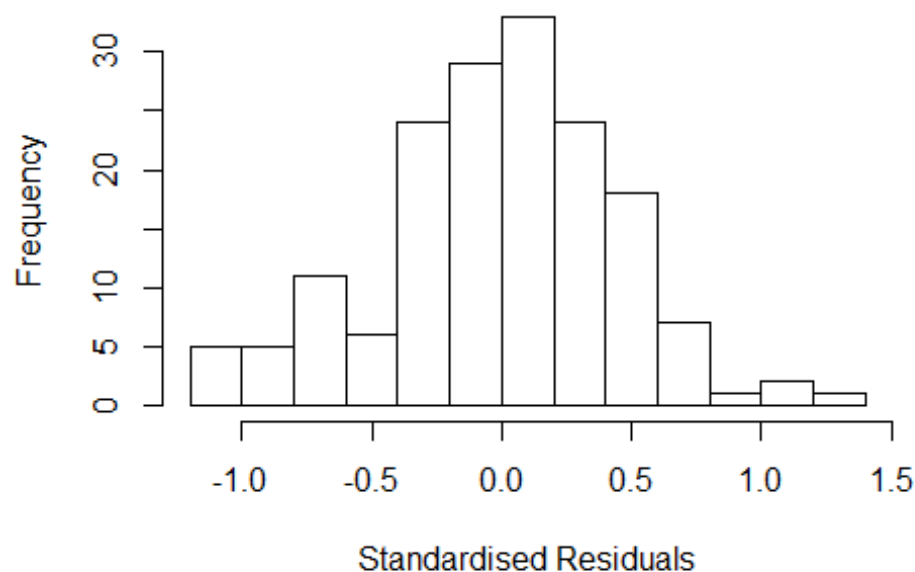
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.19025 -0.25708 -0.00439  0.22869  1.12730
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.49342    0.19762   2.50   0.0137 *
## FirstAuthorFemale1 0.11219    0.08845   1.27   0.2067
## LastAuthorFemale1 -0.00459    0.09653  -0.05   0.9621
## UniqueAuthors2    0.18813    0.10748   1.75   0.0822 .
## UniqueAuthors3    0.35178    0.14490   2.43   0.0164 *
## UniqueAuthors4    0.44644    0.14753   3.03   0.0029 **
## UniqueAuthors5    0.39542    0.14505   2.73   0.0072 **
## Year1997          0.00159    0.28198   0.01   0.9955
## Year1998          0.66932    0.97200   0.69   0.4922
## Year1999          0.16293    0.25579   0.64   0.5252
```

```

## Year2000          0.44364      0.26114      1.70      0.0915 .
## Year2001          0.40515      0.20038      2.02      0.0450 *
## Year2002          0.20997      0.24801      0.85      0.3986
## Year2003          1.11358      0.19762      5.63      9e-08 ***
## Year2004          0.18921      0.21650      0.87      0.3836
## Year2005          0.36598      0.26129      1.40      0.1635
## Year2006          0.17037      0.21796      0.78      0.4357
## Year2007          0.33851      0.19895      1.70      0.0910 .
## Year2008          0.15189      0.18986      0.80      0.4250
## Year2009          0.24593      0.21786      1.13      0.2609
## Year2010          0.22696      0.24731      0.92      0.3603
## Year2011          0.25046      0.22821      1.10      0.2743
## Year2012          0.37157      0.20579      1.81      0.0731 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.388
## Multiple R-squared:  0.235, Adjusted R-squared:  0.117
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 146 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.327  0.841  0.951  0.885  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          6.02e-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 4.475 1      2.115
## LastAuthorFemale  3.435 1      1.853
## Year              6.230 16      1.059

```

Residuals from first and last author



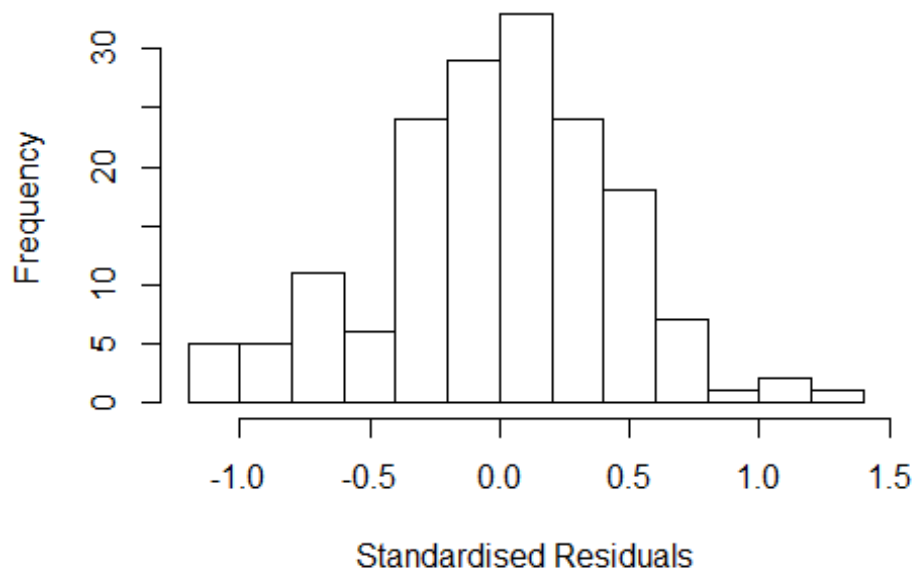
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1299 -0.2539 0.0142 0.2816 1.2387
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.684150 0.195861 3.49 0.00063 ***
## FirstAuthorFemale1 0.146088 0.092709 1.58 0.11723
## LastAuthorFemale1 -0.000168 0.090972 0.00 0.99853
## Year1997 -0.036853 0.244673 -0.15 0.88048
## Year1998 0.570434 0.581378 0.98 0.32812
## Year1999 0.180806 0.249612 0.72 0.47000
## Year2000 0.417091 0.251092 1.66 0.09882 .
## Year2001 0.354892 0.222101 1.60 0.11222
## Year2002 0.209411 0.260804 0.80 0.42330
## Year2003 0.922850 0.195861 4.71 5.6e-06 ***
## Year2004 0.246537 0.226769 1.09 0.27874
## Year2005 0.322329 0.299107 1.08 0.28296
```

```

## Year2006          0.034097    0.230242    0.15  0.88247
## Year2007          0.291212    0.218545    1.33  0.18476
## Year2008          0.111703    0.213093    0.52  0.60093
## Year2009          0.364058    0.246712    1.48  0.14218
## Year2010          0.232625    0.245248    0.95  0.34442
## Year2011          0.299807    0.252756    1.19  0.23747
## Year2012          0.507298    0.230620    2.20  0.02939 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.424
## Multiple R-squared:  0.137, Adjusted R-squared:  0.0312
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 148 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.373  0.848  0.947  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      6.02e-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 3.547 1      1.883
## Year              3.547 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
## -1.1307 -0.2538 0.0139 0.2814 1.2387
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6844 0.1932 3.54 0.00053 ***
## FirstAuthorFemale1 0.1458 0.0832 1.75 0.08190 .
## Year1997 -0.0369 0.2422 -0.15 0.87924
## Year1998 0.5701 0.5830 0.98 0.32973
## Year1999 0.1805 0.2480 0.73 0.46792
## Year2000 0.4178 0.2488 1.68 0.09519 .
## Year2001 0.3547 0.2207 1.61 0.11004
## Year2002 0.2092 0.2551 0.82 0.41352
## Year2003 0.9226 0.1932 4.78 4.3e-06 ***
## Year2004 0.2471 0.2263 1.09 0.27666
## Year2005 0.3222 0.2981 1.08 0.28155
## Year2006 0.0328 0.2302 0.14 0.88687
```



```

## Year2007          0.2908      0.2183      1.33  0.18478
## Year2008          0.1115      0.2124      0.52  0.60050
## Year2009          0.3640      0.2465      1.48  0.14195
## Year2010          0.2329      0.2461      0.95  0.34546
## Year2011          0.3005      0.2528      1.19  0.23653
## Year2012          0.5069      0.2302      2.20  0.02923 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.421
## Multiple R-squared:  0.137, Adjusted R-squared:  0.0382
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 148 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.368  0.847  0.946  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      6.02e-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 2.96 1      1.720
## Year      2.96 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
## -1.1303 -0.2581  0.0245  0.2702  1.2979

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.6715    0.1936    3.47 0.00069 ***
## LastAuthorFemale1 0.0525    0.0813    0.65 0.51939
## Year1997        0.0157    0.2385    0.07 0.94746
## Year1998        0.5567    0.6401    0.87 0.38583
## Year1999        0.2446    0.2323    1.05 0.29395
## Year2000        0.4588    0.2494    1.84 0.06788 .
## Year2001        0.4133    0.2055    2.01 0.04609 *
## Year2002        0.2923    0.2579    1.13 0.25885
## Year2003        0.9355    0.1936    4.83 3.3e-06 ***
## Year2004        0.3190    0.2210    1.44 0.15096
## Year2005        0.3931    0.3063    1.28 0.20135
## Year2006        0.0465    0.2248    0.21 0.83627
## Year2007        0.3114    0.2201    1.42 0.15915
## Year2008        0.1919    0.2007    0.96 0.34044
## Year2009        0.4312    0.2412    1.79 0.07586 .
## Year2010        0.3052    0.2344    1.30 0.19482
## Year2011        0.3935    0.2431    1.62 0.10769
## Year2012        0.5820    0.2212    2.63 0.00942 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.419
## Multiple R-squared:  0.125, Adjusted R-squared:  0.0245
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 150 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.318  0.842  0.955  0.890  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      6.02e-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: 166"
## [1] ""

```

```

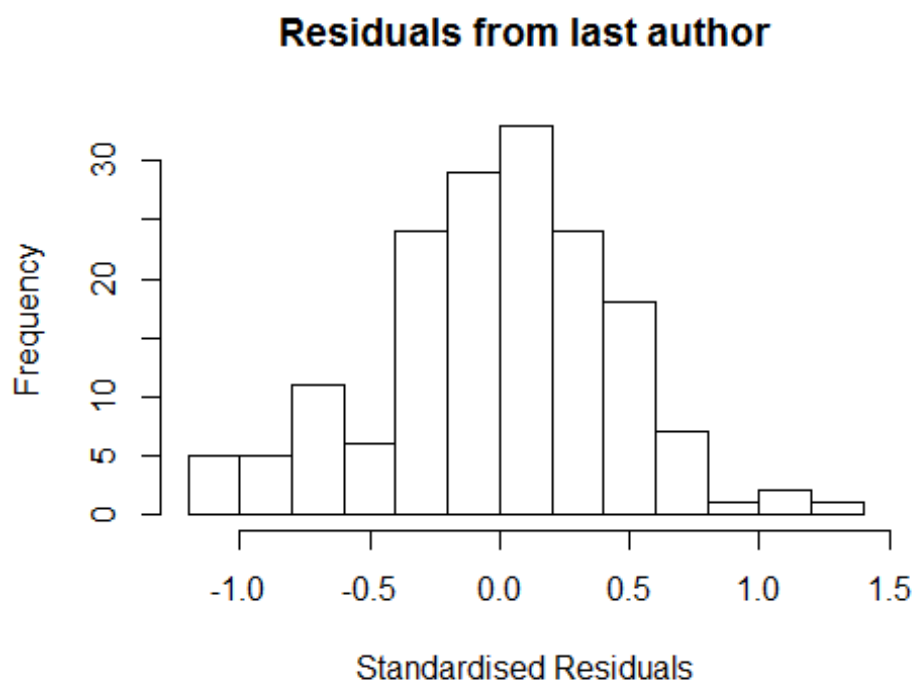
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 3204"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    1    5    8    7    3    5    4    7    8    6   15   16   20   32
## 2011 2012
##   29   26
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    1    4    5    5    2    4    3    5    7    4   14   14   19   27
## 2011 2012
##   24   24
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    1    3    5    5    1    2    3    5    7    4   12   12   15   21
## 2011 2012
##   22   21
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.79556084991966"
## [1] "Male first author team size 2018 geometric mean: 1.86554050816254"

## Warning in wilcox.test.default(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 = 0.03
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.88717908427182"
## [1] "Male last author team size 2018 geometric mean: 1.71526732051997"

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

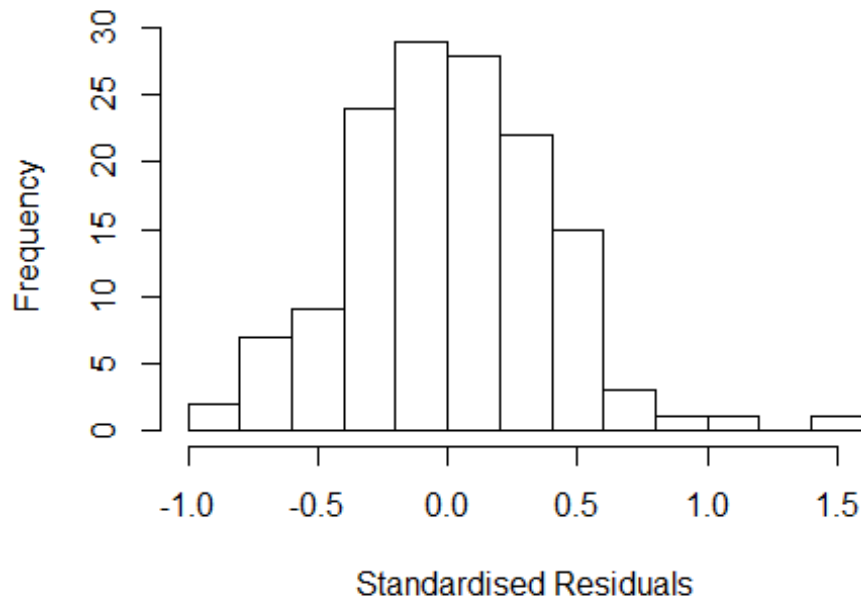
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 400, p-value = 0.005
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
```

| | GVIF | Df | GVIF ^{1/(2*Df)} |
|----------------------|------------|----|--------------------------|
| ## FirstAuthorFemale | -5.745e+14 | 1 | NaN |
| ## LastAuthorFemale | -5.288e+14 | 1 | NaN |
| ## UniqueAuthors | 3.322e+05 | 4 | 4.9 |
| ## Year | -8.470e+20 | 16 | NaN |

Residuals from first and last author and team size



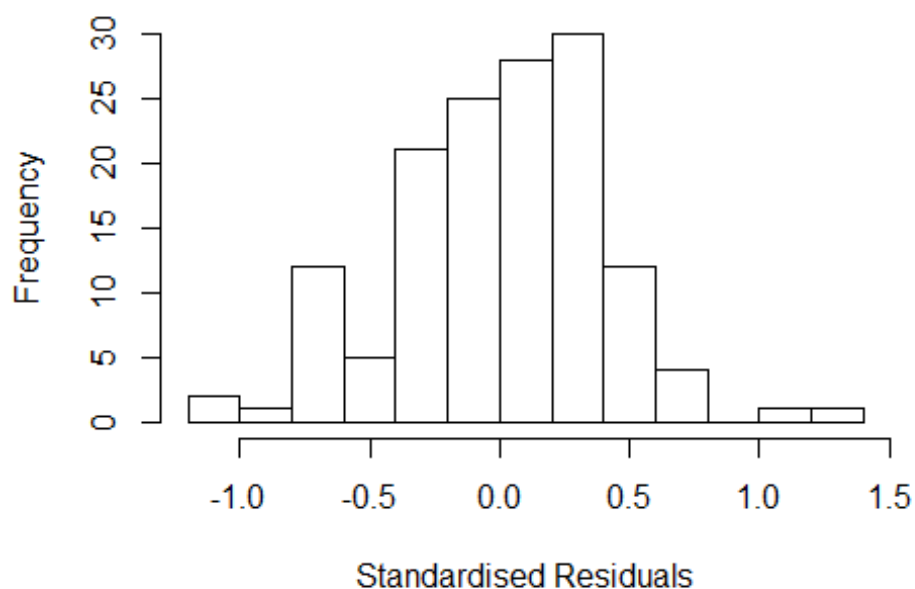
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.924288 -0.241658  0.000188  0.254171  1.458956
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9131    0.1386    6.59 1.3e-09 ***
## FirstAuthorFemale1 -0.0574    0.0874   -0.66  0.5124
## LastAuthorFemale1 -0.0871    0.0839   -1.04  0.3010
## UniqueAuthors2    -0.0449    0.1052   -0.43  0.6706
## UniqueAuthors3     0.4094    0.1286    3.18  0.0019 **
## UniqueAuthors4     0.2788    0.1381    2.02  0.0458 *
## UniqueAuthors5     0.0755    0.1155    0.65  0.5149
## Year1997          -0.6581    0.1386   -4.75 5.8e-06 ***
## Year1998           0.4461    0.2623    1.70  0.0916 .
## Year1999           0.3164    0.2504    1.26  0.2087
```

```

## Year2000          0.4509      0.1786      2.52      0.0129 *
## Year2001         -0.2176      0.1731     -1.26      0.2111
## Year2002         -0.0908      0.1515     -0.60      0.5503
## Year2003         -0.1168      0.1916     -0.61      0.5432
## Year2004         -0.0961      0.2302     -0.42      0.6771
## Year2005          0.0356      0.1662      0.21      0.8309
## Year2006         -0.3568      0.2189     -1.63      0.1058
## Year2007          0.1135      0.2147      0.53      0.5981
## Year2008          0.1970      0.1619      1.22      0.2260
## Year2009          0.1710      0.1901      0.90      0.3703
## Year2010          0.2610      0.1768      1.48      0.1425
## Year2011          0.1303      0.1982      0.66      0.5120
## Year2012          0.3326      0.1837      1.81      0.0727 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.39
## Multiple R-squared:  0.249, Adjusted R-squared:  0.11
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 127 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.132  0.879  0.952  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      7.04e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 4.139e+14 1      2.034e+07
## LastAuthorFemale  4.834e+14 1      2.199e+07
## Year              9.714e+14 16      2.940e+00

```

Residuals from first and last author



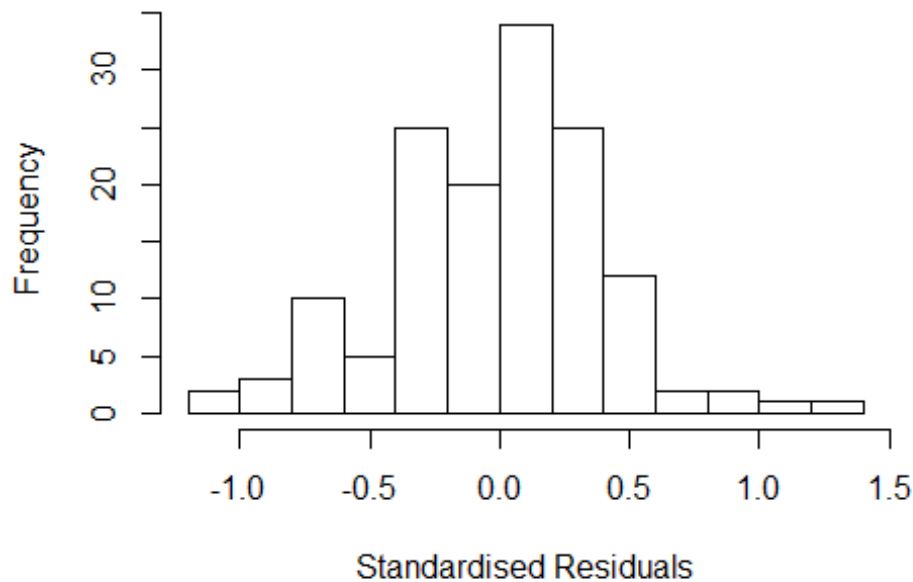
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.111 -0.266 0.031 0.248 1.263
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.90781 0.14370 6.32 4.4e-09 ***
## FirstAuthorFemale1 0.03412 0.08301 0.41 0.682
## LastAuthorFemale1 -0.07735 0.08971 -0.86 0.390
## Year1997 -0.65281 0.14370 -4.54 1.3e-05 ***
## Year1998 0.41838 0.29051 1.44 0.152
## Year1999 0.21613 0.25307 0.85 0.395
## Year2000 0.35028 0.17326 2.02 0.045 *
## Year2001 -0.31358 0.17317 -1.81 0.073 .
## Year2002 0.11437 0.21248 0.54 0.591
## Year2003 -0.00934 0.21475 -0.04 0.965
## Year2004 -0.07863 0.21727 -0.36 0.718
## Year2005 0.13521 0.18982 0.71 0.478
```

```

## Year2006      -0.38961    0.20413   -1.91    0.059 .
## Year2007      0.16902    0.19830    0.85    0.396
## Year2008      0.14263    0.16652    0.86    0.393
## Year2009      0.28089    0.19593    1.43    0.154
## Year2010      0.25167    0.18531    1.36    0.177
## Year2011      0.18504    0.19861    0.93    0.353
## Year2012      0.32343    0.19079    1.70    0.093 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.408
## Multiple R-squared:  0.148, Adjusted R-squared:  0.0238
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 128 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.317  0.892  0.955  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      7.04e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -6.133e+14 1      NaN
## Year      -6.133e+14 16      NaN

```


Residuals from first author



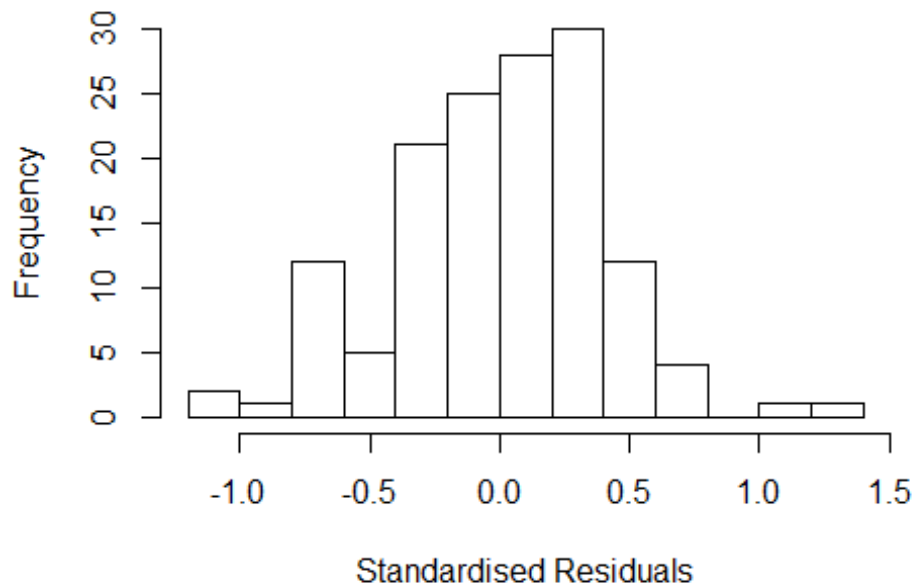
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0709 -0.2610 0.0334 0.2508 1.2421
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.91598 0.13512 6.78 4.4e-10 ***
## FirstAuthorFemale1 0.00638 0.07835 0.08 0.935
## Year1997 -0.66098 0.13512 -4.89 3.0e-06 ***
## Year1998 0.40986 0.30557 1.34 0.182
## Year1999 0.20668 0.25916 0.80 0.427
## Year2000 0.31911 0.16196 1.97 0.051 .
## Year2001 -0.37136 0.15012 -2.47 0.015 *
## Year2002 0.06752 0.17904 0.38 0.707
## Year2003 -0.02485 0.20350 -0.12 0.903
## Year2004 -0.07651 0.21642 -0.35 0.724
## Year2005 0.10419 0.17674 0.59 0.557
## Year2006 -0.40306 0.18896 -2.13 0.035 *
```

```

## Year2007          0.14321    0.19173    0.75    0.457
## Year2008          0.11574    0.15598    0.74    0.459
## Year2009          0.24461    0.17678    1.38    0.169
## Year2010          0.20727    0.16974    1.22    0.224
## Year2011          0.14857    0.18923    0.79    0.434
## Year2012          0.28874    0.17532    1.65    0.102
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.415
## Multiple R-squared:  0.141, Adjusted R-squared:  0.0236
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 126 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.350  0.884  0.957  0.908  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      7.04e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.282e+15 1      3.581e+07
## Year              1.282e+15 16      2.966e+00

```

Residuals from last author



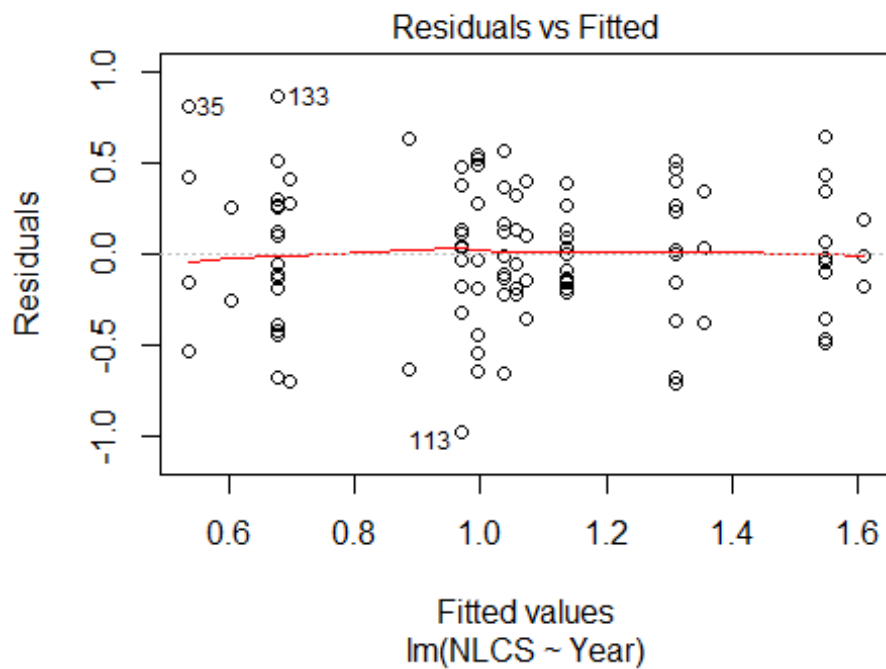
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0808 -0.2626 0.0372 0.2450 1.2683
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.918156 0.133446 6.88 2.6e-10 ***
## LastAuthorFemale1 -0.065753 0.084874 -0.77 0.440
## Year1997 -0.663156 0.133446 -4.97 2.2e-06 ***
## Year1998 0.427808 0.279416 1.53 0.128
## Year1999 0.231309 0.236777 0.98 0.331
## Year2000 0.360110 0.160755 2.24 0.027 *
## Year2001 -0.301403 0.159077 -1.89 0.060 .
## Year2002 0.098221 0.200898 0.49 0.626
## Year2003 0.000514 0.209803 0.00 0.998
## Year2004 -0.071236 0.209322 -0.34 0.734
## Year2005 0.133516 0.183097 0.73 0.467
## Year2006 -0.385785 0.194906 -1.98 0.050 *
```

```

## Year2007      0.162660    0.189690    0.86    0.393
## Year2008      0.148813    0.155518    0.96    0.340
## Year2009      0.286488    0.187594    1.53    0.129
## Year2010      0.254395    0.175056    1.45    0.149
## Year2011      0.192341    0.191527    1.00    0.317
## Year2012      0.333647    0.181839    1.83    0.069 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.405
## Multiple R-squared:  0.148, Adjusted R-squared:  0.0308
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 130 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.305  0.892  0.954  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      7.04e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 142"
## [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
##    6    5    2    3    3    7    3    3   13    9    7   10   14   12    4
## 2011 2012
##   12   12
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    2    0    3    3    5    3    2   12    9    5   10   12   12    4
## 2011 2012

```

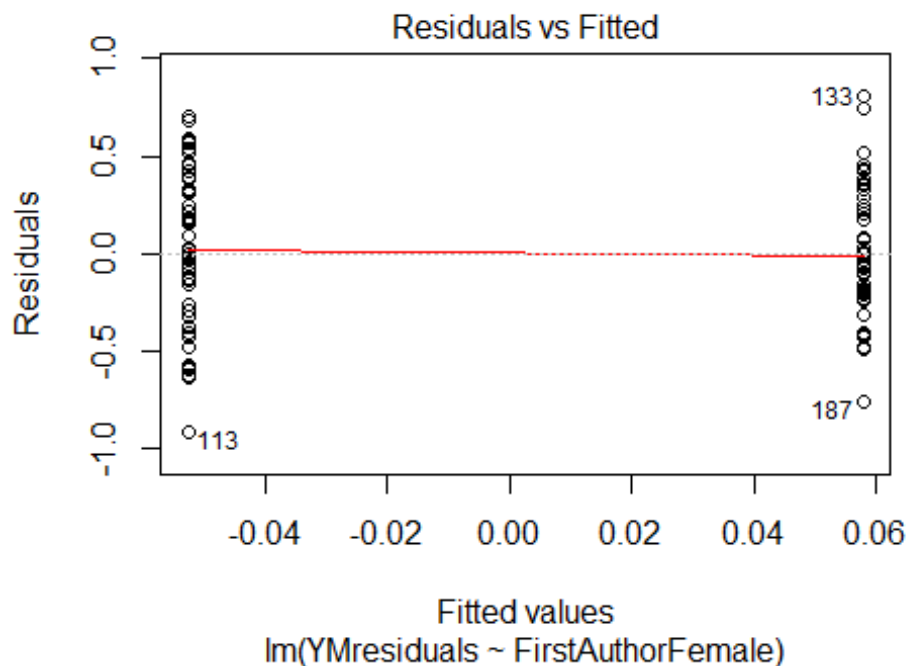
```
## 11 11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3 2 0 2 3 5 2 1 11 8 4 9 12 12 4
## 2011 2012
## 10 10
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 16, df = 15, p-value = 0.4
```



```
##
## 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: 2.39492974229681"
## [1] "Male first author team size 2018 geometric mean: 1.79349486065659"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

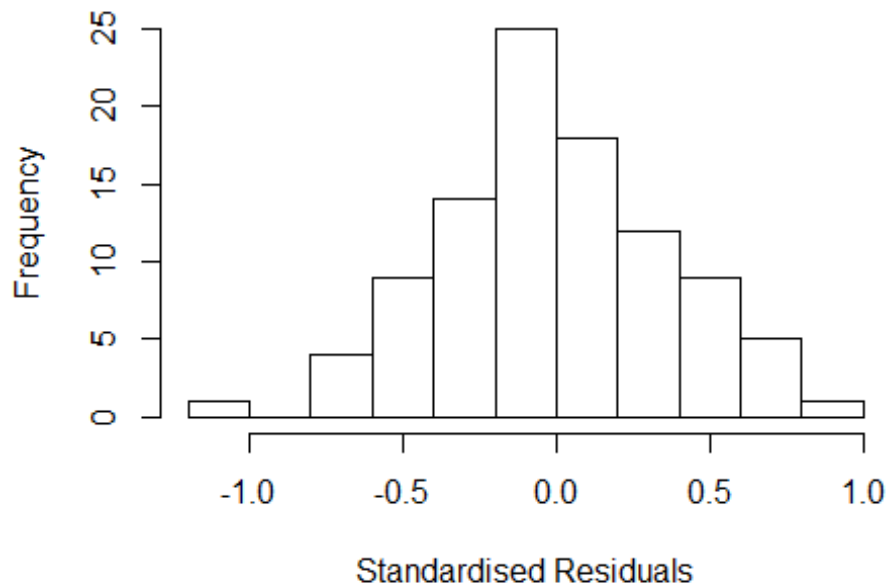
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 51, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.70586577923534"
## [1] "Male last 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 = 63, 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  6.429  1      2.536
## LastAuthorFemale  11.078  1      3.328
## UniqueAuthors    65.103  4      1.685
## Year             132.091 15      1.177
```

Residuals from first and last author and team size



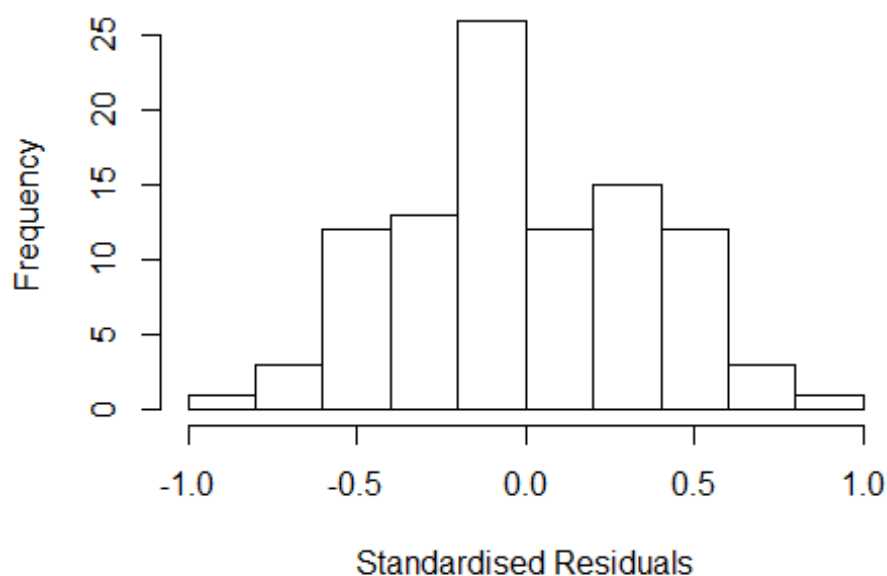
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0173 -0.2237 -0.0303 0.2409 0.8055
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1343 0.1879 6.04 5.4e-08 ***
## FirstAuthorFemale1 0.0998 0.1124 0.89 0.3771
## LastAuthorFemale1 -0.0107 0.1191 -0.09 0.9287
## UniqueAuthors2 0.0710 0.1202 0.59 0.5565
## UniqueAuthors3 0.1440 0.1368 1.05 0.2960
## UniqueAuthors4 0.3979 0.1285 3.10 0.0027 **
## UniqueAuthors5 0.2158 0.1566 1.38 0.1722
## Year1997 -0.2488 1.0631 -0.23 0.8156
## Year1999 -0.6909 0.4659 -1.48 0.1422
## Year2000 0.3414 0.2126 1.61 0.1125
```

```

## Year2001          -0.7326      0.2986   -2.45   0.0164 *
## Year2002          -0.4648      0.3808   -1.22   0.2260
## Year2003          -0.3605      0.1948   -1.85   0.0682 .
## Year2004          -0.1544      0.1848   -0.84   0.4062
## Year2005          -0.2681      0.2651   -1.01   0.3151
## Year2006          -0.3492      0.1784   -1.96   0.0540 .
## Year2007          -0.3194      0.1915   -1.67   0.0995 .
## Year2008          -0.2503      0.2023   -1.24   0.2199
## Year2009          -0.6014      0.2076   -2.90   0.0049 **
## Year2010          -0.1760      0.2401   -0.73   0.4657
## Year2011           0.1841      0.1968    0.94   0.3524
## Year2012          -0.0604      0.2487   -0.24   0.8087
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.38
## Multiple R-squared:  0.415, Adjusted R-squared:  0.253
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 91 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.454  0.880  0.960  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          1.02e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 4.063 1 2.016
## LastAuthorFemale 5.473 1 2.339
## Year 5.490 15 1.058

```


Residuals from first and last author



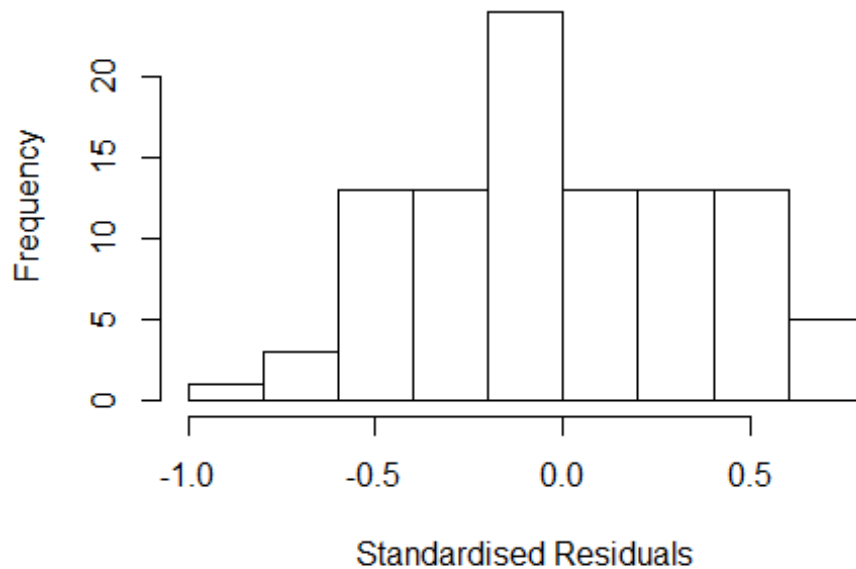
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8937 -0.2340 -0.0224 0.2867 0.8082
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2567 0.1973 6.37 1.1e-08 ***
## FirstAuthorFemale1 0.1614 0.1004 1.61 0.1120
## LastAuthorFemale1 -0.0676 0.1008 -0.67 0.5046
## Year1997 -0.3712 0.8891 -0.42 0.6774
## Year1999 -0.8156 0.4472 -1.82 0.0719 .
## Year2000 0.2871 0.2337 1.23 0.2229
## Year2001 -0.8115 0.3202 -2.53 0.0132 *
## Year2002 -0.5152 0.4585 -1.12 0.2645
## Year2003 -0.4875 0.2110 -2.31 0.0235 *
## Year2004 -0.1791 0.2037 -0.88 0.3818
## Year2005 -0.3123 0.3051 -1.02 0.3091
## Year2006 -0.3242 0.2339 -1.39 0.1696
```

```

## Year2007          -0.2579      0.2263    -1.14    0.2579
## Year2008          -0.2954      0.2209    -1.34    0.1849
## Year2009          -0.6157      0.2281    -2.70    0.0085 **
## Year2010          -0.2094      0.2591    -0.81    0.4213
## Year2011           0.1765      0.2193     0.80    0.4232
## Year2012          -0.0225      0.2570    -0.09    0.9305
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.409
## Multiple R-squared:  0.374, Adjusted R-squared:  0.241
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 90 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.613  0.894  0.952  0.924  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.02e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.008 1          1.734
## Year              3.008 15          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
## -0.9399 -0.2266 -0.0372 0.3057 0.7909
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2747 0.1935 6.59 4.2e-09 ***
## FirstAuthorFemale1 0.1328 0.0873 1.52 0.1321
## Year1997 -0.3892 0.9205 -0.42 0.6736
## Year1999 -0.8531 0.4227 -2.02 0.0468 *
## Year2000 0.2422 0.2300 1.05 0.2956
## Year2001 -0.8224 0.3291 -2.50 0.0145 *
## Year2002 -0.5332 0.4593 -1.16 0.2491
## Year2003 -0.5445 0.1977 -2.75 0.0073 **
## Year2004 -0.2024 0.1994 -1.02 0.3130
## Year2005 -0.3464 0.2907 -1.19 0.2369
## Year2006 -0.3274 0.2295 -1.43 0.1574
## Year2007 -0.2923 0.2240 -1.31 0.1955
```

```

## Year2008          -0.3349      0.2171   -1.54    0.1269
## Year2009          -0.6555      0.2197   -2.98    0.0038 **
## Year2010          -0.2631      0.2454   -1.07    0.2868
## Year2011           0.1508      0.2163    0.70    0.4877
## Year2012          -0.0524      0.2572   -0.20    0.8390
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.372, Adjusted R-squared:  0.247
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 90 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.564  0.888  0.950  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      1.02e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##  nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##  trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 4.416  1          2.101
## Year            4.416 15          1.051

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.00863 -0.20370 -0.00194  0.27757  0.86793
##

```

```

## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3561    0.1836   7.39 1.2e-10 ***
## LastAuthorFemale1  0.0124    0.0858   0.14  0.8855
## Year1997         -0.4706    0.9097  -0.52  0.6063
## Year1999         -0.8743    0.5075  -1.72  0.0887 .
## Year2000          0.2429    0.2151   1.13  0.2622
## Year2001         -0.8691    0.3735  -2.33  0.0225 *
## Year2002         -0.6146    0.4548  -1.35  0.1803
## Year2003         -0.5055    0.2029  -2.49  0.0147 *
## Year2004         -0.2272    0.1934  -1.17  0.2436
## Year2005         -0.3534    0.2959  -1.19  0.2358
## Year2006         -0.3381    0.2146  -1.58  0.1190
## Year2007         -0.3178    0.2199  -1.45  0.1522
## Year2008         -0.3599    0.2095  -1.72  0.0896 .
## Year2009         -0.6935    0.2190  -3.17  0.0022 **
## Year2010         -0.2951    0.2422  -1.22  0.2265
## Year2011          0.1404    0.2203   0.64  0.5258
## Year2012         -0.0536    0.2468  -0.22  0.8287
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.404
## Multiple R-squared:  0.359, Adjusted R-squared:  0.232
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 86 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.513  0.880   0.936   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      1.02e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500           50           2           1           1000      200
##      trace.lev      mts      compute.rd
##      0             1000      0
##      psi          subsampling          cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 98"
## [1] ""
## [1] ""
## [1] "#####"

```

```

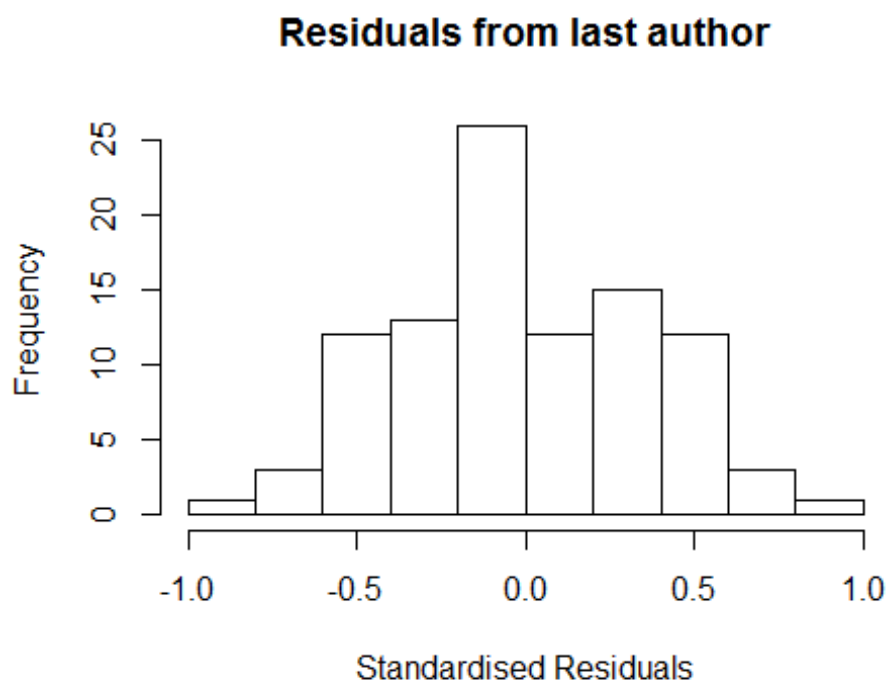
## [1] "Analysis of AJSC 3206"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    4    2    7   12    3    2    3    4    5    5    4    5    3
## 2011 2012
##   10    5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    0    2    1    1    3    2    3    4    3    5    4    5    3
## 2011 2012
##    9    5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    0    2    1    1    3    2    2    4    3    5    4    5    3
## 2011 2012
##    8    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: 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 = 1.5, p-value = 1
## 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: 2"

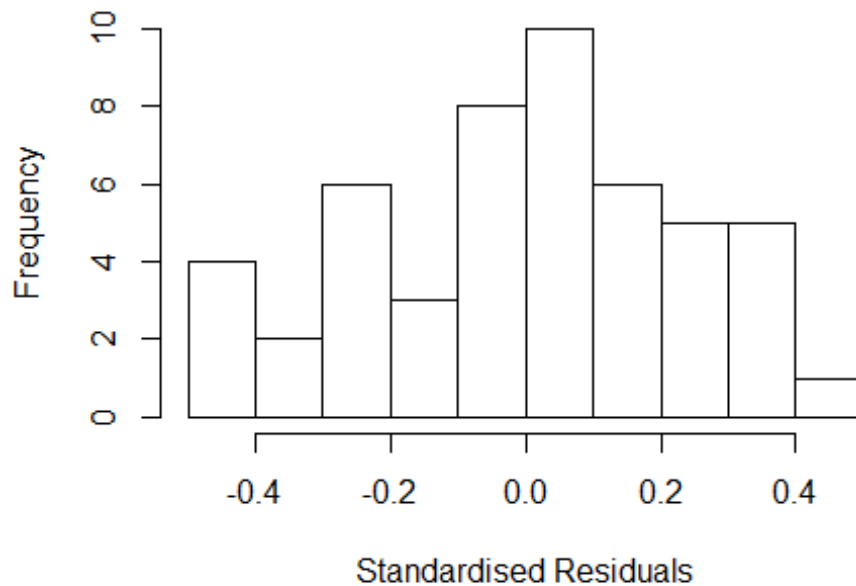
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3, 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 4.683e+13  1      6.843e+06
## LastAuthorFemale  5.092e+13  1      7.136e+06
## UniqueAuthors    2.095e+16  4      1.097e+02
## Year              1.250e+17 15      3.714e+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
## -0.48147 -0.16932 0.00715 0.17312 0.44014
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.14893 0.16808 6.84 2.0e-07 ***
## FirstAuthorFemale1 0.13634 0.08767 1.56 0.13114
## LastAuthorFemale1 -0.44000 0.09037 -4.87 4.0e-05 ***
## UniqueAuthors2 0.28991 0.14791 1.96 0.06001 .
## UniqueAuthors3 0.29227 0.16714 1.75 0.09131 .
## UniqueAuthors4 0.60059 0.15568 3.86 0.00061 ***
## UniqueAuthors5 0.13121 0.17050 0.77 0.44799
## Year1997 -0.89393 0.16808 -5.32 1.2e-05 ***
## Year1999 0.08976 0.11696 0.77 0.44924
## Year2000 0.29382 0.12272 2.39 0.02359 *
```

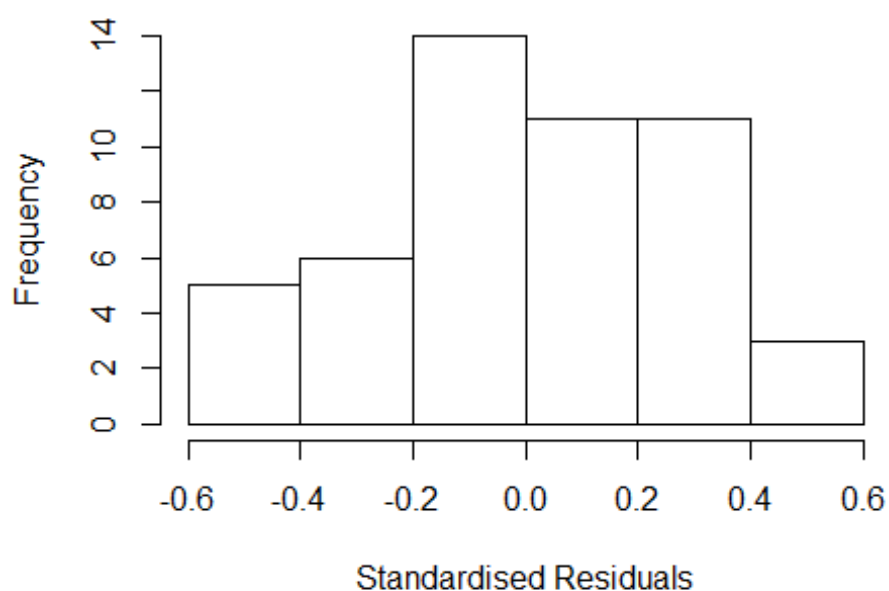


```

## Year2001      -1.14893    0.16808    -6.84    2.0e-07 ***
## Year2002      -0.55715    0.18611    -2.99    0.00571 **
## Year2003      -1.03243    0.18824    -5.48    7.4e-06 ***
## Year2004       0.00966    0.55852     0.02    0.98633
## Year2005       0.06981    0.09070     0.77    0.44797
## Year2006      -0.62369    0.17326    -3.60    0.00122 **
## Year2007      -0.00267    0.22964    -0.01    0.99080
## Year2008      -0.64215    0.16263    -3.95    0.00048 ***
## Year2009      -0.53497    0.10387    -5.15    1.8e-05 ***
## Year2010       0.09062    0.14381     0.63    0.53372
## Year2011      -0.10412    0.08151    -1.28    0.21195
## Year2012      -0.43538    0.23896    -1.82    0.07916 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.299
## Multiple R-squared:  0.752, Adjusted R-squared:  0.566
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 40 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.777  0.898  0.946  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           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 3.389e+14 1 1.841e+07
## LastAuthorFemale 4.770e+14 1 2.184e+07
## Year 2.450e+15 15 3.258e+00

```

Residuals from first and last author



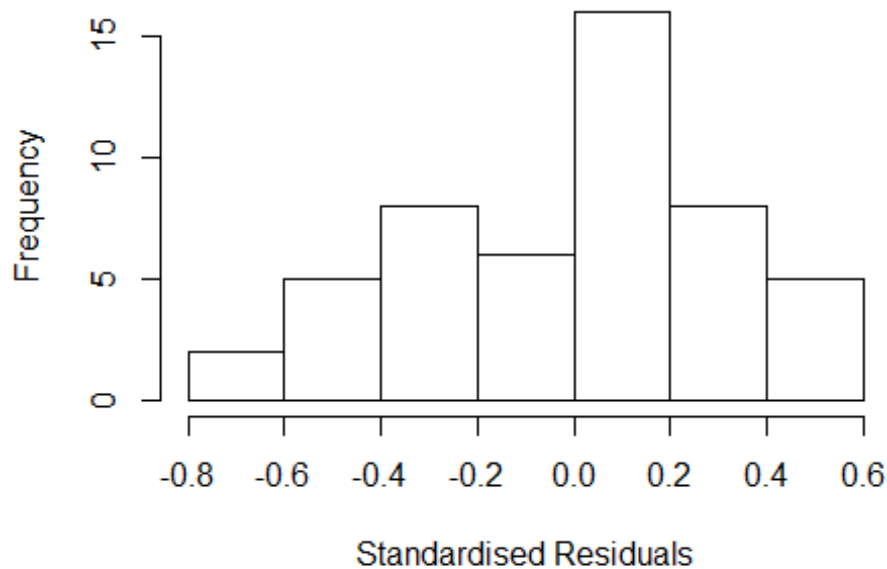
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.58477 -0.17936 0.00164 0.20572 0.55323
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4784 0.1929 7.66 9.8e-09 ***
## FirstAuthorFemale1 0.1818 0.1011 1.80 0.0816 .
## LastAuthorFemale1 -0.4126 0.1200 -3.44 0.0016 **
## Year1997 -1.2234 0.1929 -6.34 4.0e-07 ***
## Year1999 -0.1300 0.2075 -0.63 0.5353
## Year2000 0.1814 0.1929 0.94 0.3541
## Year2001 -1.4784 0.1929 -7.66 9.8e-09 ***
## Year2002 -0.7584 0.2244 -3.38 0.0019 **
## Year2003 -1.3619 0.2104 -6.47 2.8e-07 ***
## Year2004 -0.1308 0.7732 -0.17 0.8667
## Year2005 0.0373 0.1829 0.20 0.8395
## Year2006 -0.6516 0.4149 -1.57 0.1261
```

```

## Year2007          -0.2554      0.2946   -0.87    0.3924
## Year2008          -0.6913      0.2274   -3.04    0.0047 **
## Year2009          -0.6412      0.2123   -3.02    0.0049 **
## Year2010          -0.0556      0.2181   -0.25    0.8004
## Year2011          -0.1578      0.2112   -0.75    0.4604
## Year2012          -0.5888      0.2783   -2.12    0.0422 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.352
## Multiple R-squared:  0.647, Adjusted R-squared:  0.46
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 41 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.764  0.906  0.946  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      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 1.301e+14 1      1.140e+07
## Year              1.301e+14 15      2.954e+00

```

Residuals from first author



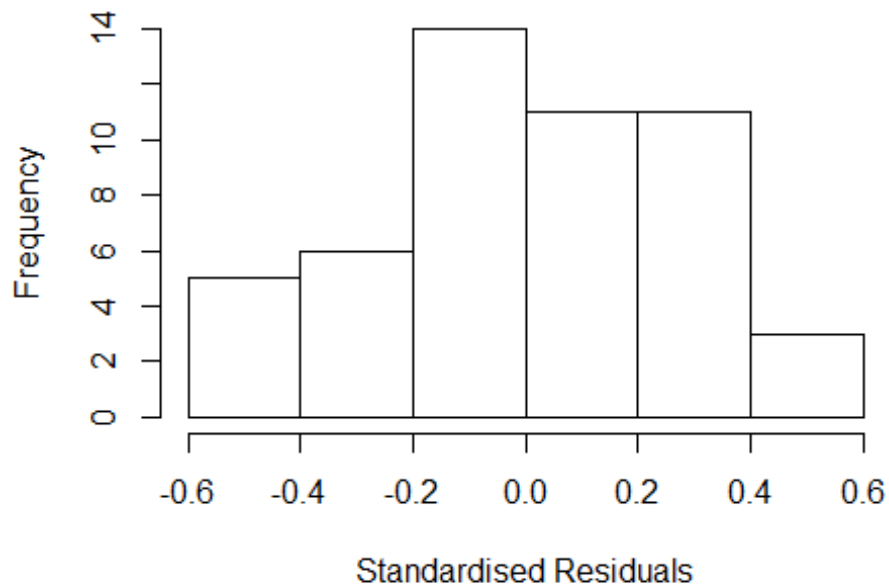
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.7453 -0.2281 0.0628 0.1993 0.5548
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3079 0.3695 3.54 0.0012 **
## FirstAuthorFemale1 0.1101 0.1224 0.90 0.3748
## Year1997 -1.0529 0.3695 -2.85 0.0075 **
## Year1999 -0.1300 0.4417 -0.29 0.7704
## Year2000 0.0109 0.3664 0.03 0.9764
## Year2001 -1.3079 0.3695 -3.54 0.0012 **
## Year2002 -0.5656 0.3905 -1.45 0.1569
## Year2003 -1.1914 0.3789 -3.14 0.0035 **
## Year2004 0.0755 0.6578 0.11 0.9093
## Year2005 0.1431 0.3740 0.38 0.7044
## Year2006 -0.4827 0.5073 -0.95 0.3482
## Year2007 -0.3041 0.4220 -0.72 0.4762
```

```

## Year2008          -0.5626      0.4470    -1.26    0.2170
## Year2009          -0.5708      0.3893    -1.47    0.1520
## Year2010          -0.0991      0.4037    -0.25    0.8076
## Year2011          -0.0672      0.3746    -0.18    0.8588
## Year2012          -0.3642      0.4198    -0.87    0.3919
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.384
## Multiple R-squared:  0.55,   Adjusted R-squared:  0.333
## 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.
##  0.686  0.902  0.954  0.930  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      2.00e-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 3.406e+14 1      1.845e+07
## Year             3.406e+14 15      3.051e+00

```

Residuals from last author



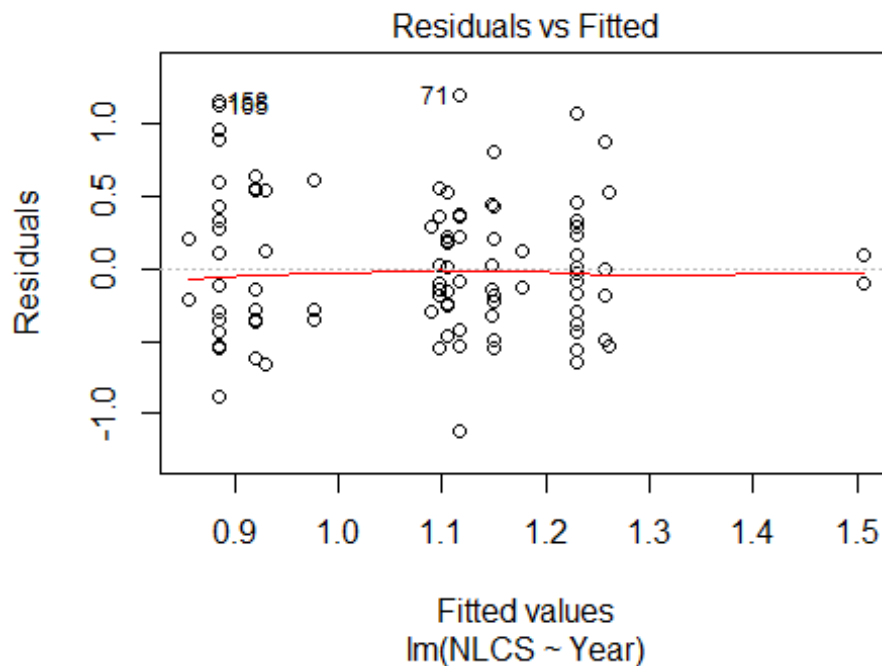
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.61584 -0.17274 0.00933 0.20934 0.55403
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5471 0.1336 11.58 3.7e-13 ***
## LastAuthorFemale1 -0.3682 0.1130 -3.26 0.00259 **
## Year1997 -1.2921 0.1336 -9.67 3.7e-11 ***
## Year1999 -0.1300 0.1318 -0.99 0.33101
## Year2000 0.2501 0.1336 1.87 0.07015 .
## Year2001 -1.5471 0.1336 -11.58 3.7e-13 ***
## Year2002 -0.7740 0.2247 -3.45 0.00157 **
## Year2003 -1.4306 0.1578 -9.07 1.8e-10 ***
## Year2004 -0.1086 0.4747 -0.23 0.82043
## Year2005 0.0483 0.1341 0.36 0.72133
## Year2006 -0.7211 0.3791 -1.90 0.06588 .
## Year2007 -0.2846 0.2435 -1.17 0.25083
```

```

## Year2008          -0.7300      0.2185    -3.34   0.00208 **
## Year2009          -0.6227      0.1699    -3.67   0.00086 ***
## Year2010          -0.0285      0.2056    -0.14   0.89069
## Year2011          -0.1593      0.1639    -0.97   0.33820
## Year2012          -0.5175      0.2751    -1.88   0.06881 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.367
## Multiple R-squared:  0.622, Adjusted R-squared:  0.439
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 42 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.759  0.893  0.940  0.933  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      2.00e-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 3207"
## [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    2    2   10    4    5    5    4    4    4   10    8   11   11   22
## 2012
##   22
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    2    2    2    8    4    5    2    3    2    3    7    8    8    9   16
## 2012
##   20

```

```
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    2    2    2    8    4    4    2    3    1    3    7    7    7    8    14
## 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 = 14, df = 15, p-value = 0.5
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.0099, df = 1, p-value = 0.9

## [1] "Female first author team size 2018 geometric mean: 1.7370729389807"
## [1] "Male first author team size 2018 geometric mean: 1.45421543344895"

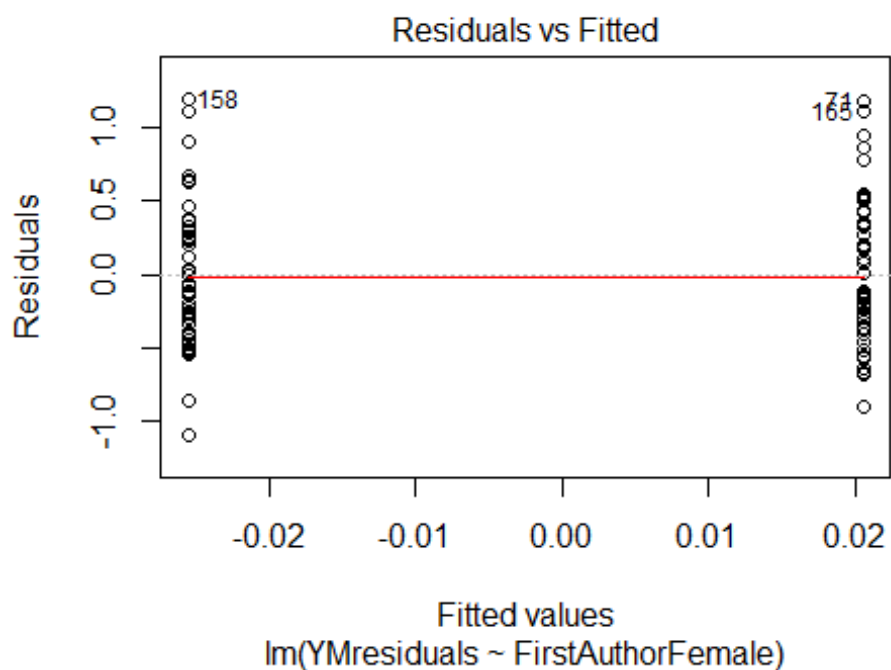
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
```



```
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 43, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.79482292137929"
## [1] "Male last author team size 2018 geometric mean: 1.47273335753469"

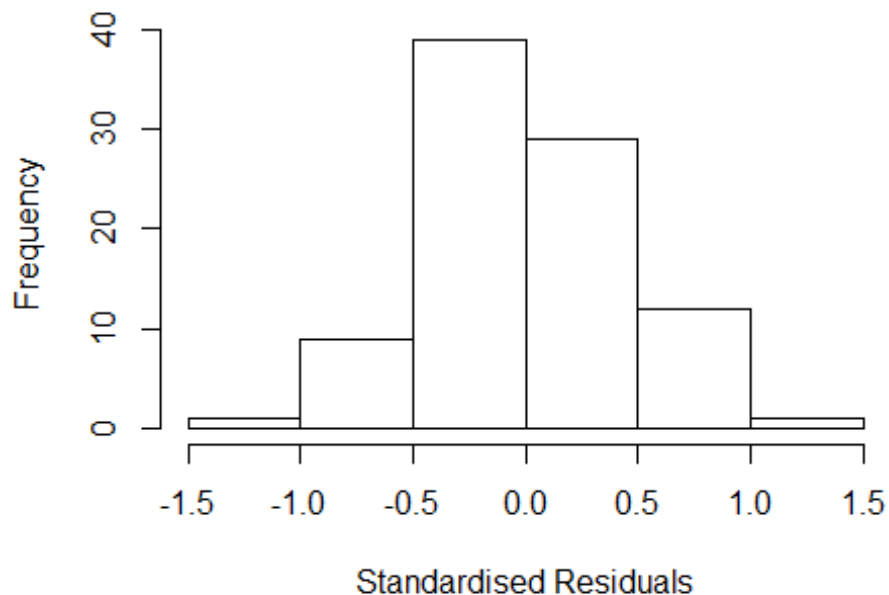
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 37, 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 | 4.475 | 1 | 2.115 |
| LastAuthorFemale | 4.383 | 1 | 2.094 |
| UniqueAuthors | 26311.092 | 4 | 3.569 |
| Year | 62933.753 | 15 | 1.445 |

Residuals from first and last author and team size



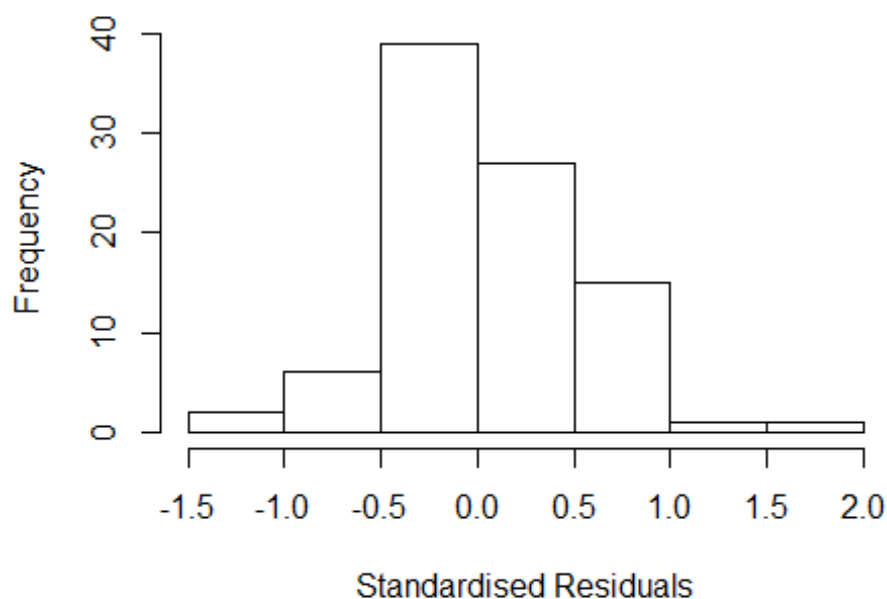
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1398 -0.3200 -0.0287 0.3302 1.2096
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0558 0.0636 16.60 <2e-16 ***
## FirstAuthorFemale1 -0.2305 0.1276 -1.81 0.075 .
## LastAuthorFemale1 0.3315 0.1611 2.06 0.043 *
## UniqueAuthors2 0.2433 0.1269 1.92 0.059 .
## UniqueAuthors3 0.4060 0.2533 1.60 0.113
## UniqueAuthors4 0.1237 0.2625 0.47 0.639
## UniqueAuthors5 -0.1510 0.2466 -0.61 0.542
## Year1998 0.1536 0.5469 0.28 0.780
## Year1999 -0.0949 0.1490 -0.64 0.526
## Year2000 -0.1816 0.1458 -1.25 0.217
```

```

## Year2001          -0.0286      0.1563   -0.18      0.856
## Year2002           0.2493      0.2990    0.83      0.407
## Year2003           0.3285      0.1629    2.02      0.048 *
## Year2004          -0.1070      0.3356   -0.32      0.751
## Year2005          -0.5335      0.2597   -2.05      0.044 *
## Year2006          -0.2500      0.1395   -1.79      0.078 .
## Year2007          -0.0368      0.2234   -0.16      0.869
## Year2008          -0.0172      0.2374   -0.07      0.943
## Year2009          -0.2104      0.2054   -1.02      0.309
## Year2010          -0.1823      0.1776   -1.03      0.308
## Year2011          -0.0173      0.1489   -0.12      0.908
## Year2012          -0.3900      0.1932   -2.02      0.047 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.527
## Multiple R-squared:  0.23,   Adjusted R-squared:  -0.00435
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 83 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.577  0.909  0.959  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           1.10e-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 4.758 1           2.181
## LastAuthorFemale 3.572 1           1.890
## Year              5.740 15           1.060

```

Residuals from first and last author



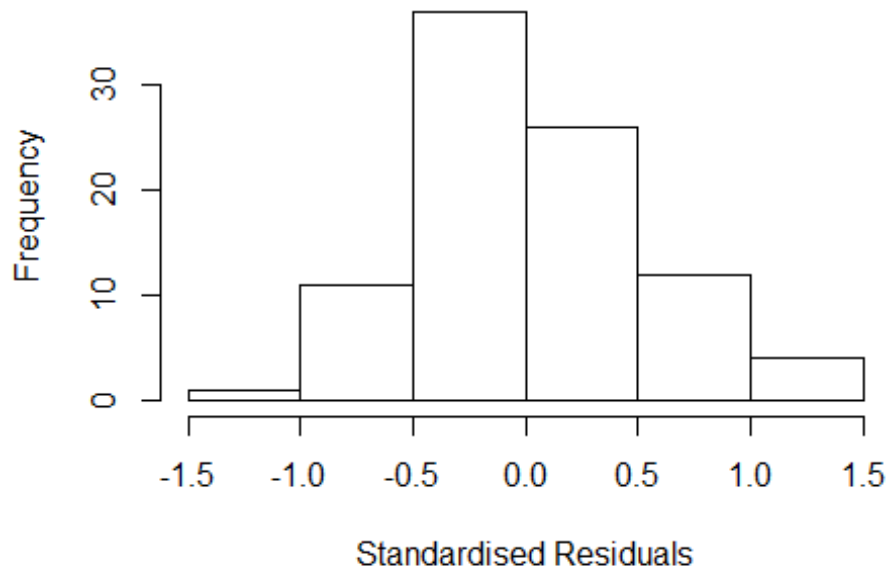
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.22032 -0.29853 -0.00252 0.26965 1.60348
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.17750 0.09202 12.80 < 2e-16 ***
## FirstAuthorFemale1 -0.20301 0.13088 -1.55 0.1252
## LastAuthorFemale1 0.37767 0.13837 2.73 0.0079 **
## Year1998 -0.00483 0.64046 -0.01 0.9940
## Year1999 0.01301 0.17755 0.07 0.9418
## Year2000 -0.08305 0.12955 -0.64 0.5235
## Year2001 -0.06574 0.16563 -0.40 0.6926
## Year2002 0.21812 0.31489 0.69 0.4907
## Year2003 0.32850 0.11683 2.81 0.0063 **
## Year2004 -0.22502 0.35180 -0.64 0.5244
## Year2005 -0.53150 0.09202 -5.78 1.7e-07 ***
## Year2006 -0.27251 0.23934 -1.14 0.2586
```

```

## Year2007          -0.10513    0.22113   -0.48    0.6359
## Year2008          -0.13184    0.26366   -0.50    0.6186
## Year2009          -0.31621    0.21095   -1.50    0.1382
## Year2010          -0.23164    0.21345   -1.09    0.2814
## Year2011           0.00209    0.16636    0.01    0.9900
## Year2012          -0.52697    0.23814   -2.21    0.0300 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.49
## Multiple R-squared:  0.2,   Adjusted R-squared:  0.0131
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 85 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.261  0.897  0.966  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      1.10e-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 3.327 1          1.824
## Year              3.327 15          1.041

```

Residuals from first author

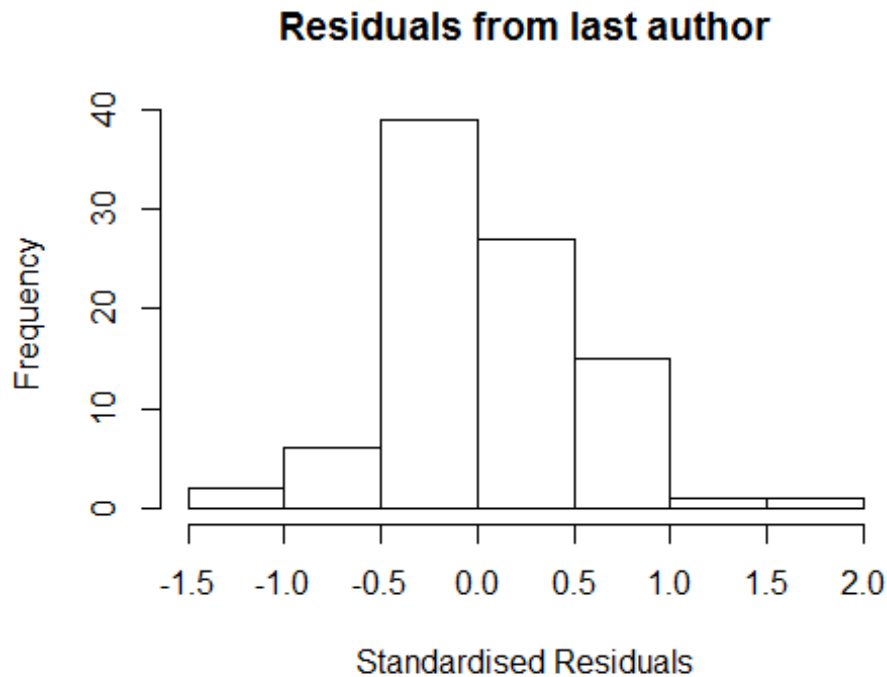


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1930 -0.3251 -0.0467 0.3431 1.2777
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.17750 0.09189 12.81 < 2e-16 ***
## FirstAuthorFemale1 -0.00705 0.12602 -0.06 0.9555
## Year1998 0.08603 0.47389 0.18 0.8564
## Year1999 -0.08497 0.24181 -0.35 0.7263
## Year2000 -0.08143 0.14738 -0.55 0.5823
## Year2001 -0.02807 0.18925 -0.15 0.8825
## Year2002 0.09184 0.29353 0.31 0.7553
## Year2003 0.32850 0.11668 2.82 0.0062 **
## Year2004 -0.22808 0.34529 -0.66 0.5110
## Year2005 -0.53150 0.09189 -5.78 1.6e-07 ***
## Year2006 -0.22253 0.29627 -0.75 0.4550
## Year2007 -0.04389 0.22126 -0.20 0.8433
```

```

## Year2008          0.02252    0.30571    0.07    0.9415
## Year2009          -0.20760    0.23341   -0.89    0.3767
## Year2010          -0.09337    0.17873   -0.52    0.6029
## Year2011           0.07825    0.16703    0.47    0.6408
## Year2012          -0.39716    0.24233   -1.64    0.1055
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.52
## Multiple R-squared:  0.12,   Adjusted R-squared:  -0.0709
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 83 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.525  0.905   0.950   0.918   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      1.10e-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 3.207 1          1.791
## Year              3.207 15          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
## -1.311 -0.282 -0.056 0.285 1.439
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1775 0.0921 12.79 < 2e-16 ***
## LastAuthorFemale1 0.2672 0.1282 2.08 0.0405 *
## Year1998 -0.0511 0.7591 -0.07 0.9465
## Year1999 -0.0885 0.2384 -0.37 0.7116
## Year2000 -0.1178 0.1382 -0.85 0.3969
## Year2001 -0.1628 0.1640 -0.99 0.3241
## Year2002 0.0784 0.2928 0.27 0.7898
## Year2003 0.3285 0.1169 2.81 0.0063 **
## Year2004 -0.2239 0.3543 -0.63 0.5293
## Year2005 -0.5315 0.0921 -5.77 1.7e-07 ***
## Year2006 -0.2991 0.2144 -1.40 0.1672
## Year2007 -0.1365 0.2200 -0.62 0.5369
```

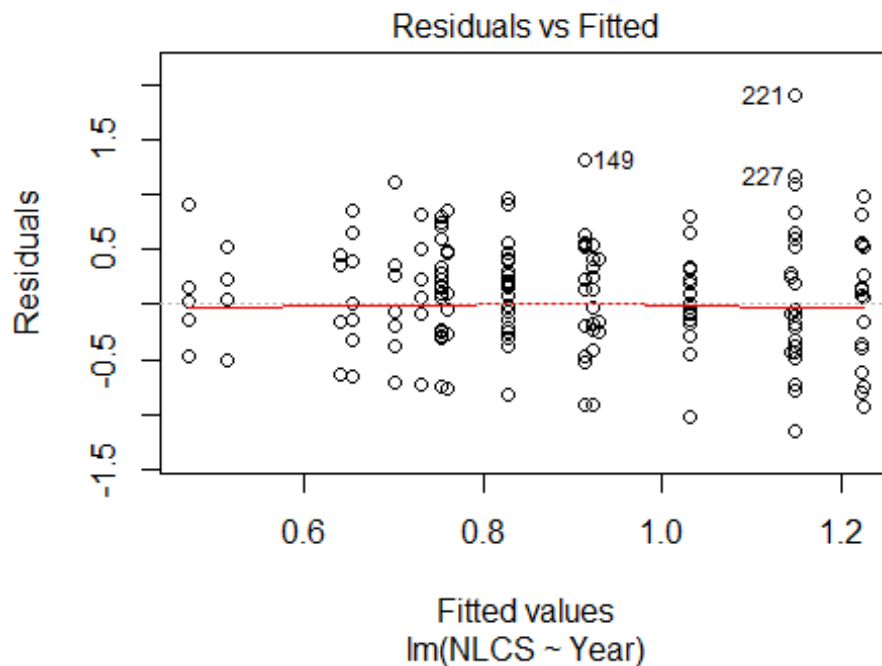


```

## Year2008          -0.1336      0.2824   -0.47    0.6376
## Year2009          -0.3239      0.2193   -1.48    0.1438
## Year2010          -0.3097      0.1969   -1.57    0.1200
## Year2011          -0.0617      0.1491   -0.41    0.6804
## Year2012          -0.5653      0.2490   -2.27    0.0261 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.48
## Multiple R-squared:  0.18,   Adjusted R-squared:  0.00284
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 87 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.349  0.894  0.967  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      1.10e-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 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
##    5    7    8    5    7    5   14   13    9   12   11   22   14   14   30
## 2011 2012
##   27   25
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    6    6    4    7    3    9   11    8    9    8   19   13   13   27
## 2011 2012
##   23   22

```

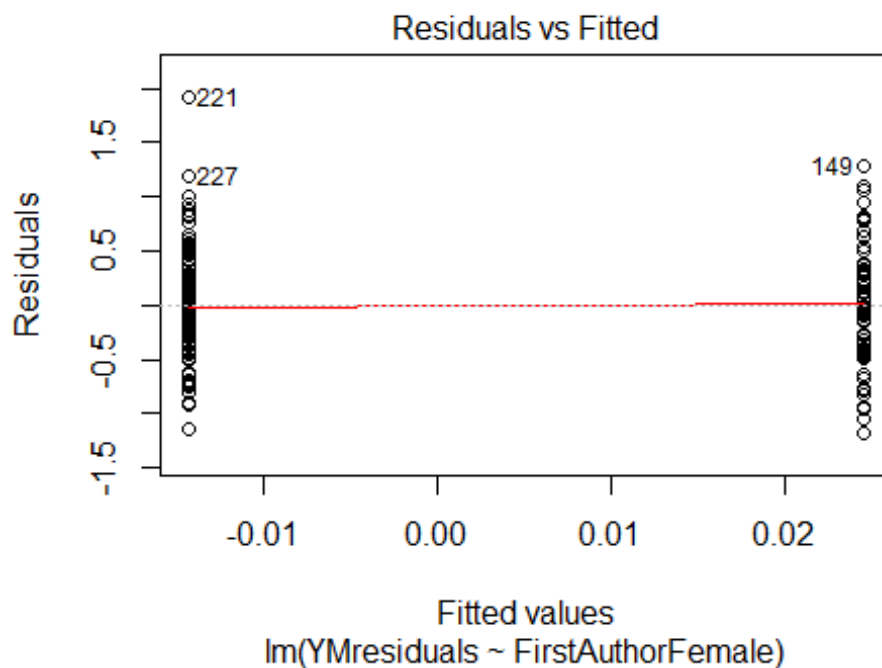
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    6    6    4    7    3    8   10    8    9    8   17   13   12   25
## 2011 2012
##   22   21
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 17, df = 16, p-value = 0.4
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.6, df = 1, p-value = 0.2
## [1] "Female first author team size 2018 geometric mean: 1.62450479271247"
## [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 = 52, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.29683955465101"
## [1] "Male last author team size 2018 geometric mean: 1.55440628177092"

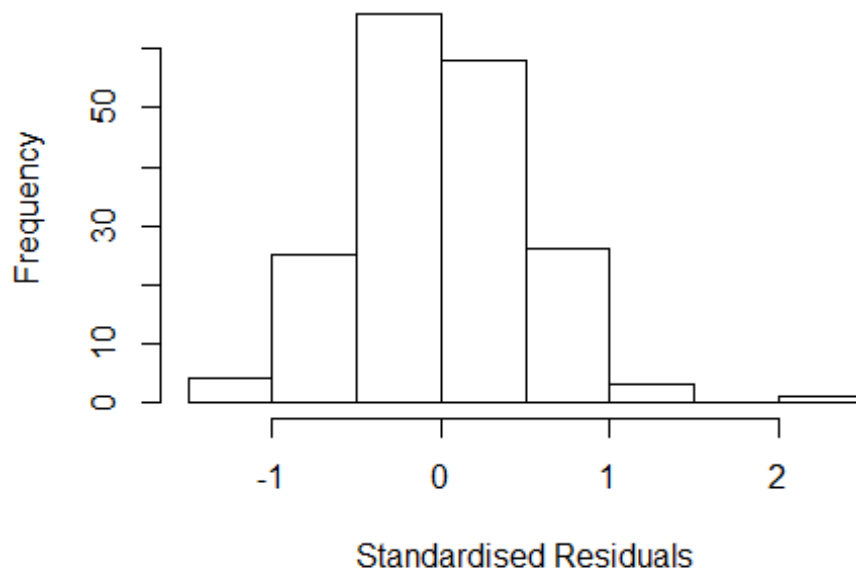
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## 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] "Regression 1: First author gender, last author gender, team size,
## Year as factors"
##
```

| | GVIF | Df | GVIF ^{1/(2*Df)} |
|-------------------|-----------|----|--------------------------|
| FirstAuthorFemale | 4.472e+14 | 1 | 2.115e+07 |
| LastAuthorFemale | 1.161e+01 | 1 | 3.407e+00 |
| UniqueAuthors | 3.763e+16 | 4 | 1.180e+02 |
| Year | 1.188e+16 | 16 | 3.179e+00 |

Residuals from first and last author and team size



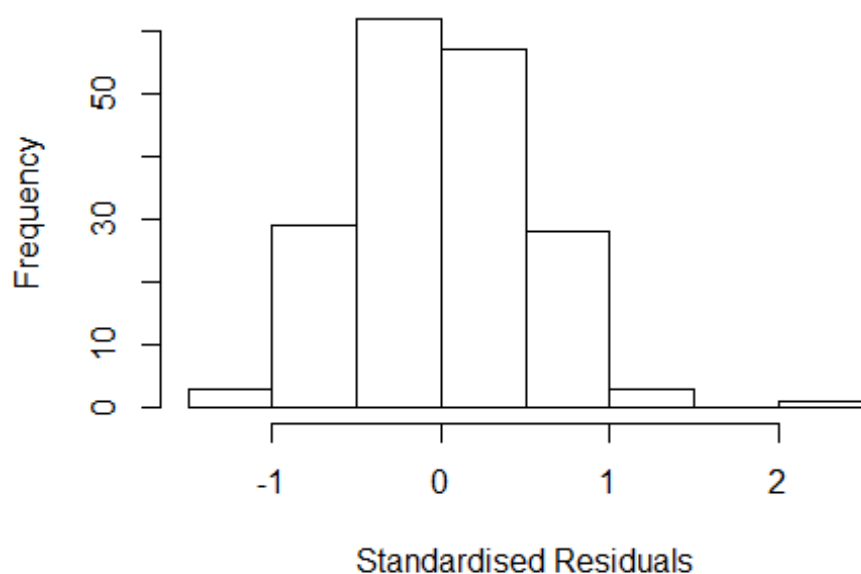
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.06943 -0.33823 -0.00642 0.31475 2.05958
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1081 0.1644 6.74 2.7e-10 ***
## FirstAuthorFemale1 0.2823 0.2421 1.17 0.2453
## LastAuthorFemale1 -0.2003 0.2343 -0.85 0.3940
## UniqueAuthors2 -0.0810 0.1495 -0.54 0.5887
## UniqueAuthors3 0.1243 0.2556 0.49 0.6274
## UniqueAuthors4 0.0818 0.2000 0.41 0.6830
## UniqueAuthors5 -0.7590 0.2759 -2.75 0.0066 **
## Year1997 -0.7024 0.2563 -2.74 0.0068 **
## Year1998 -0.6309 0.2523 -2.50 0.0134 *
## Year1999 -0.4779 0.2794 -1.71 0.0891 .
```

```

## Year2000          0.1256      0.3191      0.39      0.6943
## Year2001         -0.1545      0.2449     -0.63      0.5291
## Year2002         -0.4590      0.2343     -1.96      0.0519 .
## Year2003         -0.1284      0.2110     -0.61      0.5438
## Year2004         -0.4627      0.2459     -1.88      0.0617 .
## Year2005         -0.5219      0.2477     -2.11      0.0367 *
## Year2006         -0.3507      0.2382     -1.47      0.1428
## Year2007         -0.1470      0.1845     -0.80      0.4269
## Year2008         -0.2065      0.2991     -0.69      0.4911
## Year2009          0.0562      0.2313      0.24      0.8085
## Year2010         -0.3290      0.1959     -1.68      0.0949 .
## Year2011         -0.1207      0.2528     -0.48      0.6338
## Year2012         -0.4098      0.1939     -2.11      0.0361 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.528
## Multiple R-squared:  0.154, Adjusted R-squared:  0.0373
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 164 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0944 0.8670 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          5.46e-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 7.408 1          2.722
## LastAuthorFemale  7.420 1          2.724
## Year              1.386 16          1.010

```

Residuals from first and last author



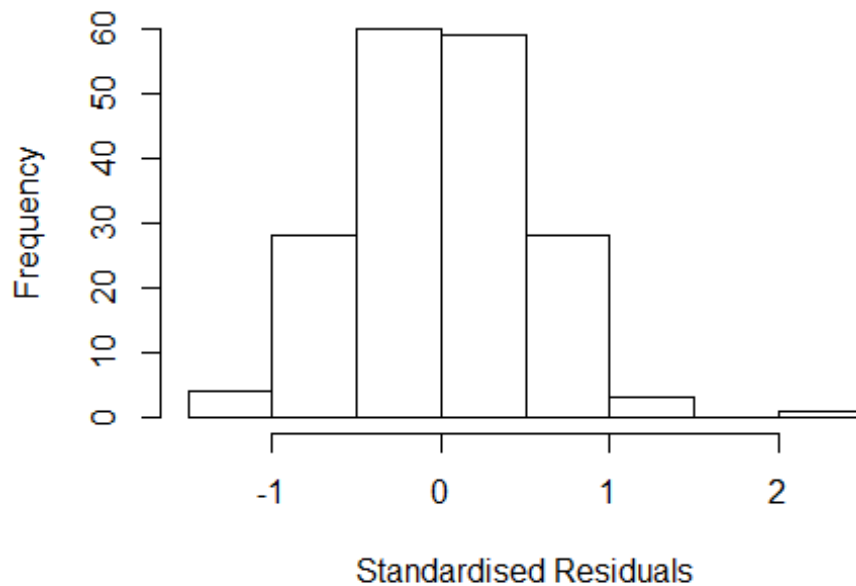
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.04857 -0.33125 -0.00857 0.32165 2.07539
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.110 0.163 6.80 1.8e-10 ***
## FirstAuthorFemale1 0.234 0.222 1.05 0.2938
## LastAuthorFemale1 -0.157 0.221 -0.71 0.4784
## Year1997 -0.699 0.255 -2.74 0.0068 **
## Year1998 -0.616 0.224 -2.75 0.0067 **
## Year1999 -0.479 0.278 -1.73 0.0864 .
## Year2000 0.107 0.312 0.34 0.7324
## Year2001 -0.185 0.239 -0.77 0.4408
## Year2002 -0.461 0.233 -1.98 0.0495 *
## Year2003 -0.205 0.218 -0.94 0.3479
## Year2004 -0.460 0.245 -1.88 0.0620 .
## Year2005 -0.519 0.245 -2.12 0.0358 *
```

```

## Year2006          -0.378      0.237   -1.60    0.1125
## Year2007          -0.149      0.184   -0.81    0.4208
## Year2008          -0.216      0.287   -0.75    0.4531
## Year2009           0.047      0.229    0.20    0.8378
## Year2010          -0.318      0.185   -1.72    0.0870 .
## Year2011          -0.138      0.239   -0.58    0.5631
## Year2012          -0.412      0.194   -2.13    0.0350 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.551
## Multiple R-squared:  0.137, Adjusted R-squared:  0.0427
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 169 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.125  0.877   0.961   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      5.46e-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 1.278 1      1.131
## Year              1.278 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
## -1.05851 -0.32835 -0.00783 0.31779 2.09031
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1004 0.1653 6.66 3.9e-10 ***
## FirstAuthorFemale1 0.0968 0.0923 1.05 0.2955
## Year1997 -0.6962 0.2579 -2.70 0.0077 **
## Year1998 -0.6133 0.2239 -2.74 0.0068 **
## Year1999 -0.4754 0.2769 -1.72 0.0879 .
## Year2000 0.1130 0.3111 0.36 0.7168
## Year2001 -0.1751 0.2405 -0.73 0.4676
## Year2002 -0.4541 0.2343 -1.94 0.0543 .
## Year2003 -0.1895 0.2145 -0.88 0.3784
## Year2004 -0.4535 0.2443 -1.86 0.0652 .
## Year2005 -0.5223 0.2465 -2.12 0.0356 *
## Year2006 -0.3889 0.2457 -1.58 0.1154
```

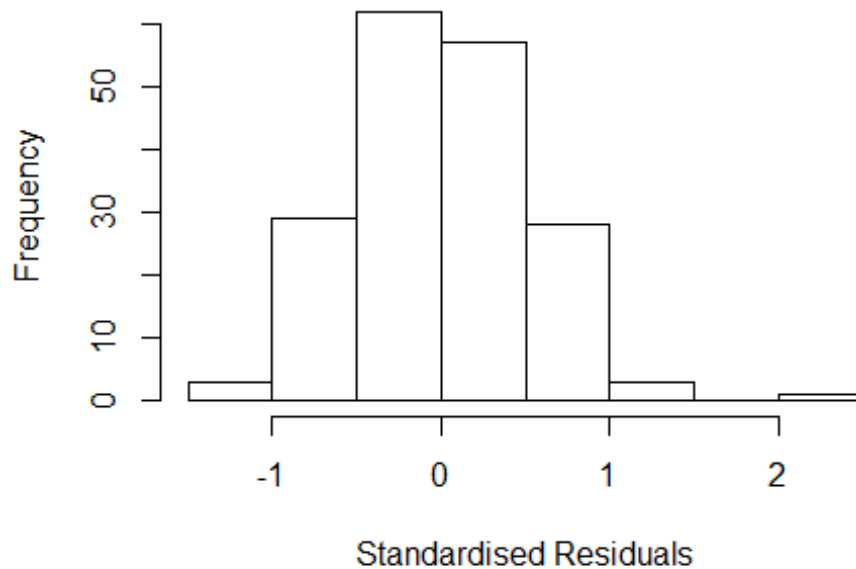


```

## Year2007          -0.1387      0.1863   -0.74   0.4576
## Year2008          -0.2248      0.2856   -0.79   0.4324
## Year2009           0.0535      0.2273    0.24   0.8142
## Year2010          -0.3153      0.1885   -1.67   0.0962 .
## Year2011          -0.1437      0.2380   -0.60   0.5469
## Year2012          -0.4086      0.1961   -2.08   0.0388 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.554
## Multiple R-squared:  0.132, Adjusted R-squared:  0.043
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 169 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.124  0.880  0.963  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      5.46e-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 1.284 1          1.133
## Year            1.284 16          1.008

```

Residuals from last author



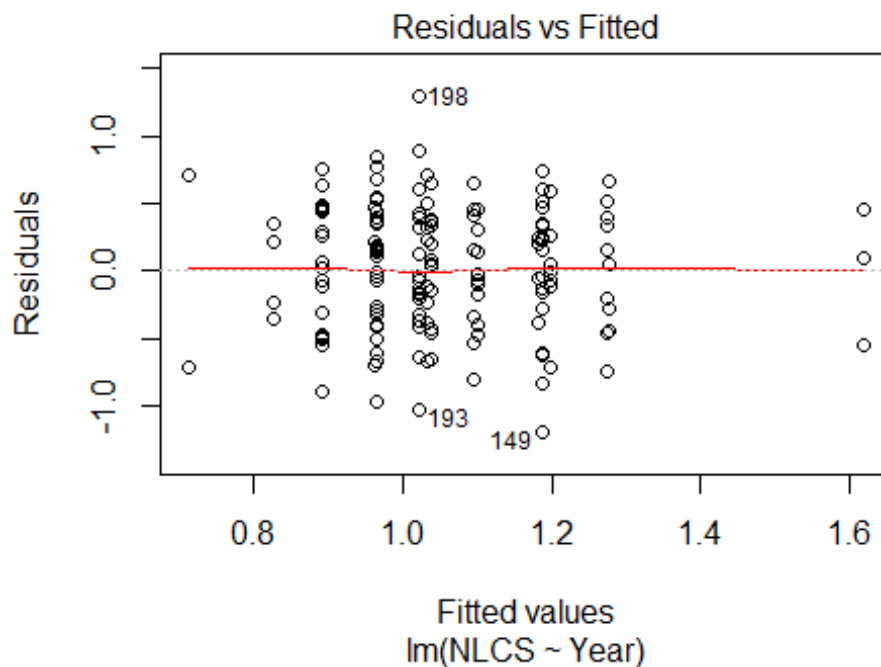
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0342 -0.3350 0.0129 0.3264 2.0650
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1252 0.1600 7.03 5.1e-11 ***
## LastAuthorFemale1 0.0451 0.0940 0.48 0.6318
## Year1997 -0.7018 0.2510 -2.80 0.0058 **
## Year1998 -0.6207 0.2254 -2.75 0.0066 **
## Year1999 -0.4861 0.2794 -1.74 0.0837 .
## Year2000 0.0970 0.3128 0.31 0.7569
## Year2001 -0.1998 0.2368 -0.84 0.3999
## Year2002 -0.4711 0.2305 -2.04 0.0425 *
## Year2003 -0.1920 0.2077 -0.92 0.3567
## Year2004 -0.4702 0.2457 -1.91 0.0574 .
## Year2005 -0.5136 0.2433 -2.11 0.0363 *
## Year2006 -0.4055 0.2497 -1.62 0.1063
```

```

## Year2007          -0.1361      0.1825    -0.75    0.4568
## Year2008          -0.2408      0.2781    -0.87    0.3878
## Year2009           0.0648      0.2199     0.29    0.7685
## Year2010          -0.3251      0.1865    -1.74    0.0832 .
## Year2011          -0.1432      0.2346    -0.61    0.5425
## Year2012          -0.4181      0.1903    -2.20    0.0294 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.555
## Multiple R-squared:  0.128, Adjusted R-squared:  0.0376
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 169 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.137  0.877  0.959  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      5.46e-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: 183"
## [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
##    3    5    9    5   10    5    9    7    4    3   11   12   17   23   23
## 2011 2012
##   24   26
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    4    7    4    9    4    7    7    3    3   10    9   13   22   19
## 2011 2012

```

```
## 23 21
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2 4 5 4 9 4 6 6 2 1 10 8 13 21 17
## 2011 2012
## 22 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 = 9.6, df = 16, p-value = 0.9
```



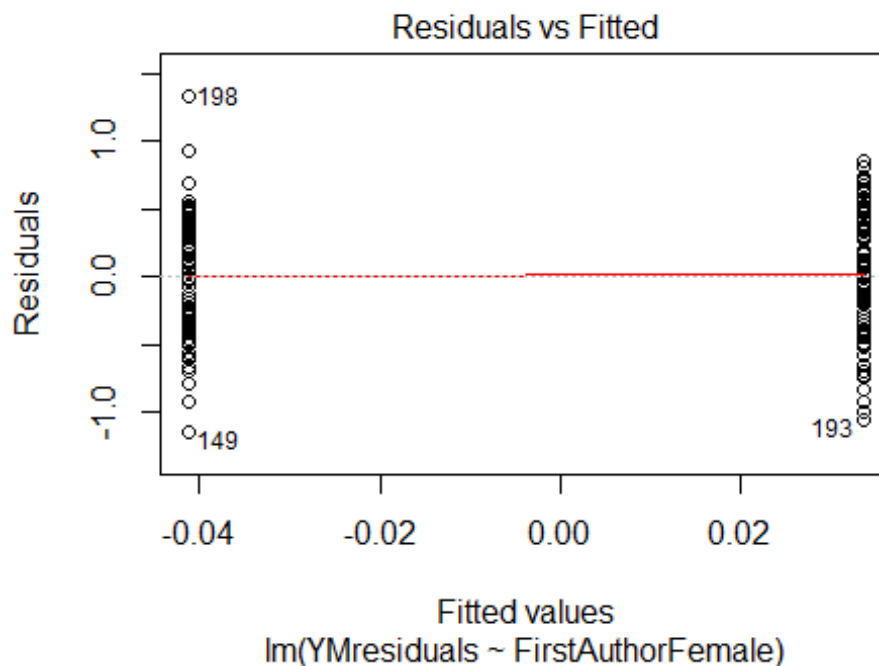
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.29, df = 1, p-value = 0.6

## [1] "Female first author team size 2018 geometric mean: 2.19176545540841"
## [1] "Male first author team size 2018 geometric mean: 1.45421543344895"

## Warning in wilcox.test.default(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 = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.15331731991423"
## [1] "Male last author team size 2018 geometric mean: 1.73552611877147"

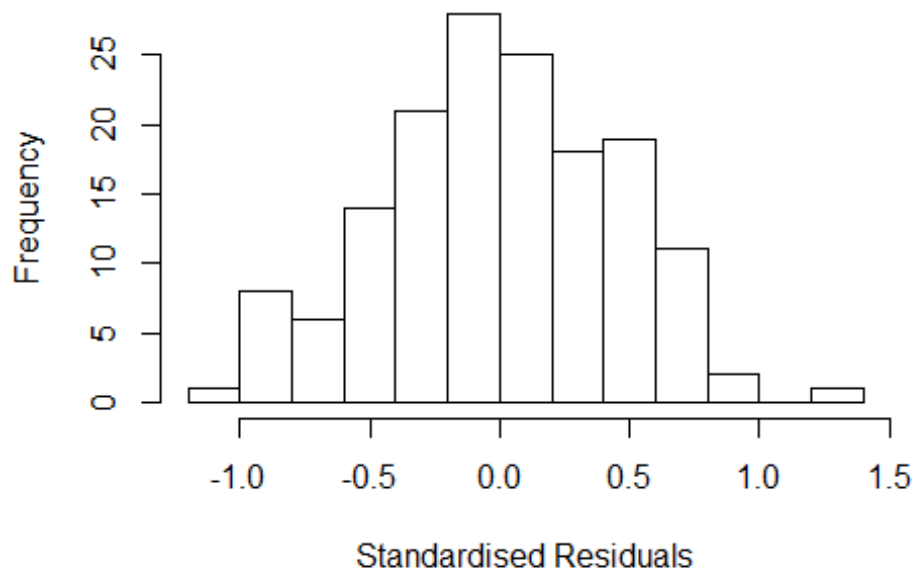
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110, 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 | 17.02 | 1 | 4.125 |
| LastAuthorFemale | 19.24 | 1 | 4.387 |
| UniqueAuthors | 59.17 | 3 | 1.974 |
| Year | 134.64 | 16 | 1.166 |

Residuals from first and last author and team size



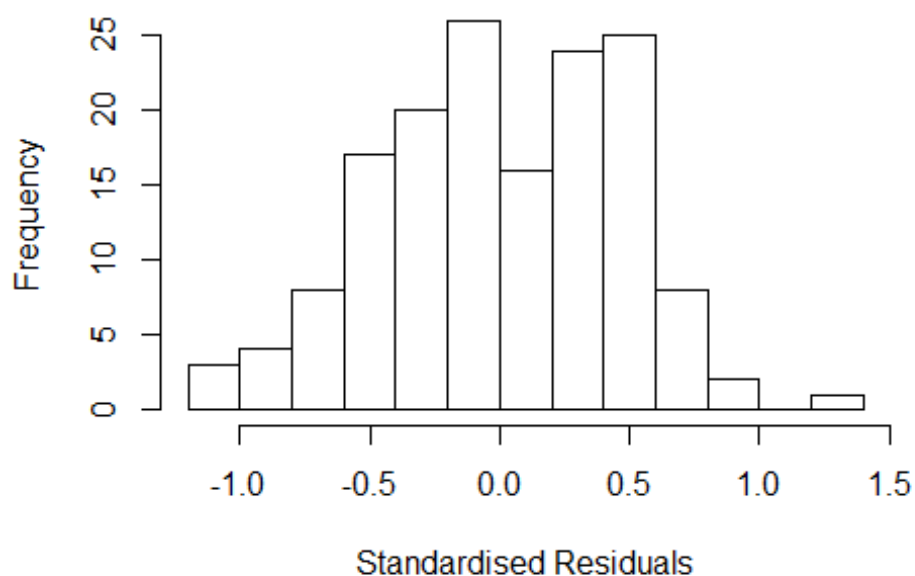
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.01891 -0.33356 -0.00203 0.35061 1.27840
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7120 0.8385 0.85 0.397
## FirstAuthorFemale1 -0.0505 0.1030 -0.49 0.625
## LastAuthorFemale1 -0.0832 0.1124 -0.74 0.461
## UniqueAuthors2 0.1872 0.0986 1.90 0.060 .
## UniqueAuthors3 0.3876 0.2016 1.92 0.057 .
## UniqueAuthors4 -0.1498 0.1231 -1.22 0.226
## Year1997 0.0971 0.8641 0.11 0.911
## Year1998 0.4622 0.8599 0.54 0.592
## Year1999 0.3229 0.8528 0.38 0.706
## Year2000 0.2009 0.8615 0.23 0.816
```

```

## Year2001      0.5537      0.8743      0.63      0.528
## Year2002      0.2501      0.8714      0.29      0.775
## Year2003      0.5938      0.8594      0.69      0.491
## Year2004      0.0788      0.9618      0.08      0.935
## Year2005      1.3610      0.8385      1.62      0.107
## Year2006      0.3568      0.8416      0.42      0.672
## Year2007      0.3452      0.8535      0.40      0.687
## Year2008      0.3432      0.8453      0.41      0.685
## Year2009      0.2799      0.8466      0.33      0.741
## Year2010      0.4406      0.8477      0.52      0.604
## Year2011      0.2691      0.8488      0.32      0.752
## Year2012      0.1985      0.8520      0.23      0.816
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.504
## Multiple R-squared:  0.13,    Adjusted R-squared:  -0.00785
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 138 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.500  0.894  0.952  0.923  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      6.49e-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 4.209 1      2.052
## LastAuthorFemale 3.921 1      1.980
## Year              4.541 16      1.048

```

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
## -1.1150 -0.3556 -0.0101  0.3648  1.3777
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.7120     0.7822   0.91   0.364
## FirstAuthorFemale1 0.0554     0.1077   0.51   0.608
## LastAuthorFemale1 -0.1292     0.1212  -1.07   0.288
## Year1997          0.1314     0.8000   0.16   0.870
## Year1998          0.5264     0.8047   0.65   0.514
## Year1999          0.4310     0.7972   0.54   0.590
## Year2000          0.2750     0.8029   0.34   0.732
## Year2001          0.5553     0.8169   0.68   0.498
## Year2002          0.2705     0.8167   0.33   0.741
## Year2003          0.5526     0.8068   0.68   0.495
## Year2004          0.0489     0.8884   0.06   0.956
## Year2005          1.3610     0.7822   1.74   0.084 .
```

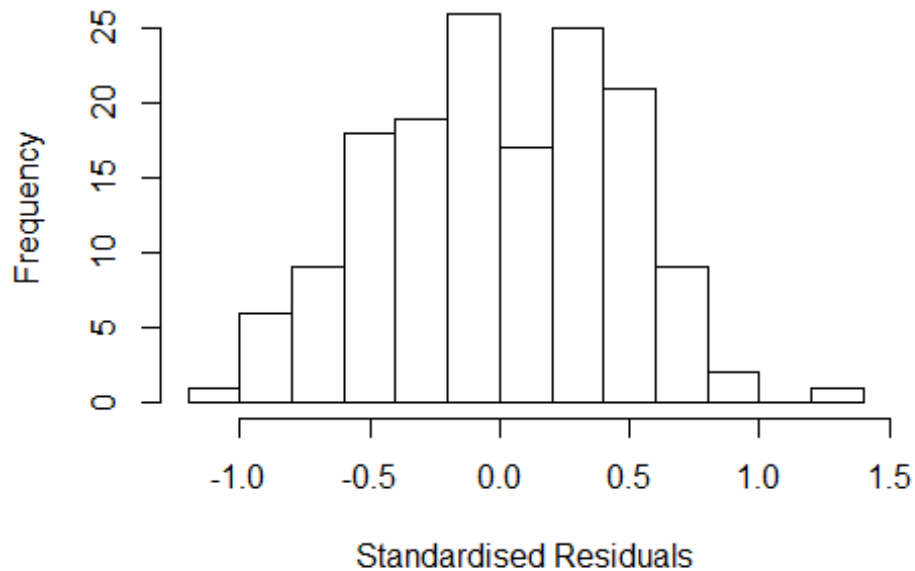


```

## Year2006          0.3852      0.7869      0.49      0.625
## Year2007          0.3817      0.7996      0.48      0.634
## Year2008          0.3399      0.7899      0.43      0.668
## Year2009          0.2872      0.7903      0.36      0.717
## Year2010          0.4768      0.7907      0.60      0.548
## Year2011          0.2971      0.7944      0.37      0.709
## Year2012          0.2385      0.7969      0.30      0.765
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.531
## Multiple R-squared:  0.0891, Adjusted R-squared:  -0.0323
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 142 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.481  0.907  0.952  0.929  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      6.49e-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 2.641 1      1.625
## Year              2.641 16      1.031

```

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
## -1.131 -0.355 -0.020 0.371 1.374
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7120 0.7739 0.92 0.359
## FirstAuthorFemale1 -0.0461 0.0890 -0.52 0.605
## Year1997 0.1249 0.7909 0.16 0.875
## Year1998 0.5149 0.7989 0.64 0.520
## Year1999 0.5066 0.7859 0.64 0.520
## Year2000 0.2715 0.7952 0.34 0.733
## Year2001 0.5722 0.8056 0.71 0.479
## Year2002 0.3053 0.8067 0.38 0.706
## Year2003 0.5337 0.7996 0.67 0.506
## Year2004 0.0351 0.8729 0.04 0.968
## Year2005 1.3610 0.7739 1.76 0.081 .
## Year2006 0.4015 0.7799 0.51 0.608
```

```

## Year2007          0.3777      0.7917      0.48      0.634
## Year2008          0.3488      0.7821      0.45      0.656
## Year2009          0.2765      0.7820      0.35      0.724
## Year2010          0.4650      0.7826      0.59      0.553
## Year2011          0.2735      0.7849      0.35      0.728
## Year2012          0.2208      0.7877      0.28      0.780
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.536
## Multiple R-squared:  0.0854, Adjusted R-squared:  -0.0289
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 140 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.491  0.906  0.953  0.929  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      6.49e-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 2.457 1      1.567
## Year      2.457 16      1.028

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.1117 -0.3616 -0.0118  0.3583  1.3835

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.7120     0.7903   0.90   0.369
## LastAuthorFemale1 -0.0802     0.0972  -0.82   0.411
## Year1997          0.1329     0.8081   0.16   0.870
## Year1998          0.5286     0.8131   0.65   0.517
## Year1999          0.4724     0.8005   0.59   0.556
## Year2000          0.2760     0.8107   0.34   0.734
## Year2001          0.5700     0.8217   0.69   0.489
## Year2002          0.2936     0.8187   0.36   0.720
## Year2003          0.5570     0.8142   0.68   0.495
## Year2004          0.0521     0.8979   0.06   0.954
## Year2005          1.3610     0.7903   1.72   0.087 .
## Year2006          0.3963     0.7954   0.50   0.619
## Year2007          0.3827     0.8075   0.47   0.636
## Year2008          0.3509     0.7973   0.44   0.661
## Year2009          0.2897     0.7983   0.36   0.717
## Year2010          0.4798     0.7987   0.60   0.549
## Year2011          0.2976     0.8024   0.37   0.711
## Year2012          0.2429     0.8047   0.30   0.763
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.527
## Multiple R-squared:  0.0887, Adjusted R-squared:  -0.0253
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 141 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.470  0.907   0.952   0.927   0.985   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      6.49e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##   nResample    max.it    best.r.s    k.fast.s    k.max maxit.scale
##      500         50         2         1        1000         200
##   trace.lev    mts    compute.rd
##      0        1000         0
##           psi          subsampling          cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 154"
## [1] ""

```

```

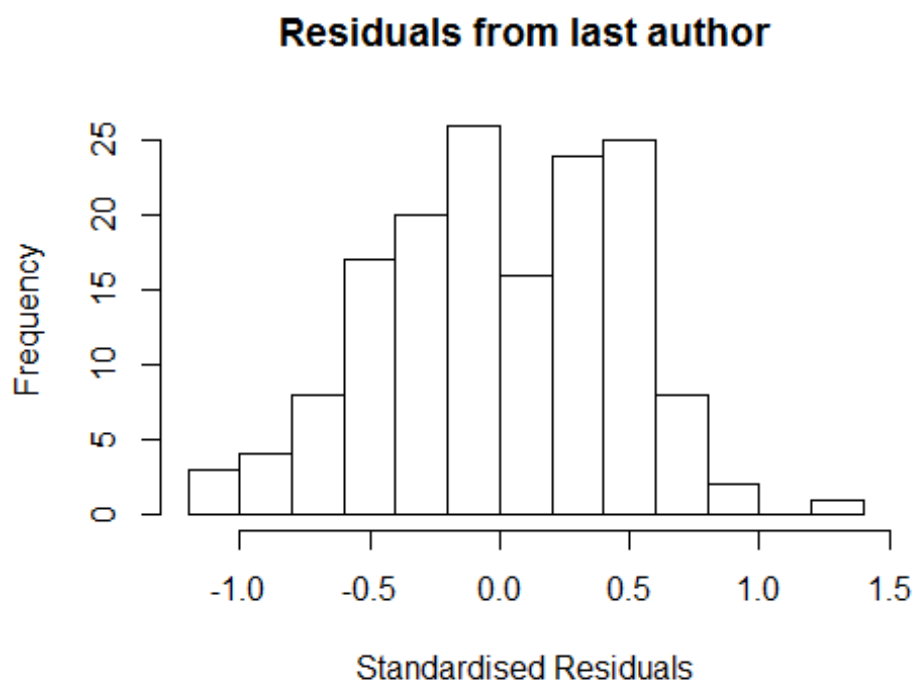
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 3302"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    3    2    3    4    2    4    8    4    3    9   10   17    8    9
## 2011 2012
##   23    9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    2    2    2    4    0    3    5    1    2    8    9   11    6    9
## 2011 2012
##   18    7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    2    2    2    4    0    3    5    1    2    8    9   11    6    9
## 2011 2012
##   17    7
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.09050773266526"
## [1] "Male first author team size 2018 geometric mean: 1.09050773266526"

## Warning in wilcox.test.default(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 = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.09050773266526"
## [1] "Male last author team size 2018 geometric mean: 1.09050773266526"

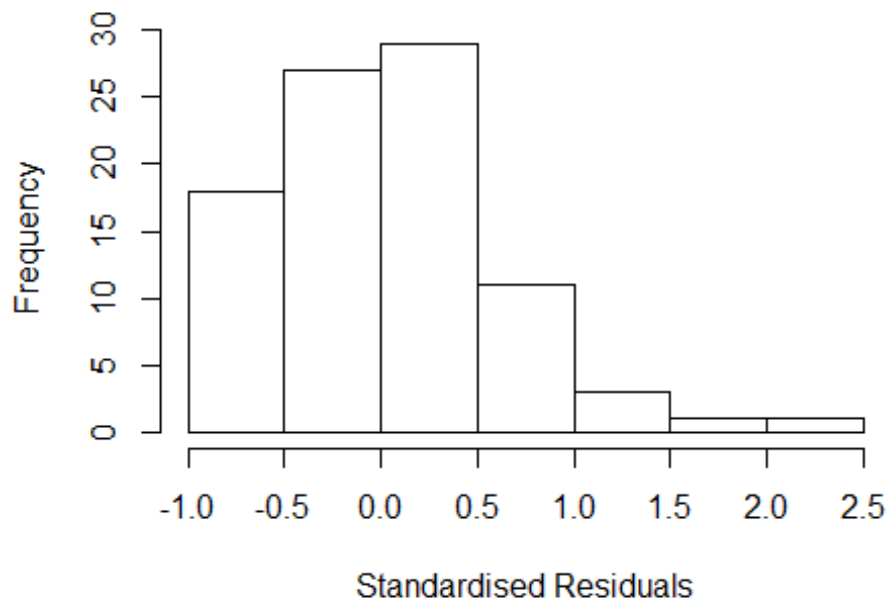
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 32, p-value = 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 7.482e+14  1      2.735e+07
## LastAuthorFemale  5.917e+14  1      2.432e+07
## UniqueAuthors    1.434e+17  3      7.235e+02
## Year              8.074e+17 15      3.953e+00
```

Residuals from first and last author and team size



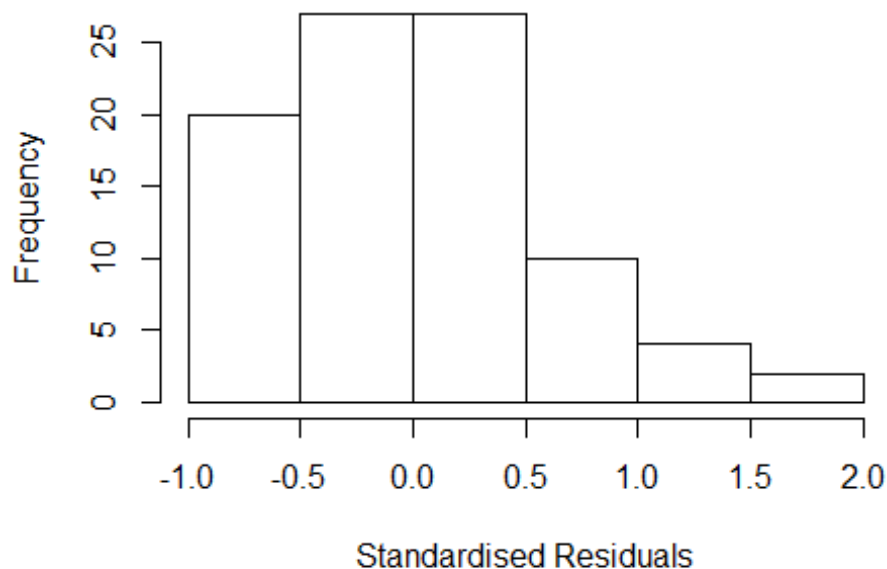
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.95967 -0.41943 0.00515 0.37255 2.06330
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2260 0.1430 8.57 1.8e-12 ***
## FirstAuthorFemale1 0.4939 0.2322 2.13 0.03696 *
## LastAuthorFemale1 -0.3911 0.2065 -1.89 0.06240 .
## UniqueAuthors2 0.3800 0.2636 1.44 0.15402
## UniqueAuthors3 0.6764 0.4590 1.47 0.14508
## UniqueAuthors4 -0.0650 0.2765 -0.23 0.81491
## Year1997 -0.3162 0.3090 -1.02 0.30977
## Year1998 0.2240 0.1526 1.47 0.14660
## Year1999 0.2471 0.3573 0.69 0.49165
## Year2000 -0.4173 0.3553 -1.17 0.24429
```

```

## Year2002          -0.5385      0.3317    -1.62   0.10907
## Year2003          -0.3099      0.3522    -0.88   0.38199
## Year2004           0.7850      0.1430     5.49   6.3e-07 ***
## Year2005          -0.0914      0.2192    -0.42   0.67785
## Year2006          -0.1866      0.3438    -0.54   0.58918
## Year2007           0.0122      0.2622     0.05   0.96311
## Year2008          -0.8066      0.2197    -3.67   0.00047 ***
## Year2009           0.0167      0.2979     0.06   0.95546
## Year2010          -0.7594      0.1949    -3.90   0.00022 ***
## Year2011          -0.6463      0.2455    -2.63   0.01045 *
## Year2012          -0.6779      0.2752    -2.46   0.01626 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.586
## Multiple R-squared:  0.324, Adjusted R-squared:  0.128
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 76 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.190  0.880  0.952  0.908  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      1.11e-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 15.796 1      3.974
## LastAuthorFemale 13.887 1      3.726
## Year              3.478 15      1.042

```


Residuals from first and last author



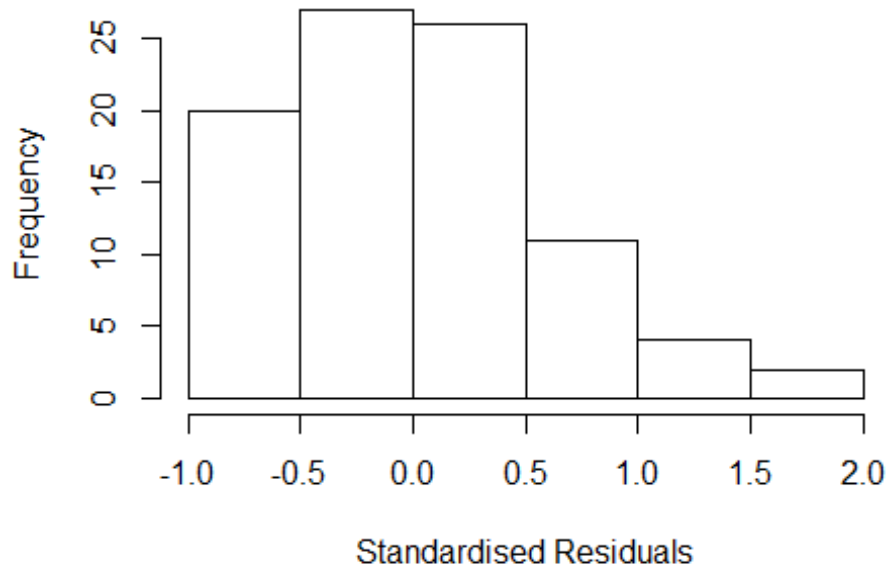
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9188 -0.4298 -0.0181 0.3659 1.9410
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2260 0.1427 8.59 1.2e-12 ***
## FirstAuthorFemale1 0.2447 0.3728 0.66 0.51360
## LastAuthorFemale1 -0.1663 0.3453 -0.48 0.63154
## Year1997 0.0220 0.4752 0.05 0.96320
## Year1998 0.2240 0.1523 1.47 0.14577
## Year1999 0.2593 0.3646 0.71 0.47927
## Year2000 -0.4126 0.3546 -1.16 0.24840
## Year2002 -0.5406 0.3293 -1.64 0.10497
## Year2003 -0.3072 0.3438 -0.89 0.37456
## Year2004 0.7850 0.1427 5.50 5.5e-07 ***
## Year2005 0.1108 0.2230 0.50 0.62087
## Year2006 -0.0976 0.2882 -0.34 0.73575
```

```

## Year2007          0.0355      0.2776      0.13  0.89863
## Year2008         -0.7962      0.2179     -3.65  0.00049 ***
## Year2009          0.0298      0.2934      0.10  0.91932
## Year2010         -0.6368      0.1885     -3.38  0.00118 **
## Year2011         -0.5240      0.2552     -2.05  0.04370 *
## Year2012         -0.6741      0.2468     -2.73  0.00794 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.612
## Multiple R-squared:  0.276, Adjusted R-squared:  0.105
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 82 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.294  0.890  0.956  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      1.11e-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 2.54 1      1.594
## Year              2.54 15      1.032

```

Residuals from first author



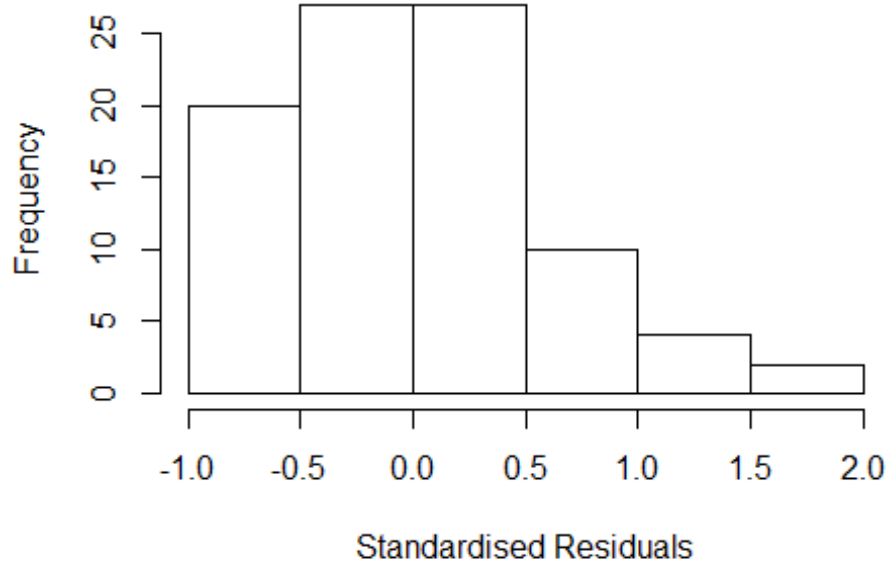
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9160 -0.4263 -0.0246 0.3693 1.9635
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2260 0.1427 8.59 1.1e-12 ***
## FirstAuthorFemale1 0.0907 0.1571 0.58 0.56548
## Year1997 0.0220 0.4746 0.05 0.96315
## Year1998 0.2240 0.1523 1.47 0.14566
## Year1999 0.2531 0.3567 0.71 0.48020
## Year2000 -0.4160 0.3525 -1.18 0.24173
## Year2002 -0.5408 0.3291 -1.64 0.10460
## Year2003 -0.3100 0.3453 -0.90 0.37220
## Year2004 0.7850 0.1427 5.50 5.3e-07 ***
## Year2005 0.1046 0.2265 0.46 0.64542
## Year2006 -0.1264 0.2692 -0.47 0.63999
## Year2007 0.0310 0.2779 0.11 0.91155
```

```

## Year2008          -0.7997      0.2181    -3.67   0.00046 ***
## Year2009           0.0225      0.2938      0.08   0.93927
## Year2010          -0.6260      0.1910    -3.28   0.00161 **
## Year2011          -0.5465      0.2635    -2.07   0.04161 *
## Year2012          -0.6803      0.2472    -2.75   0.00747 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.615
## Multiple R-squared:  0.274, Adjusted R-squared:  0.114
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 82 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.286  0.892   0.957   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      1.11e-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"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 2.461 1          1.569
## Year             2.461 15          1.030

```

Residuals from last author



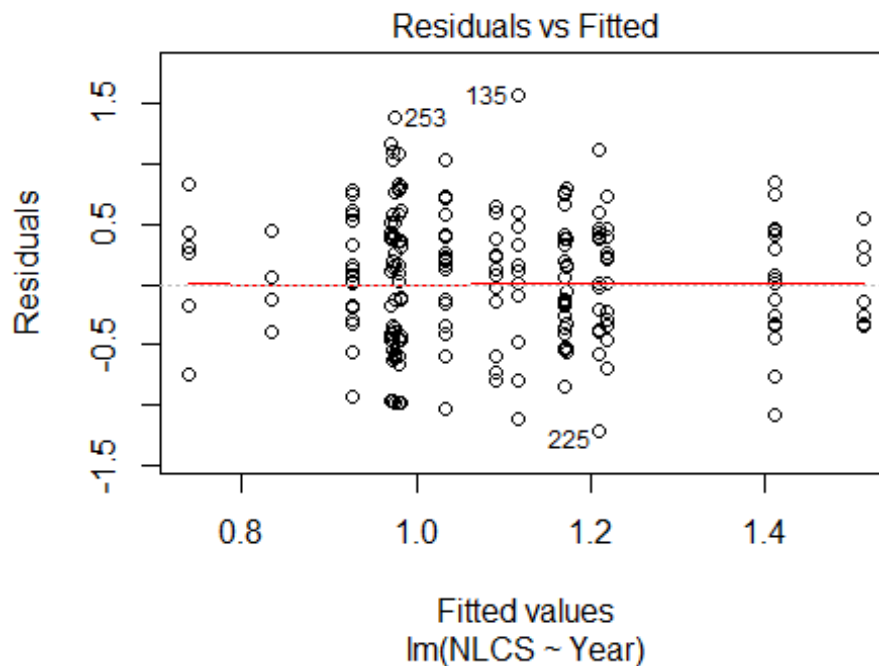
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.92554 -0.43806 -0.00807 0.35722 1.96075
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2260 0.1427 8.59 1.1e-12 ***
## LastAuthorFemale1 0.0477 0.1445 0.33 0.74214
## Year1997 0.0220 0.4752 0.05 0.96320
## Year1998 0.2240 0.1523 1.47 0.14571
## Year1999 0.2746 0.3762 0.73 0.46767
## Year2000 -0.4045 0.3588 -1.13 0.26319
## Year2002 -0.5406 0.3292 -1.64 0.10487
## Year2003 -0.3005 0.3394 -0.89 0.37890
## Year2004 0.7850 0.1427 5.50 5.3e-07 ***
## Year2005 0.1261 0.2133 0.59 0.55616
## Year2006 -0.1260 0.2742 -0.46 0.64717
## Year2007 0.0484 0.2769 0.17 0.86164
```

```

## Year2008          -0.7868      0.2164    -3.64  0.00051 ***
## Year2009           0.0479      0.2879      0.17  0.86838
## Year2010          -0.6030      0.1928    -3.13  0.00252 **
## Year2011          -0.5437      0.2601    -2.09  0.04003 *
## Year2012          -0.6583      0.2463    -2.67  0.00929 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.612
## Multiple R-squared:  0.273, Adjusted R-squared:  0.113
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 82 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.284  0.890  0.954  0.918  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      1.11e-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 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
##   10   10   9   19   25   17   21   14   20   4   16   12   14   19   18
## 2011 2012
##   31   25
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    8    7    8   16   16   13   18   13   14    4   12    9   11   15   15
## 2011 2012
##   26   22

```

```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7    7    7   16   16    8   18   12   13    4   12    9   11   14   15
## 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 = 13, df = 16, p-value = 0.7
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.14, df = 1, p-value = 0.7

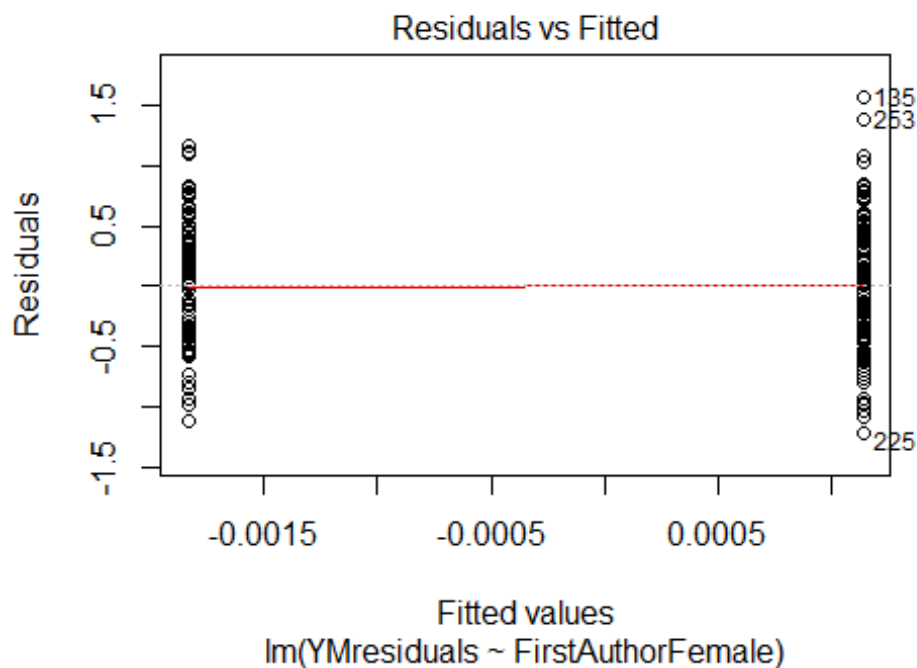
## [1] "Female first author team size 2018 geometric mean: 2"
## [1] "Male first author team size 2018 geometric mean: 1.56508458007329"

## Warning in wilcox.test.default(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 = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.8881750225898"
## [1] "Male last author team size 2018 geometric mean: 1.65453202244151"

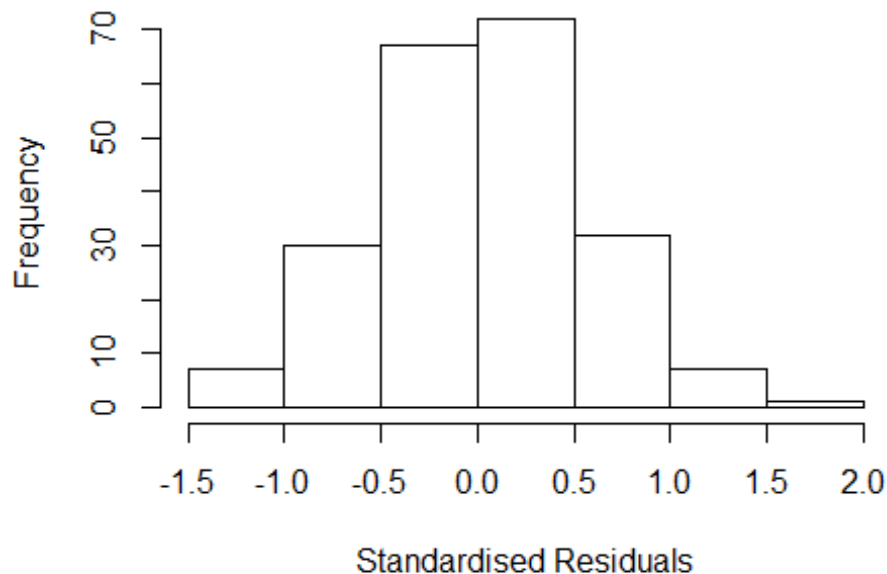
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 40, p-value = 0.6
## 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 | 2.089 | 1 | 1.445 |
| LastAuthorFemale | 1.941 | 1 | 1.393 |
| UniqueAuthors | 2.090 | 3 | 1.131 |
| Year | 3.888 | 16 | 1.043 |

Residuals from first and last author and team size



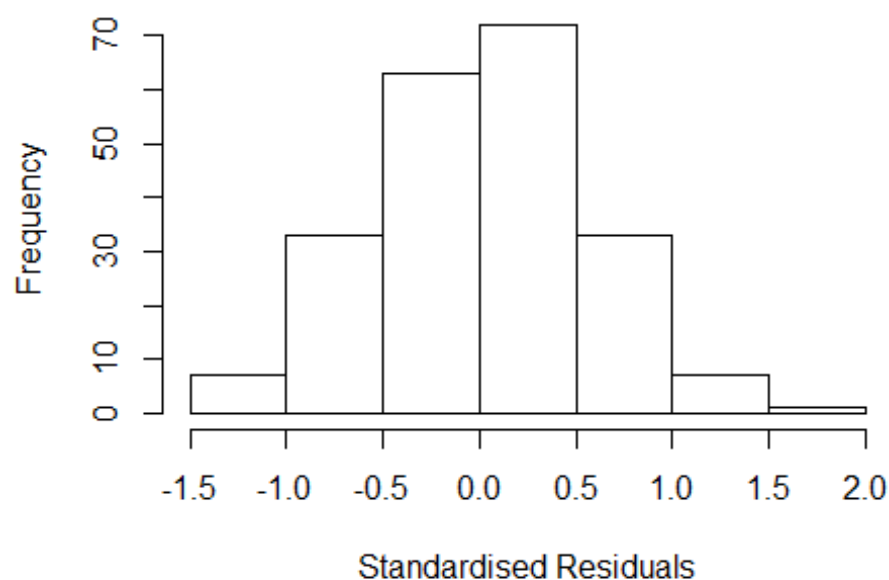
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2933 -0.3905 0.0475 0.3803 1.6471
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.03418 0.13788 7.50 2.2e-12 ***
## FirstAuthorFemale1 0.02275 0.10452 0.22 0.8279
## LastAuthorFemale1 -0.01369 0.10466 -0.13 0.8961
## UniqueAuthors2 0.05964 0.09558 0.62 0.5334
## UniqueAuthors3 -0.12665 0.13264 -0.95 0.3409
## UniqueAuthors4 0.61728 0.20339 3.03 0.0027 **
## Year1997 0.45515 0.18620 2.44 0.0154 *
## Year1998 -0.19886 0.23016 -0.86 0.3887
## Year1999 -0.07757 0.23882 -0.32 0.7457
## Year2000 0.37021 0.20706 1.79 0.0753 .
```

```

## Year2001      0.23954      0.16388      1.46      0.1454
## Year2002      0.09137      0.17095      0.53      0.5936
## Year2003      0.00269      0.25292      0.01      0.9915
## Year2004     -0.05470      0.24740     -0.22      0.8253
## Year2005     -0.20733      0.20554     -1.01      0.3144
## Year2006      0.04975      0.18878      0.26      0.7924
## Year2007     -0.07945      0.24114     -0.33      0.7422
## Year2008     -0.13507      0.25011     -0.54      0.5898
## Year2009      0.25910      0.23176      1.12      0.2650
## Year2010     -0.14732      0.21911     -0.67      0.5021
## Year2011     -0.13282      0.16664     -0.80      0.4264
## Year2012      0.00406      0.17864      0.02      0.9819
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.605
## Multiple R-squared:  0.104, Adjusted R-squared:  0.00751
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 204 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.439  0.902  0.959  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      4.63e-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 1.995 1      1.413
## LastAuthorFemale  1.907 1      1.381
## Year              1.986 16      1.022

```

Residuals from first and last author



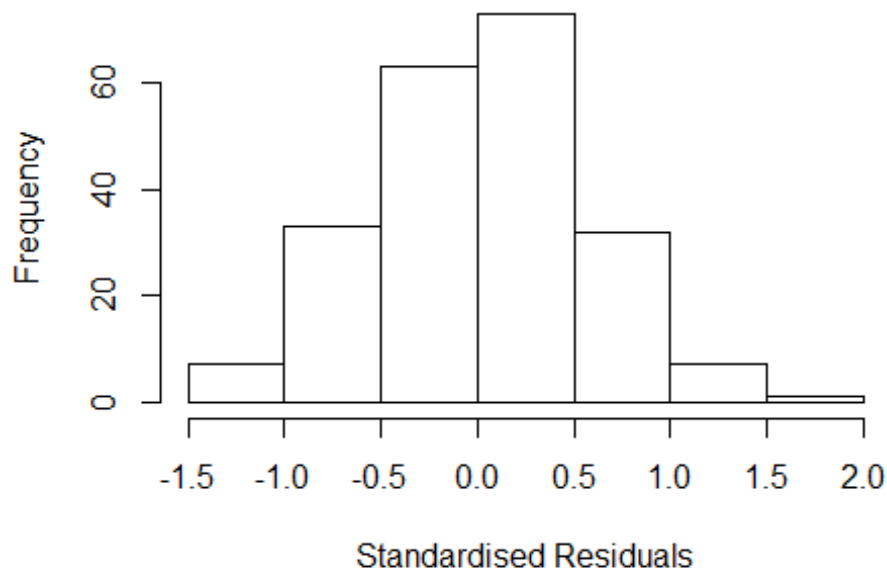
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2808 -0.4114 0.0578 0.3797 1.6468
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.05434 0.13742 7.67 7.6e-13 ***
## FirstAuthorFemale1 0.00413 0.10506 0.04 0.969
## LastAuthorFemale1 0.01450 0.10554 0.14 0.891
## Year1997 0.45120 0.18732 2.41 0.017 *
## Year1998 -0.20833 0.23473 -0.89 0.376
## Year1999 -0.08982 0.24136 -0.37 0.710
## Year2000 0.36169 0.20531 1.76 0.080 .
## Year2001 0.18175 0.16733 1.09 0.279
## Year2002 0.10560 0.17788 0.59 0.553
## Year2003 -0.01715 0.26302 -0.07 0.948
## Year2004 -0.07832 0.24305 -0.32 0.748
## Year2005 -0.23195 0.20473 -1.13 0.259
```

```

## Year2006          0.03953      0.19411      0.20      0.839
## Year2007          -0.07680      0.24706     -0.31      0.756
## Year2008          -0.08767      0.24846     -0.35      0.725
## Year2009          0.22647      0.23021      0.98      0.326
## Year2010          -0.12274      0.22805     -0.54      0.591
## Year2011          -0.15109      0.17255     -0.88      0.382
## Year2012          -0.01857      0.18435     -0.10      0.920
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.609
## Multiple R-squared:  0.0838, Adjusted R-squared:  0.000135
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 206 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
## 0.444 0.904 0.959 0.925 0.985 0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      4.63e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.342 1      1.159
## Year              1.342 16      1.009

```

Residuals from first author



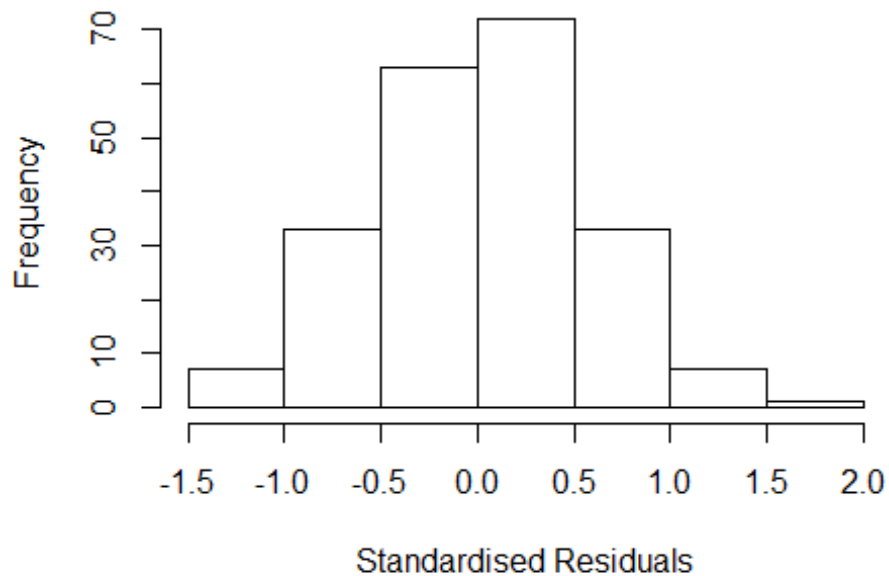
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2843 -0.4136 0.0537 0.3781 1.6483
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0572 0.1355 7.80 3.4e-13 ***
## FirstAuthorFemale1 0.0137 0.0878 0.16 0.876
## Year1997 0.4489 0.1866 2.41 0.017 *
## Year1998 -0.2114 0.2316 -0.91 0.362
## Year1999 -0.0917 0.2431 -0.38 0.706
## Year2000 0.3566 0.1983 1.80 0.074 .
## Year2001 0.1819 0.1679 1.08 0.280
## Year2002 0.1019 0.1768 0.58 0.565
## Year2003 -0.0216 0.2572 -0.08 0.933
## Year2004 -0.0822 0.2425 -0.34 0.735
## Year2005 -0.2326 0.2050 -1.13 0.258
## Year2006 0.0404 0.1950 0.21 0.836
```

```

## Year2007          -0.0762      0.2474   -0.31    0.759
## Year2008          -0.0877      0.2500   -0.35    0.726
## Year2009           0.2271      0.2309    0.98    0.327
## Year2010          -0.1245      0.2291   -0.54    0.587
## Year2011          -0.1531      0.1714   -0.89    0.373
## Year2012          -0.0194      0.1851   -0.11    0.916
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.602
## Multiple R-squared:  0.0841, Adjusted R-squared:  0.00549
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 205 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.434  0.901  0.958  0.924  0.984  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      4.63e-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 1.288 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
## -1.2806 -0.4091 0.0586 0.3787 1.6454
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0541 0.1372 7.68 7.1e-13 ***
## LastAuthorFemale1 0.0174 0.0887 0.20 0.845
## Year1997 0.4516 0.1869 2.42 0.017 *
## Year1998 -0.2071 0.2309 -0.90 0.371
## Year1999 -0.0886 0.2413 -0.37 0.714
## Year2000 0.3637 0.1895 1.92 0.056 .
## Year2001 0.1820 0.1668 1.09 0.277
## Year2002 0.1070 0.1761 0.61 0.544
## Year2003 -0.0155 0.2571 -0.06 0.952
## Year2004 -0.0770 0.2405 -0.32 0.749
## Year2005 -0.2311 0.2038 -1.13 0.258
## Year2006 0.0400 0.1937 0.21 0.837
```

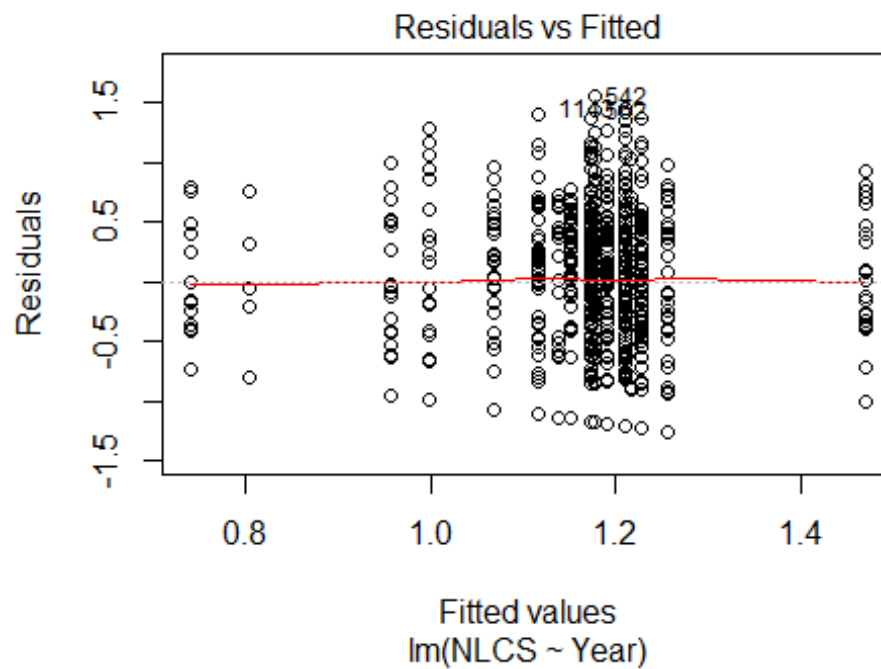
```

## Year2007          -0.0756      0.2416   -0.31    0.755
## Year2008          -0.0875      0.2490   -0.35    0.726
## Year2009           0.2265      0.2304    0.98    0.327
## Year2010          -0.1218      0.2274   -0.54    0.593
## Year2011          -0.1504      0.1714   -0.88    0.381
## Year2012          -0.0179      0.1822   -0.10    0.922
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.61
## Multiple R-squared:  0.0838, Adjusted R-squared:  0.00513
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 207 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.447  0.905  0.960  0.926  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      4.63e-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: 216"
## [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
##    7   29   18   35   40   22   30   38   35   29   60   62   96  109  133
## 2011 2012
##  135  159
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5   22   14   20   29   20   28   31   30   27   49   47   80   92  101
## 2011 2012

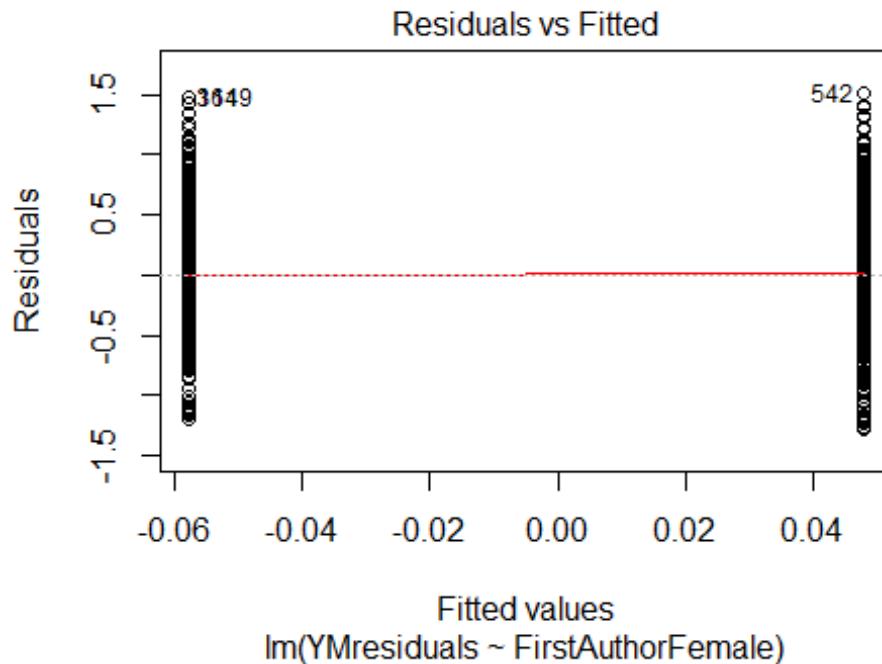
```



```
## 110 127
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4 22 14 17 28 18 25 26 28 26 46 43 74 82 95
## 2011 2012
## 101 113
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 9.3, df = 16, p-value = 0.9
```

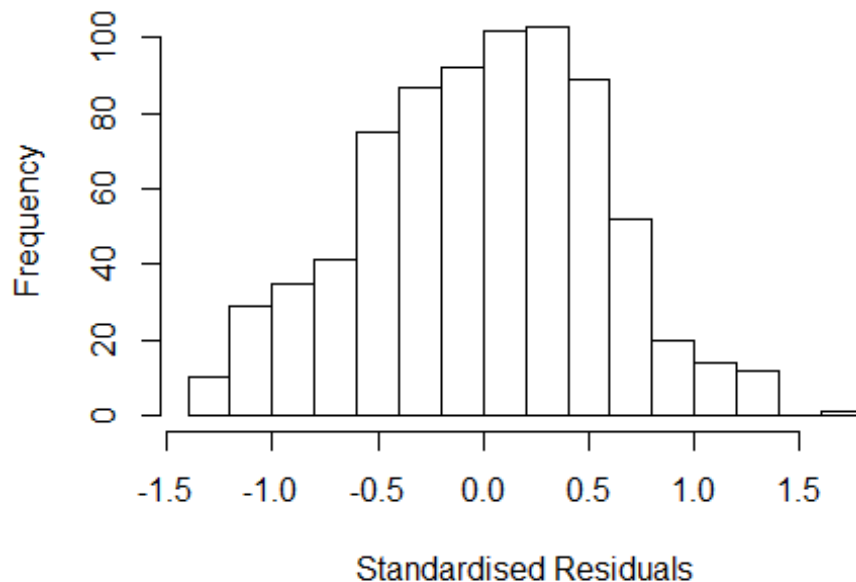


```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.32, df = 1, p-value = 0.6
```



```
## [1] "Female first author team size 2018 geometric mean: 2.03400180834285"
## [1] "Male first author team size 2018 geometric mean: 1.61765365589512"
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3600, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.87187867203252"
## [1] "Male last author team size 2018 geometric mean: 1.83886438529196"
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3100, 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 1.831 1      1.353
## LastAuthorFemale  1.758 1      1.326
## UniqueAuthors    1.506 4      1.052
## Year              1.708 16     1.017
```

Residuals from first and last author and team size



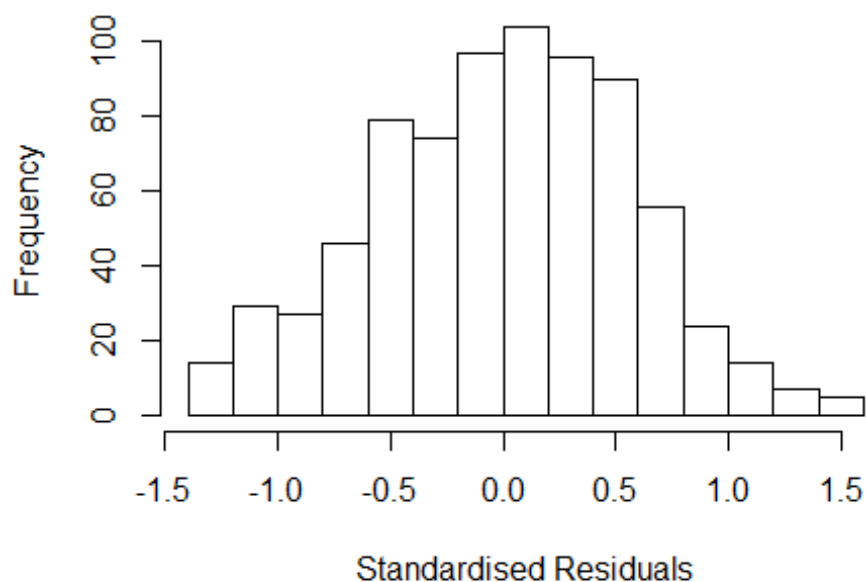
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3190 -0.3972 0.0329 0.3965 1.6354
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6492 0.3239 2.00 0.0454 *
## FirstAuthorFemale1 0.0829 0.0564 1.47 0.1419
## LastAuthorFemale1 0.0590 0.0548 1.08 0.2822
## UniqueAuthors2 0.1658 0.0515 3.22 0.0013 **
## UniqueAuthors3 0.2181 0.0701 3.11 0.0019 **
## UniqueAuthors4 0.2319 0.0915 2.54 0.0114 *
## UniqueAuthors5 0.2357 0.0965 2.44 0.0148 *
## Year1997 0.2203 0.3858 0.57 0.5681
## Year1998 0.0149 0.3395 0.04 0.9649
## Year1999 0.3253 0.3558 0.91 0.3608
```

```

## Year2000          0.3620      0.3369      1.07      0.2829
## Year2001          0.1198      0.3439      0.35      0.7276
## Year2002          0.3145      0.3380      0.93      0.3524
## Year2003          0.2114      0.3427      0.62      0.5375
## Year2004          0.4564      0.3461      1.32      0.1877
## Year2005          0.5782      0.3365      1.72      0.0861 .
## Year2006          0.3348      0.3332      1.00      0.3154
## Year2007          0.3621      0.3344      1.08      0.2792
## Year2008          0.3398      0.3290      1.03      0.3021
## Year2009          0.3852      0.3288      1.17      0.2417
## Year2010          0.4211      0.3273      1.29      0.1986
## Year2011          0.3222      0.3262      0.99      0.3236
## Year2012          0.3384      0.3257      1.04      0.2991
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.573
## Multiple R-squared:  0.0721, Adjusted R-squared:  0.0445
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 57 weights are ~= 1. The remaining 705 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.396  0.878  0.948  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          1.31e-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 1.813 1          1.347
## LastAuthorFemale 1.729 1          1.315
## Year          1.205 16          1.006

```

Residuals from first and last author



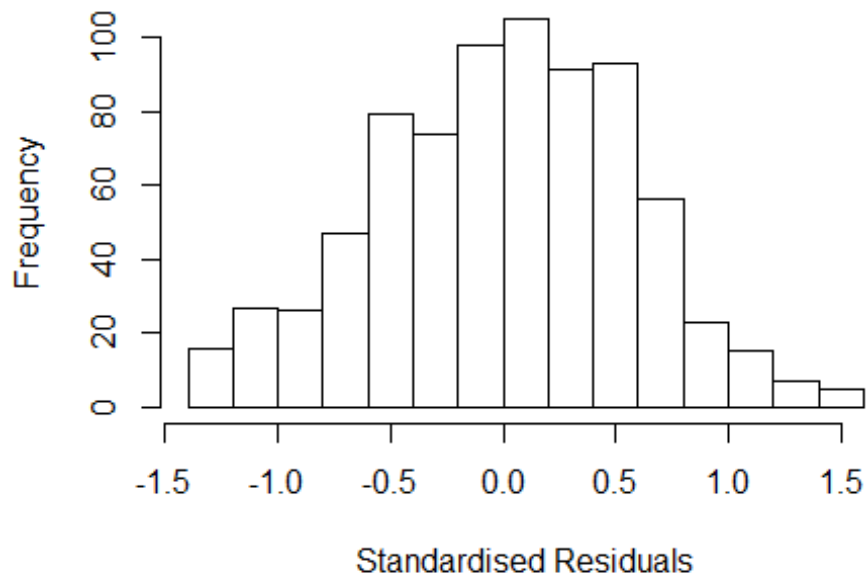
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3110 -0.4036  0.0295  0.4115  1.5305
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.7956    0.3504   2.27   0.023 *
## FirstAuthorFemale1 0.1080    0.0571   1.89   0.059 .
## LastAuthorFemale1  0.0380    0.0553   0.69   0.492
## Year1997         0.1385    0.4066   0.34   0.733
## Year1998        -0.0945    0.3706  -0.25   0.799
## Year1999         0.2350    0.3851   0.61   0.542
## Year2000         0.3069    0.3639   0.84   0.399
## Year2001         0.0661    0.3717   0.18   0.859
## Year2002         0.2532    0.3681   0.69   0.492
## Year2003         0.1413    0.3705   0.38   0.703
## Year2004         0.3908    0.3735   1.05   0.296
## Year2005         0.5499    0.3660   1.50   0.133
```

```

## Year2006          0.2505      0.3617      0.69      0.489
## Year2007          0.3073      0.3607      0.85      0.395
## Year2008          0.2950      0.3574      0.83      0.409
## Year2009          0.3537      0.3570      0.99      0.322
## Year2010          0.3695      0.3550      1.04      0.298
## Year2011          0.2848      0.3547      0.80      0.422
## Year2012          0.2969      0.3547      0.84      0.403
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.582
## Multiple R-squared:  0.0459, Adjusted R-squared:  0.0227
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 68 weights are ~= 1. The remaining 694 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.469  0.876  0.949  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      1.31e-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 1.146 1          1.070
## Year              1.146 16          1.004

```

Residuals from first author



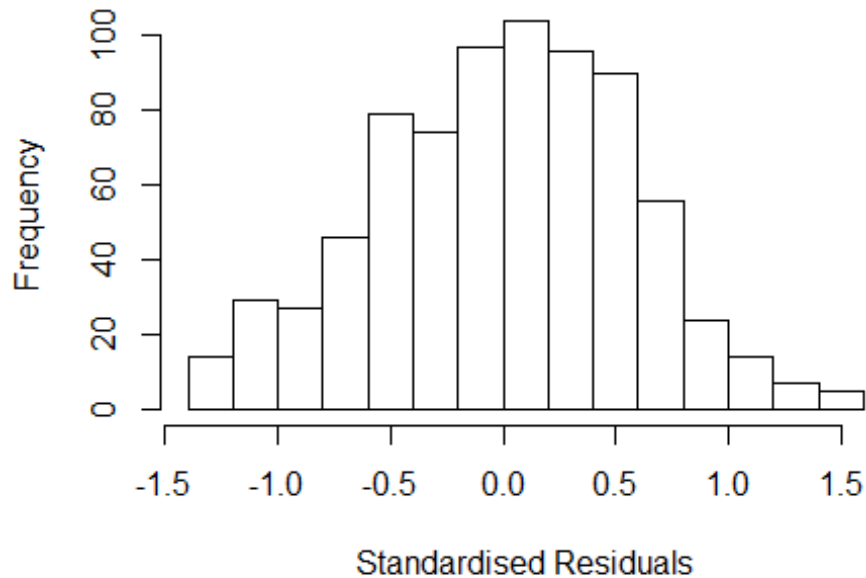
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3038 -0.4059 0.0263 0.4118 1.5252
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7881 0.3535 2.23 0.0261 *
## FirstAuthorFemale1 0.1317 0.0455 2.90 0.0039 **
## Year1997 0.1511 0.4077 0.37 0.7110
## Year1998 -0.0866 0.3734 -0.23 0.8167
## Year1999 0.2447 0.3867 0.63 0.5271
## Year2000 0.3175 0.3662 0.87 0.3863
## Year2001 0.0797 0.3746 0.21 0.8316
## Year2002 0.2638 0.3710 0.71 0.4773
## Year2003 0.1529 0.3734 0.41 0.6823
## Year2004 0.4044 0.3758 1.08 0.2823
## Year2005 0.5605 0.3683 1.52 0.1285
## Year2006 0.2640 0.3639 0.73 0.4684
```

```

## Year2007          0.3199      0.3631      0.88      0.3785
## Year2008          0.3114      0.3597      0.87      0.3869
## Year2009          0.3681      0.3591      1.03      0.3056
## Year2010          0.3840      0.3573      1.07      0.2829
## Year2011          0.3009      0.3565      0.84      0.3989
## Year2012          0.3097      0.3570      0.87      0.3860
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.582
## Multiple R-squared:  0.0452, Adjusted R-squared:  0.0234
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 68 weights are ~= 1. The remaining 694 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.472  0.878  0.949  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      1.31e-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 1.089 1      1.043
## Year      1.089 16      1.003

```


Residuals from last author



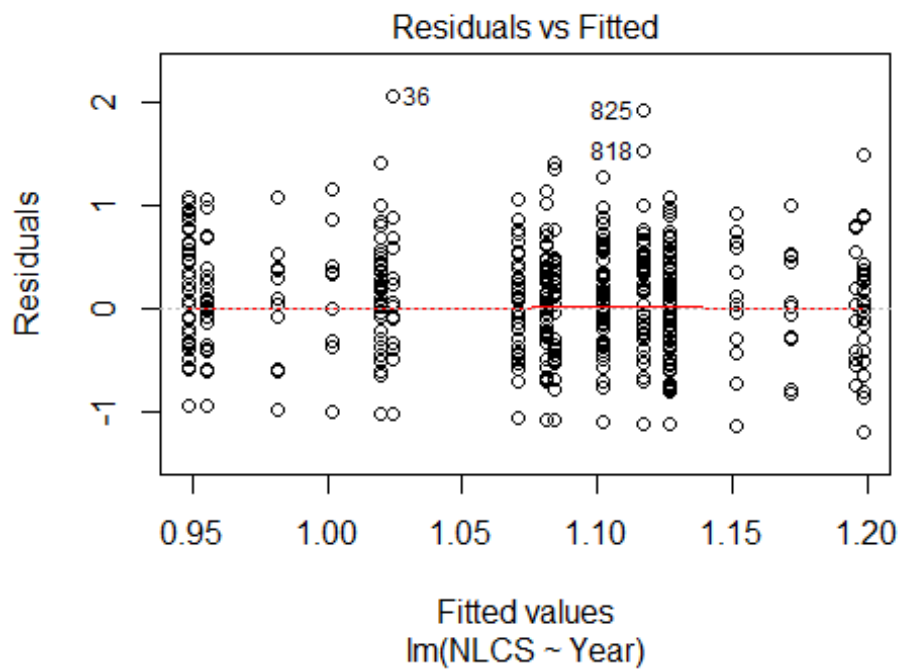
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2972 -0.4019 0.0123 0.4094 1.5219
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8300 0.3341 2.48 0.013 *
## LastAuthorFemale1 0.1034 0.0441 2.35 0.019 *
## Year1997 0.1097 0.3907 0.28 0.779
## Year1998 -0.1162 0.3570 -0.33 0.745
## Year1999 0.2392 0.3735 0.64 0.522
## Year2000 0.3024 0.3507 0.86 0.389
## Year2001 0.0507 0.3556 0.14 0.887
## Year2002 0.2654 0.3536 0.75 0.453
## Year2003 0.1297 0.3553 0.37 0.715
## Year2004 0.3843 0.3587 1.07 0.284
## Year2005 0.5513 0.3520 1.57 0.118
## Year2006 0.2441 0.3475 0.70 0.483
```

```

## Year2007          0.3117      0.3458      0.90      0.368
## Year2008          0.2767      0.3423      0.81      0.419
## Year2009          0.3376      0.3415      0.99      0.323
## Year2010          0.3638      0.3401      1.07      0.285
## Year2011          0.2729      0.3398      0.80      0.422
## Year2012          0.2965      0.3403      0.87      0.384
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.583
## Multiple R-squared:  0.041, Adjusted R-squared:  0.0191
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 68 weights are ~= 1. The remaining 694 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.475  0.876  0.949  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      1.31e-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: 762"
## [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
##   11   16   22   28   27   20   22   30   43   40   46   59   56   70   62
## 2011 2012
##   81   78
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   10   13   19   18   11   12   18   26   32   35   38   45   43   63   51
## 2011 2012

```

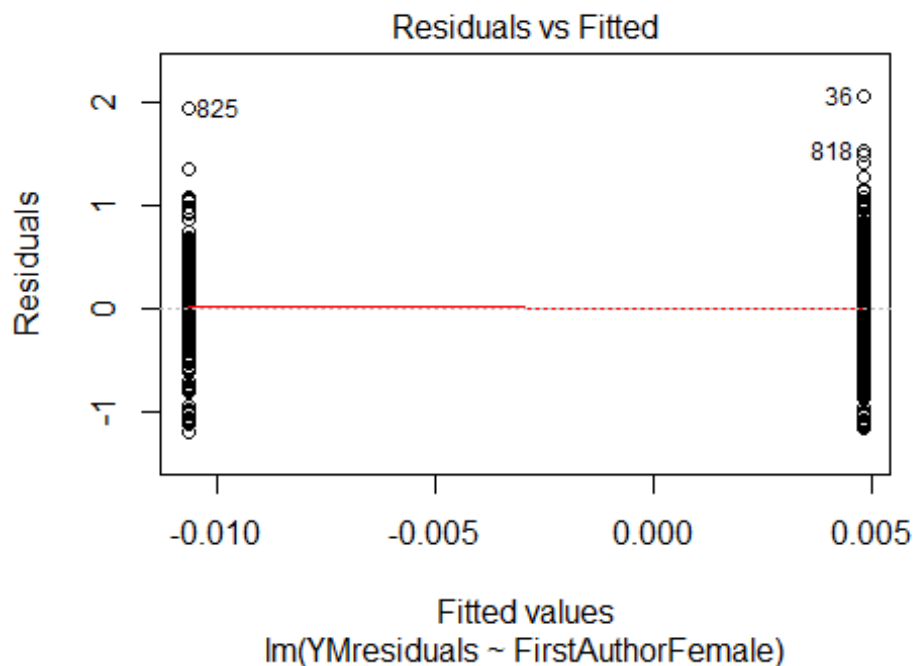
```
## 67 59
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 9 12 17 17 10 10 17 26 30 33 37 42 42 60 49
## 2011 2012
## 64 55
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 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: 1.64809606116061"
## [1] "Male first author team size 2018 geometric mean: 1.83552615153557"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

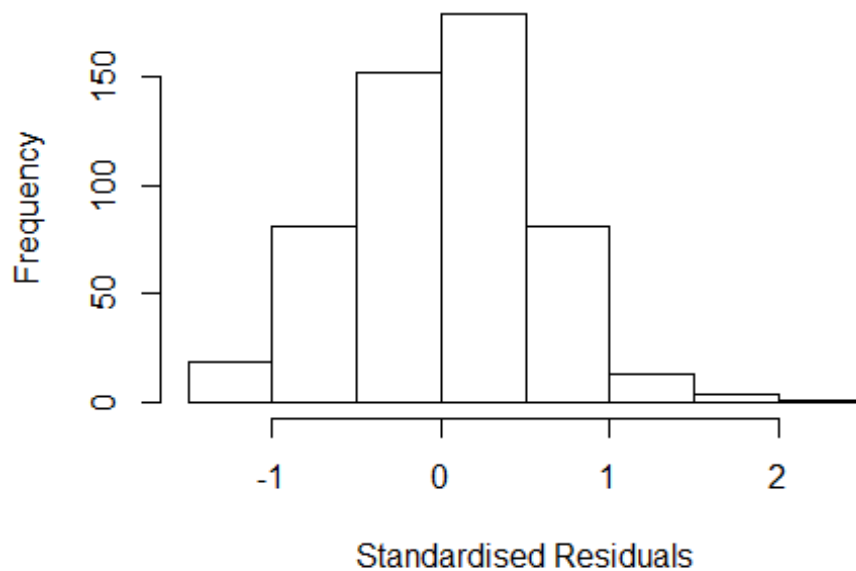
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 410, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.63795872675789"
## [1] "Male last author team size 2018 geometric mean: 1.83196231948013"

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



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 410, 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 1.676 1      1.295
## LastAuthorFemale  1.720 1      1.311
## UniqueAuthors    1.987 4      1.090
## Year              2.208 16     1.025
```

Residuals from first and last author and team size



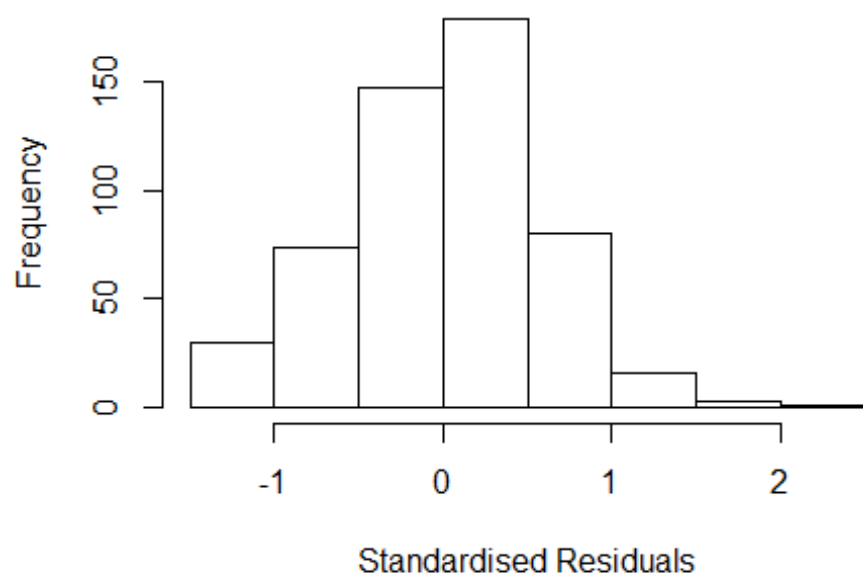
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2371 -0.3684 0.0378 0.3686 2.0955
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.01711 0.16193 6.28 7.2e-10 ***
## FirstAuthorFemale1 -0.02403 0.06532 -0.37 0.71309
## LastAuthorFemale1 0.04804 0.06561 0.73 0.46438
## UniqueAuthors2 0.23465 0.06503 3.61 0.00034 ***
## UniqueAuthors3 0.17076 0.07646 2.23 0.02595 *
## UniqueAuthors4 0.27889 0.13594 2.05 0.04072 *
## UniqueAuthors5 -0.06744 0.18293 -0.37 0.71252
## Year1997 0.15103 0.24081 0.63 0.53084
## Year1998 -0.03256 0.22458 -0.14 0.88478
## Year1999 -0.09306 0.21603 -0.43 0.66682
```

```

## Year2000      -0.11461    0.31258   -0.37  0.71403
## Year2001      0.10684    0.23686    0.45  0.65214
## Year2002      0.02311    0.19106    0.12  0.90379
## Year2003      0.06556    0.19504    0.34  0.73691
## Year2004     -0.07815    0.18947   -0.41  0.68018
## Year2005     -0.03032    0.18409   -0.16  0.86924
## Year2006     -0.03916    0.17908   -0.22  0.82698
## Year2007     -0.19495    0.19168   -1.02  0.30961
## Year2008     -0.09663    0.18471   -0.52  0.60111
## Year2009      0.00716    0.17460    0.04  0.96733
## Year2010     -0.04512    0.17237   -0.26  0.79361
## Year2011     -0.03487    0.17169   -0.20  0.83913
## Year2012      0.02128    0.17959    0.12  0.90574
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.558
## Multiple R-squared:  0.0507, Adjusted R-squared:  0.00947
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 41 weights are ~= 1. The remaining 489 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.128  0.872  0.952  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      1.89e-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 1.571 1      1.254
## LastAuthorFemale  1.671 1      1.293
## Year              1.222 16      1.006

```

Residuals from first and last author



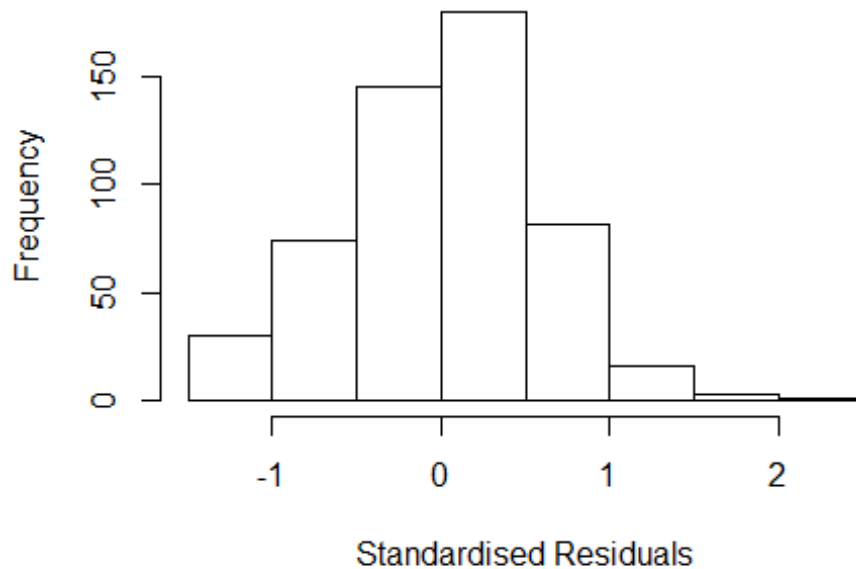
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2382 -0.3809 0.0451 0.3879 2.0872
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.08840 0.18183 5.99 4.1e-09 ***
## FirstAuthorFemale1 -0.02053 0.06565 -0.31 0.75
## LastAuthorFemale1 0.04020 0.06613 0.61 0.54
## Year1997 0.14980 0.26305 0.57 0.57
## Year1998 -0.09565 0.23579 -0.41 0.69
## Year1999 -0.12060 0.22989 -0.52 0.60
## Year2000 -0.15145 0.30997 -0.49 0.63
## Year2001 0.04705 0.25827 0.18 0.86
## Year2002 0.02140 0.21048 0.10 0.92
## Year2003 0.09257 0.21298 0.43 0.66
## Year2004 -0.12530 0.21023 -0.60 0.55
## Year2005 -0.04445 0.20431 -0.22 0.83
```

```

## Year2006      -0.02435      0.20147      -0.12      0.90
## Year2007      -0.20334      0.20985      -0.97      0.33
## Year2008      -0.08584      0.20525      -0.42      0.68
## Year2009       0.03806      0.19595       0.19      0.85
## Year2010      -0.03112      0.19468      -0.16      0.87
## Year2011      -0.00333      0.19211      -0.02      0.99
## Year2012       0.00571      0.20124       0.03      0.98
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.58
## Multiple R-squared:  0.0192, Adjusted R-squared:  -0.0154
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 47 weights are ~= 1. The remaining 483 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.168  0.874  0.954  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      1.89e-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 1.1 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
## -1.241 -0.379 0.045 0.384 2.081
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.091768 0.182178 5.99 3.9e-09 ***
## FirstAuthorFemale1 0.005459 0.055154 0.10 0.92
## Year1997 0.149328 0.263860 0.57 0.57
## Year1998 -0.092612 0.235812 -0.39 0.69
## Year1999 -0.122600 0.229375 -0.53 0.59
## Year2000 -0.154223 0.309466 -0.50 0.62
## Year2001 0.046473 0.258203 0.18 0.86
## Year2002 0.018225 0.210290 0.09 0.93
## Year2003 0.090503 0.213429 0.42 0.67
## Year2004 -0.121668 0.210506 -0.58 0.56
## Year2005 -0.046631 0.204575 -0.23 0.82
## Year2006 -0.020959 0.201887 -0.10 0.92
```

```

## Year2007          -0.204469    0.210608   -0.97     0.33
## Year2008          -0.086711    0.205910   -0.42     0.67
## Year2009           0.039343    0.196570    0.20     0.84
## Year2010          -0.027969    0.194897   -0.14     0.89
## Year2011          -0.000178    0.192385    0.00     1.00
## Year2012           0.005987    0.201893    0.03     0.98
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.579
## Multiple R-squared:  0.0186, Adjusted R-squared:  -0.014
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 46 weights are ~= 1. The remaining 484 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.169  0.873  0.954  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      1.89e-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 1.172 1      1.083
## Year              1.172 16      1.005
##
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.237 -0.380  0.048  0.389  2.090

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.08646    0.18180    5.98 4.3e-09 ***
## LastAuthorFemale1 0.02657    0.05583    0.48  0.63
## Year1997        0.15056    0.26315    0.57  0.57
## Year1998       -0.09654    0.23556   -0.41  0.68
## Year1999       -0.12115    0.23023   -0.53  0.60
## Year2000       -0.14996    0.31045   -0.48  0.63
## Year2001        0.04758    0.25874    0.18  0.85
## Year2002        0.01910    0.21073    0.09  0.93
## Year2003        0.09201    0.21324    0.43  0.67
## Year2004       -0.12536    0.21070   -0.59  0.55
## Year2005       -0.04568    0.20451   -0.22  0.82
## Year2006       -0.02344    0.20166   -0.12  0.91
## Year2007       -0.20444    0.21024   -0.97  0.33
## Year2008       -0.08622    0.20570   -0.42  0.68
## Year2009        0.03805    0.19637    0.19  0.85
## Year2010       -0.03232    0.19509   -0.17  0.87
## Year2011       -0.00329    0.19246   -0.02  0.99
## Year2012        0.00490    0.20171    0.02  0.98
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.579
## Multiple R-squared:  0.0191, Adjusted R-squared:  -0.0135
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 47 weights are ~= 1. The remaining 483 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.165  0.875  0.954  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      1.89e-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: 530"
## [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
##    3    5    3    1    6    2    8    6    5   16   17   16   18   23   43
## 2011 2012
##   38   38
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    2    3    0    1    2    4    3    4   12   13   11   16   21   34
## 2011 2012
##   31   35
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    2    2    0    1    2    4    3    4   12   13   11   16   21   29
## 2011 2012
##   31   34
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.14330750297728"
## [1] "Male first author team size 2018 geometric mean: 1.86536360324116"

## Warning in wilcox.test.default(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 = 0.01
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.97683233070989"
## [1] "Male last author team size 2018 geometric mean: 2.06285203796338"

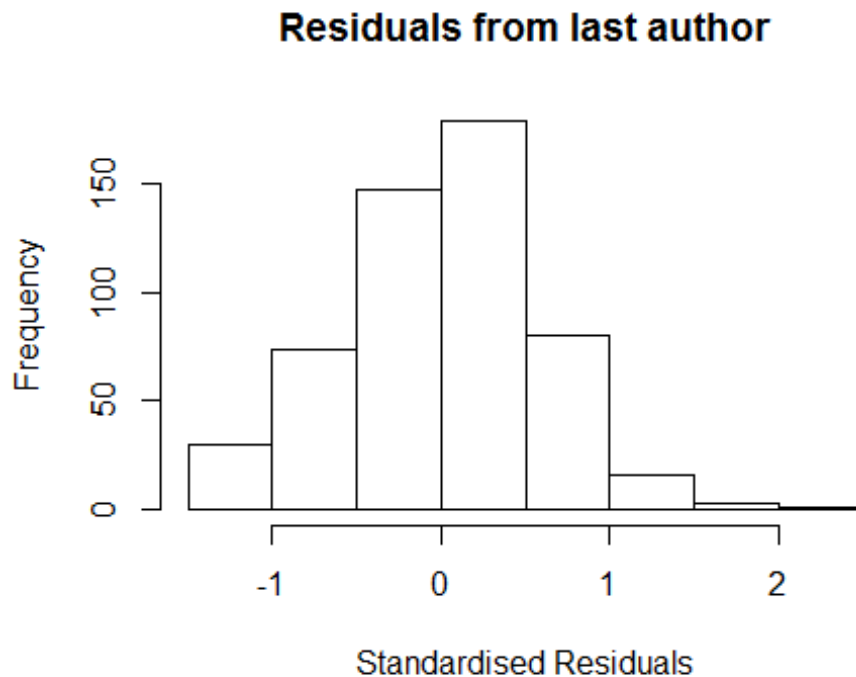
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 210, p-value = 0.09
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"

```

```
## 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 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
## -1.6382 -0.3013  0.0131  0.3145  1.4816
##
```

```

## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3120    0.0000      Inf <2e-16 ***
## FirstAuthorFemale1 -0.1483    0.0881    -1.68  0.0941 .
## LastAuthorFemale1  0.1164    0.0802     1.45  0.1483
## UniqueAuthors2     0.0732    0.1089     0.67  0.5022
## UniqueAuthors3     0.1226    0.1170     1.05  0.2961
## UniqueAuthors4     0.0515    0.1360     0.38  0.7056
## UniqueAuthors5     0.3105    0.1460     2.13  0.0350 *
## Year1997          -0.2905    0.1567    -1.85  0.0655 .
## Year1998          -0.7540    0.4669    -1.62  0.1082
## Year2000           0.9000    0.0000      Inf <2e-16 ***
## Year2001           0.3260    0.2331     1.40  0.1638
## Year2002          -0.4670    0.1401    -3.33  0.0011 **
## Year2003           0.1809    0.1574     1.15  0.2522
## Year2004          -0.0449    0.4543    -0.10  0.9213
## Year2005          -0.1944    0.2016    -0.96  0.3363
## Year2006          -0.1111    0.2113    -0.53  0.6000
## Year2007           0.0674    0.1533     0.44  0.6607
## Year2008          -0.1093    0.1358    -0.80  0.4222
## Year2009          -0.1706    0.1220    -1.40  0.1639
## Year2010          -0.1598    0.1385    -1.15  0.2501
## Year2011          -0.1850    0.1142    -1.62  0.1073
## Year2012          -0.1007    0.1375    -0.73  0.4650
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.517
## Multiple R-squared:  0.103, Adjusted R-squared:  -0.0123
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 171 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.295  0.885  0.960  0.910  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          5.38e-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"

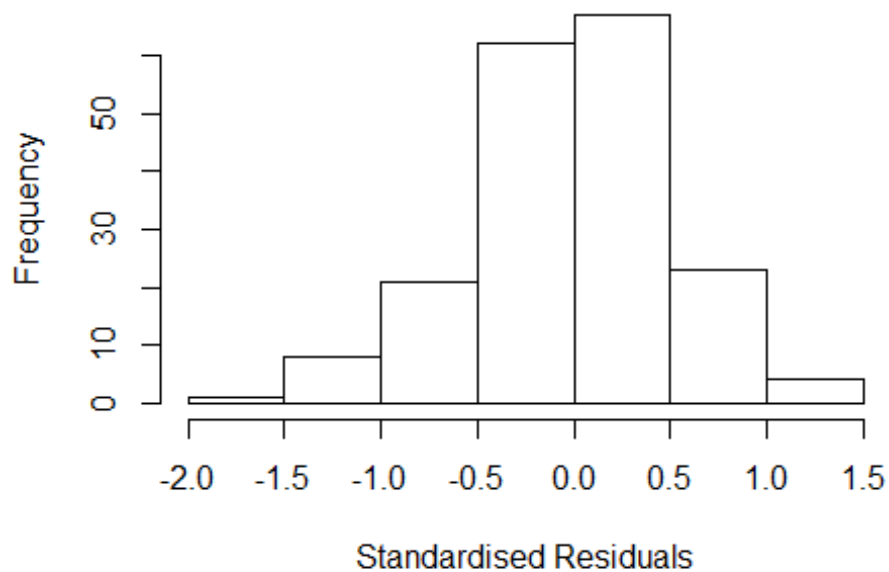
```

```
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"

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

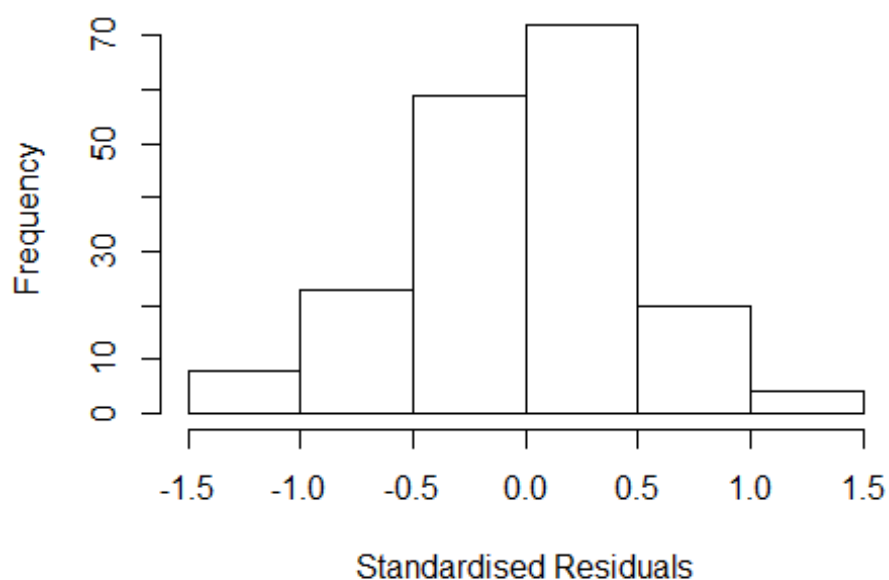
## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

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
## -1.3914 -0.2998  0.0142  0.3196  1.3900
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3120     0.0000      Inf < 2e-16 ***
## FirstAuthorFemale1 -0.1203     0.0838    -1.44  0.15307
## LastAuthorFemale1  0.1176     0.0827     1.42  0.15694
## Year1997          -0.2298     0.1050    -2.19  0.03006 *
## Year1998          -0.7686     0.4907    -1.57  0.11916
## Year2000           0.9000     0.0000      Inf < 2e-16 ***
## Year2001           0.3260     0.2336     1.40  0.16476
## Year2002          -0.4613     0.1375    -3.35  0.00098 ***
## Year2003           0.2074     0.1999     1.04  0.30084
## Year2004           0.0202     0.4851     0.04  0.96678
## Year2005          -0.1911     0.1912    -1.00  0.31921
## Year2006          -0.0702     0.1966    -0.36  0.72161
```

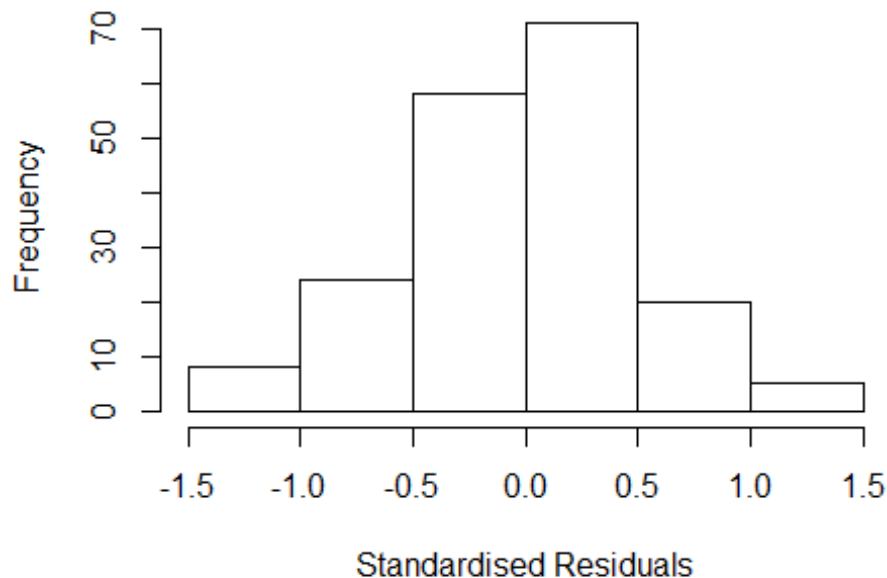


```

## Year2007          0.0614      0.1523      0.40  0.68730
## Year2008         -0.0678      0.1291     -0.52  0.60049
## Year2009         -0.0931      0.1085     -0.86  0.39238
## Year2010         -0.1237      0.1255     -0.99  0.32544
## Year2011         -0.1518      0.1048     -1.45  0.14927
## Year2012         -0.0382      0.1249     -0.31  0.76010
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.509
## Multiple R-squared:  0.0847, Adjusted R-squared:  -0.00797
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 167 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.435  0.883  0.958  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      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 3.017 1      1.737
## Year              3.017 15      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
## -1.3109 -0.3115 0.0146 0.3276 1.4114
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.31e+00 1.25e-07 1.05e+07 <2e-16 ***
## FirstAuthorFemale1 -6.12e-02 8.59e-02 -7.10e-01 0.4767
## Year1997 -1.71e-01 5.40e-02 -3.17e+00 0.0018 **
## Year1998 -7.39e-01 4.47e-01 -1.65e+00 0.0998 .
## Year2000 9.00e-01 1.25e-07 7.21e+06 <2e-16 ***
## Year2001 3.26e-01 2.33e-01 1.40e+00 0.1639
## Year2002 -4.47e-01 1.44e-01 -3.09e+00 0.0023 **
## Year2003 2.28e-01 1.89e-01 1.21e+00 0.2294
## Year2004 4.07e-02 4.30e-01 9.00e-02 0.9248
## Year2005 -1.59e-01 1.94e-01 -8.20e-01 0.4135
## Year2006 -4.17e-02 1.97e-01 -2.10e-01 0.8328
## Year2007 1.02e-01 1.43e-01 7.10e-01 0.4789
```

```

## Year2008          -5.74e-02   1.32e-01 -4.40e-01   0.6636
## Year2009          -7.25e-02   1.08e-01 -6.70e-01   0.5041
## Year2010          -1.06e-01   1.26e-01 -8.40e-01   0.4021
## Year2011          -1.12e-01   9.51e-02 -1.17e+00   0.2419
## Year2012          -1.15e-03   1.21e-01 -1.00e-02   0.9924
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.516
## Multiple R-squared:  0.0764, Adjusted R-squared:  -0.011
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 171 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.435  0.887   0.959   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      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 2.92  1          1.709
## Year              2.92 15          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
## -1.2903 -0.2908  0.0329  0.3162  1.3707
##

```

```

## Coefficients:
##              Estimate Std. Error   t value Pr(>|t|)
## (Intercept)   1.31e+00   8.03e-08  1.63e+07  <2e-16 ***
## LastAuthorFemale1 6.03e-02   8.41e-02  7.20e-01  0.4746
## Year1997      -2.01e-01   8.66e-02 -2.32e+00  0.0213 *
## Year1998      -8.00e-01   5.39e-01 -1.49e+00  0.1392
## Year2000       9.00e-01   1.03e-07  8.76e+06  <2e-16 ***
## Year2001       3.26e-01   2.34e-01  1.39e+00  0.1650
## Year2002      -4.77e-01   1.29e-01 -3.69e+00  0.0003 ***
## Year2003       1.85e-01   2.10e-01  8.80e-01  0.3779
## Year2004      -4.55e-03   4.35e-01 -1.00e-02  0.9917
## Year2005      -2.26e-01   1.80e-01 -1.26e+00  0.2111
## Year2006      -9.82e-02   1.87e-01 -5.30e-01  0.6002
## Year2007      -6.61e-04   1.41e-01  0.00e+00  0.9963
## Year2008      -1.08e-01   1.33e-01 -8.20e-01  0.4148
## Year2009      -1.32e-01   1.02e-01 -1.29e+00  0.1984
## Year2010      -1.75e-01   1.17e-01 -1.49e+00  0.1372
## Year2011      -1.76e-01   1.07e-01 -1.64e+00  0.1019
## Year2012      -8.20e-02   1.12e-01 -7.30e-01  0.4636
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.507
## Multiple R-squared:  0.0765, Adjusted R-squared:  -0.011
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 174 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.445  0.887  0.961  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      5.38e-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 3307"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1997 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    3    1    2    2    1    2    5    1    6    2    2    7    2    6
##
## 1997 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    2    1    2    1    1    1    4    1    4    2    1    6    2    4
##
## 1997 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    2    0    2    0    1    0    3    1    3    2    0    6    1    3
## [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: 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 = 6, 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: 2.1689435423954"

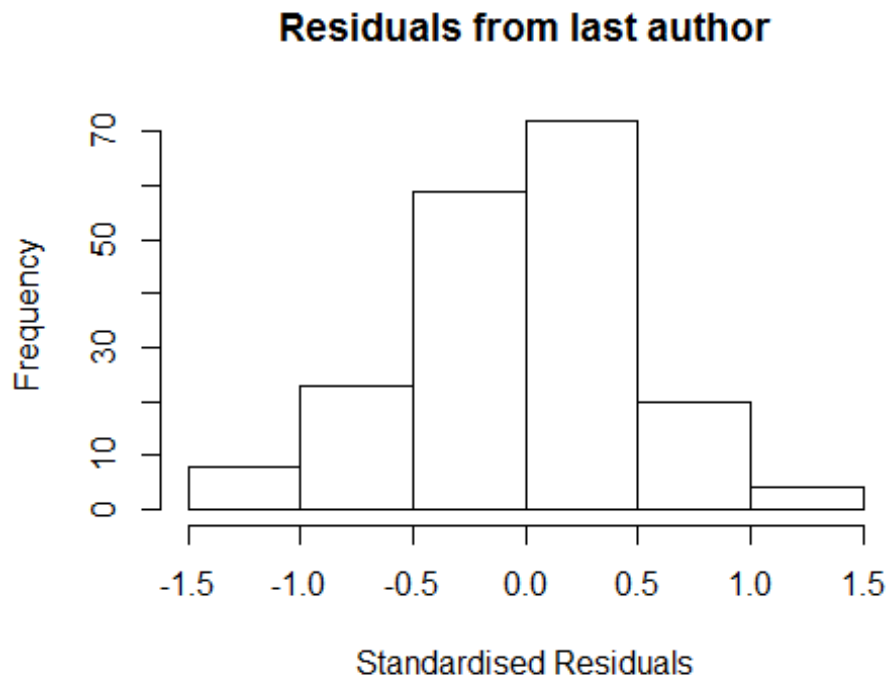
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4.5, 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"
## [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
```



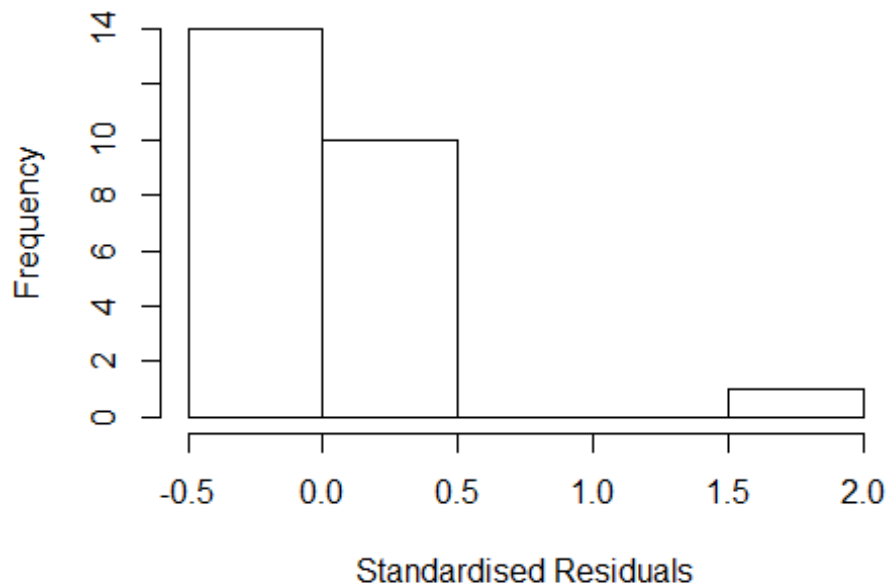
```
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1          NaN
## Year            NaN 10          NaN

## [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

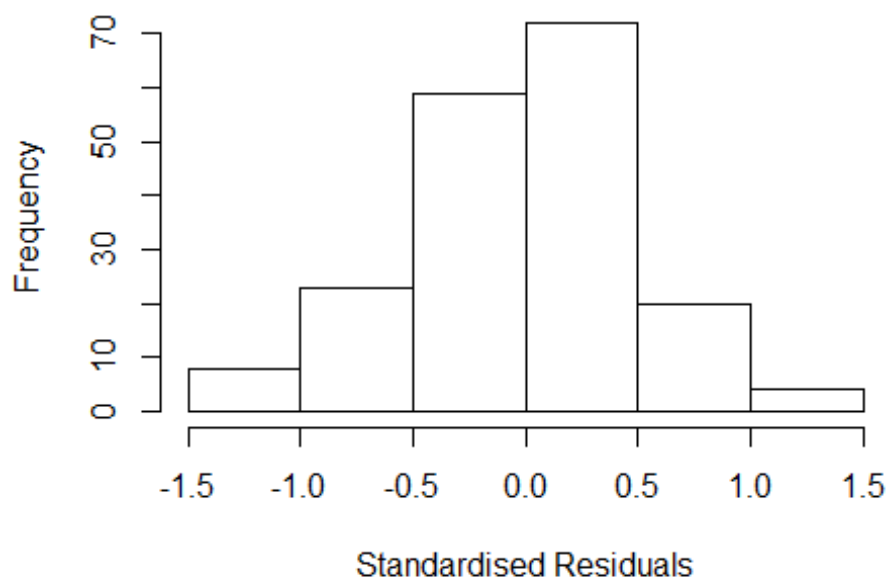
## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

Residuals from first author



| | | | |
|---------------------|------|----|-------------------------|
| ## | GVIF | Df | $GVIF^{1/(2 \cdot Df)}$ |
| ## LastAuthorFemale | NaN | 1 | NaN |
| ## Year | NaN | 10 | NaN |

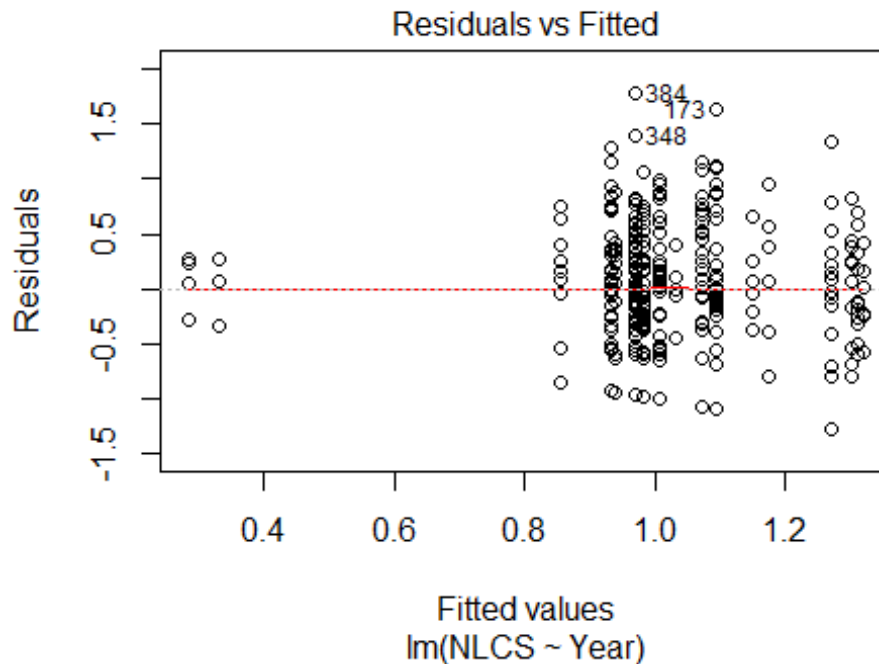
Residuals from last author



```

## [1] "Sample size for the above analysis: 25"
## [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
##    5   13    8    6   11   11   13    8   14   18   16   39   33   48   45
## 2011 2012
##   77   47
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3   10    5    5    7    7   10    7   12   14   14   34   30   44   37
## 2011 2012
##   68   37
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2   10    5    5    7    7    9    7   12   14   13   33   28   41   37
## 2011 2012
##   66   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 = 20, df = 16, p-value = 0.2

```

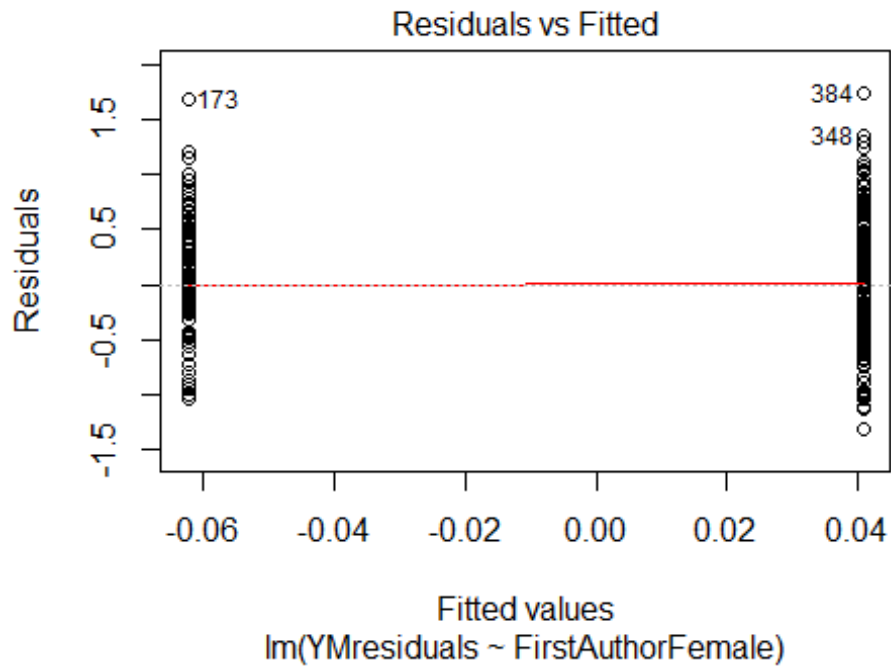
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.9e-08, df = 1, p-value = 1

## [1] "Female first author team size 2018 geometric mean: 1.19310427635346"
## [1] "Male first author team size 2018 geometric mean: 1.30798929674694"

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

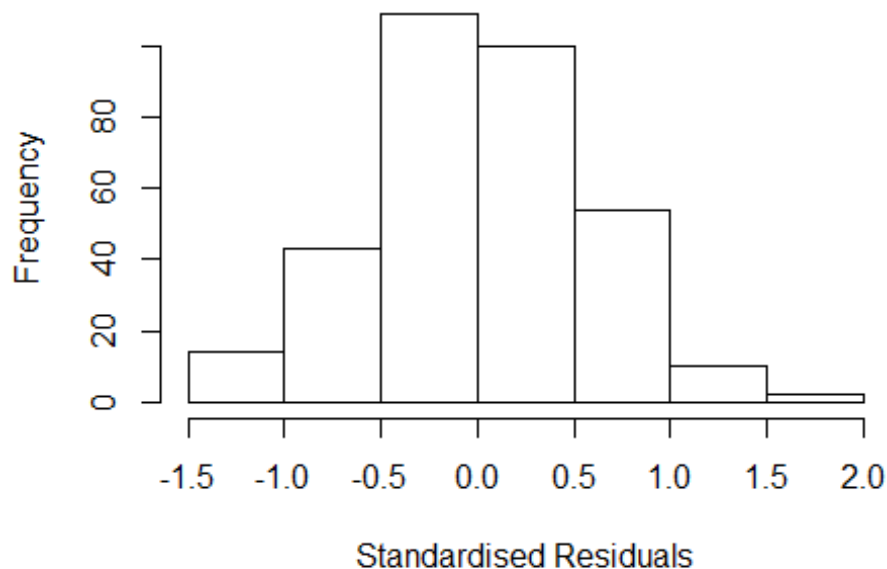
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 140, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.04427378242741"
## [1] "Male last author team size 2018 geometric mean: 1.44920251097144"

## Warning in wilcox.test.default(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 = 0.06
## 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 3.525 1          1.877
## LastAuthorFemale  3.882 1          1.970
## UniqueAuthors    4.503 4          1.207
## Year              7.267 16         1.064
```

Residuals from first and last author and team size



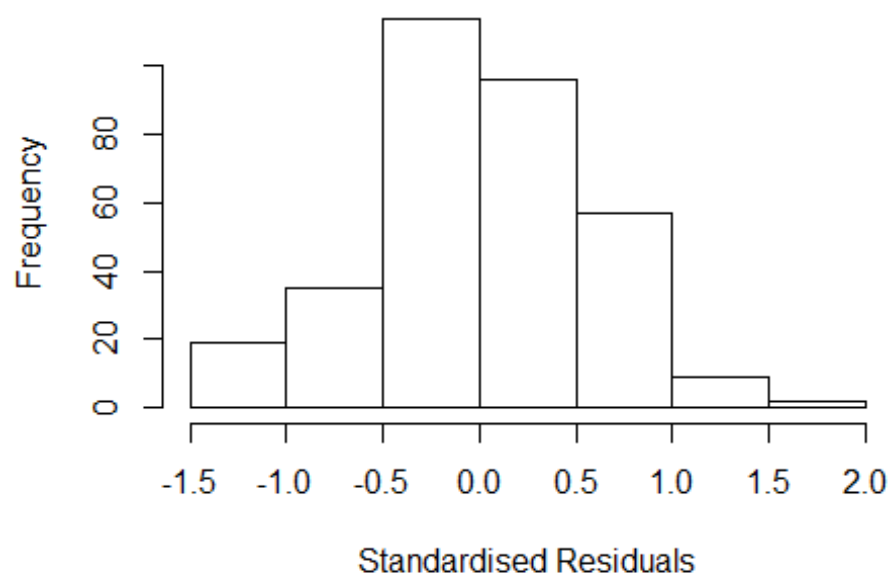
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.22727 -0.33951 0.00392 0.35427 1.68751
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.3374 0.1306 2.58 0.0103 *
## FirstAuthorFemale1 -0.0832 0.1121 -0.74 0.4582
## LastAuthorFemale1 -0.1080 0.1203 -0.90 0.3699
## UniqueAuthors2 0.2275 0.0850 2.68 0.0078 **
## UniqueAuthors3 0.0277 0.1242 0.22 0.8235
## UniqueAuthors4 0.1699 0.1494 1.14 0.2564
## UniqueAuthors5 -0.0277 0.2984 -0.09 0.9262
## Year1997 0.5402 0.2415 2.24 0.0260 *
## Year1998 -0.0554 0.1348 -0.41 0.6816
## Year1999 0.6486 0.2130 3.05 0.0025 **
```

```

## Year2000      0.9628      0.1602      6.01  5.2e-09 ***
## Year2001      0.7904      0.3456      2.29  0.0229 *
## Year2002      0.9600      0.2406      3.99  8.3e-05 ***
## Year2003      0.7805      0.1751      4.46  1.2e-05 ***
## Year2004      0.9535      0.1755      5.43  1.1e-07 ***
## Year2005      0.8899      0.2238      3.98  8.7e-05 ***
## Year2006      0.6574      0.1612      4.08  5.8e-05 ***
## Year2007      0.7214      0.1782      4.05  6.5e-05 ***
## Year2008      0.8581      0.1803      4.76  3.0e-06 ***
## Year2009      0.6224      0.1403      4.44  1.3e-05 ***
## Year2010      0.6326      0.1589      3.98  8.6e-05 ***
## Year2011      0.6502      0.1353      4.81  2.4e-06 ***
## Year2012      0.7129      0.1527      4.67  4.6e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.545
## Multiple R-squared:  0.119, Adjusted R-squared:  0.0558
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 33 weights are ~= 1. The remaining 299 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.318  0.860  0.953  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      3.01e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500      50      2      1      1000      200
## trace.lev      mts      compute.rd
##      0      1000      0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.385 1      1.840
## LastAuthorFemale  3.777 1      1.943
## Year              2.292 16      1.026

```

Residuals from first and last author



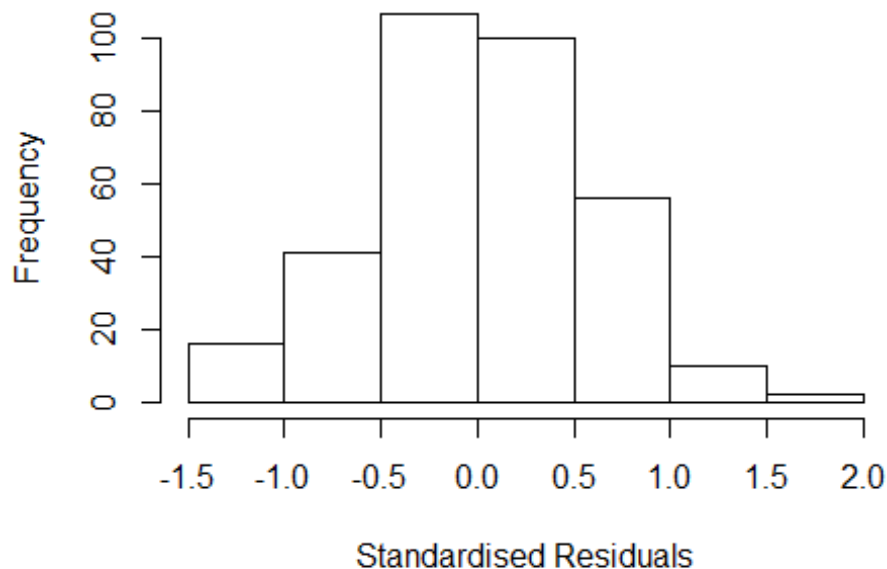
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.31151 -0.36590 -0.00236 0.37539 1.90920
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.49745 0.09098 5.47 9.3e-08 ***
## FirstAuthorFemale1 -0.00591 0.11258 -0.05 0.95820
## LastAuthorFemale1 -0.19609 0.11884 -1.65 0.09994 .
## Year1997 0.47051 0.22483 2.09 0.03718 *
## Year1998 -0.17333 0.13008 -1.33 0.18366
## Year1999 0.53779 0.15251 3.53 0.00048 ***
## Year2000 0.85183 0.14337 5.94 7.5e-09 ***
## Year2001 0.68260 0.33423 2.04 0.04196 *
## Year2002 0.84570 0.22936 3.69 0.00027 ***
## Year2003 0.70599 0.15873 4.45 1.2e-05 ***
## Year2004 0.82646 0.15244 5.42 1.2e-07 ***
## Year2005 0.81406 0.19583 4.16 4.2e-05 ***
```

```

## Year2006          0.54773      0.15435      3.55  0.00045 ***
## Year2007          0.59972      0.15105      3.97  8.9e-05 ***
## Year2008          0.74140      0.16861      4.40  1.5e-05 ***
## Year2009          0.49948      0.10997      4.54  8.0e-06 ***
## Year2010          0.52503      0.13815      3.80  0.00017 ***
## Year2011          0.54043      0.12127      4.46  1.2e-05 ***
## Year2012          0.58242      0.13240      4.40  1.5e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.547
## Multiple R-squared:  0.0997, Adjusted R-squared:  0.048
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 35 weights are ~= 1. The remaining 297 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.198  0.854  0.951  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      3.01e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.327 1      1.152
## Year              1.327 16      1.009

```

Residuals from first author



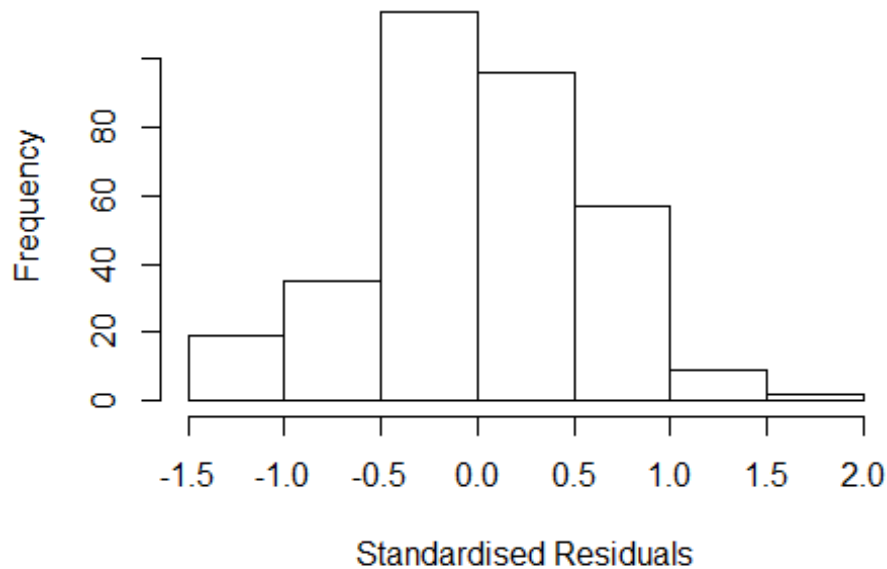
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.33727 -0.35844 0.00884 0.37116 1.78856
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5684 0.1284 4.43 1.3e-05 ***
## FirstAuthorFemale1 -0.1479 0.0713 -2.07 0.03893 *
## Year1997 0.3386 0.2235 1.51 0.13088
## Year1998 -0.2554 0.1587 -1.61 0.10853
## Year1999 0.4668 0.1773 2.63 0.00891 **
## Year2000 0.7966 0.1882 4.23 3.0e-05 ***
## Year2001 0.6117 0.3405 1.80 0.07342 .
## Year2002 0.7658 0.2488 3.08 0.00227 **
## Year2003 0.6323 0.1901 3.33 0.00098 ***
## Year2004 0.7652 0.1779 4.30 2.3e-05 ***
## Year2005 0.7688 0.2106 3.65 0.00031 ***
## Year2006 0.4634 0.1813 2.56 0.01105 *
```

```

## Year2007          0.5159      0.1723      2.99  0.00297 **
## Year2008          0.6572      0.1919      3.42  0.00070 ***
## Year2009          0.4367      0.1453      3.00  0.00287 **
## Year2010          0.4294      0.1607      2.67  0.00794 **
## Year2011          0.4450      0.1420      3.13  0.00190 **
## Year2012          0.5083      0.1636      3.11  0.00206 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.554
## Multiple R-squared:  0.0892, Adjusted R-squared:  0.0399
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 34 weights are ~= 1. The remaining 298 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.276  0.852  0.947  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      3.01e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.486 1      1.219
## Year              1.486 16      1.012

```


Residuals from last author



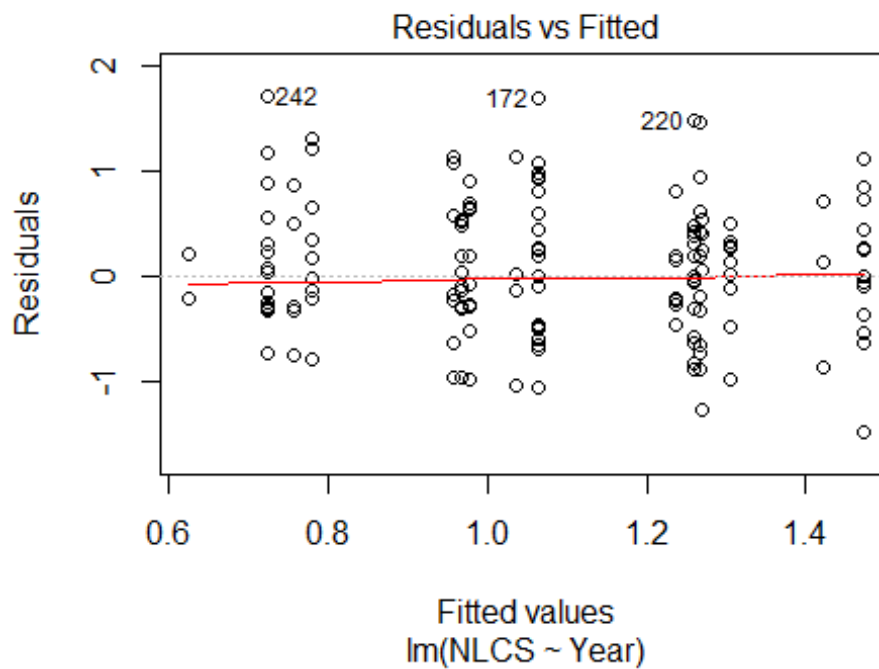
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.30998 -0.36518 -0.00194 0.37582 1.91483
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.4945 0.0693 7.13 6.9e-12 ***
## LastAuthorFemale1 -0.2011 0.0747 -2.69 0.00750 **
## Year1997 0.4749 0.2052 2.31 0.02130 *
## Year1998 -0.1705 0.1181 -1.44 0.14974
## Year1999 0.5408 0.1407 3.84 0.00015 ***
## Year2000 0.8538 0.1370 6.23 1.5e-09 ***
## Year2001 0.6853 0.3275 2.09 0.03721 *
## Year2002 0.8480 0.2206 3.84 0.00015 ***
## Year2003 0.7080 0.1496 4.73 3.3e-06 ***
## Year2004 0.8287 0.1469 5.64 3.8e-08 ***
## Year2005 0.8155 0.1957 4.17 4.0e-05 ***
## Year2006 0.5508 0.1462 3.77 0.00020 ***
```

```

## Year2007          0.6019      0.1372      4.39  1.6e-05 ***
## Year2008          0.7440      0.1630      4.56  7.2e-06 ***
## Year2009          0.5013      0.1041      4.81  2.3e-06 ***
## Year2010          0.5271      0.1320      3.99  8.1e-05 ***
## Year2011          0.5428      0.1096      4.95  1.2e-06 ***
## Year2012          0.5845      0.1284      4.55  7.6e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.544
## Multiple R-squared:  0.1, Adjusted R-squared:  0.0514
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 35 weights are ~= 1. The remaining 297 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.190  0.852  0.951  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      3.01e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 332"
## [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
##    3    7    5    3   11    6    6   10   11    8   15   10   16   30   19
## 2011 2012
##   17   22
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    5    4    2   10    4    5   10   11    7   14    9   11   26   14
## 2011 2012

```

```
## 14 18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2 5 4 2 10 4 5 9 11 6 14 9 10 23 13
## 2011 2012
## 11 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 = 9.9, df = 16, p-value = 0.9
```



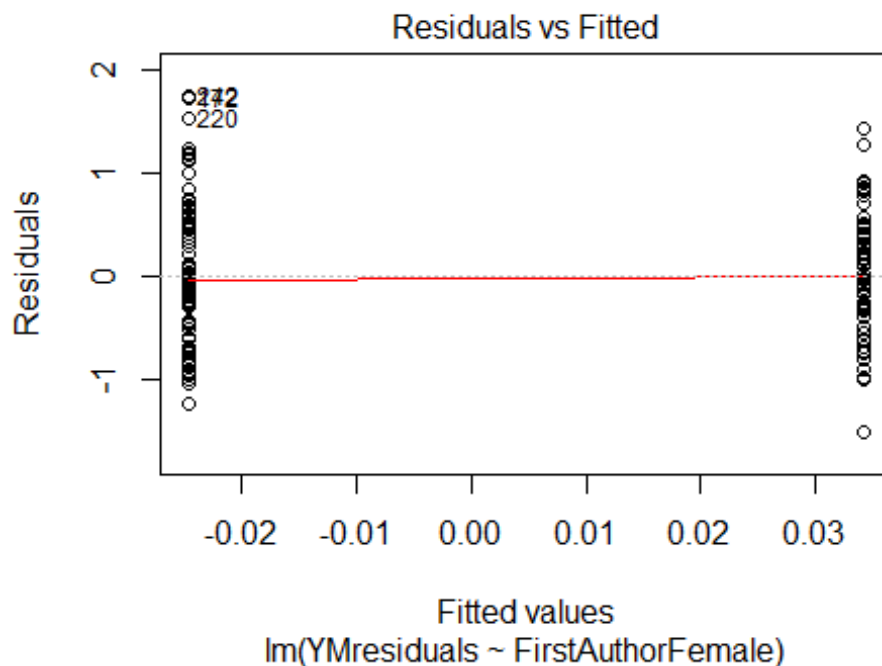
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.6, df = 1, p-value = 0.1

## [1] "Female first author team size 2018 geometric mean: 1.5130857494229"
## [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
```

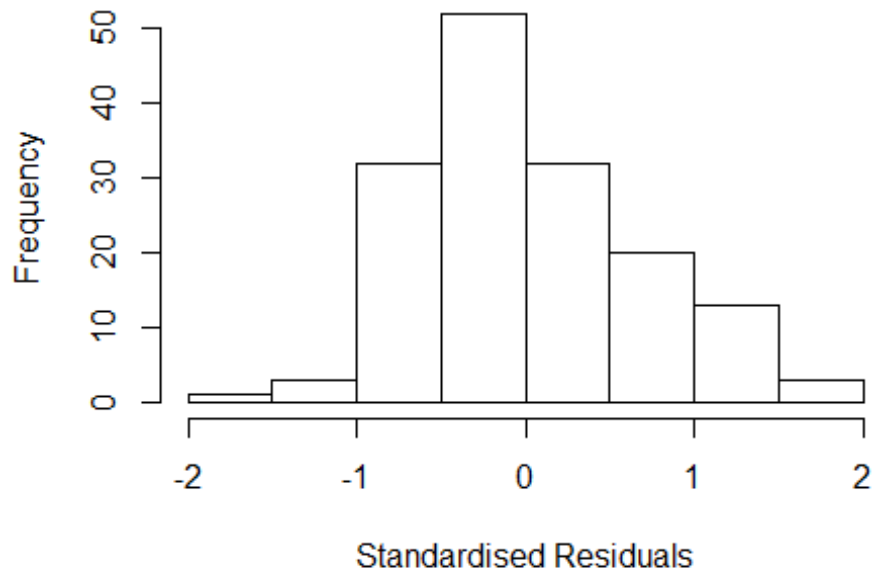
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 14, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.62238960361098"
## [1] "Male last 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 = 16, 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 2.084 1      1.444
## LastAuthorFemale  1.902 1      1.379
## UniqueAuthors    8.489 4      1.306
## Year             10.853 16      1.077
```

Residuals from first and last author and team size



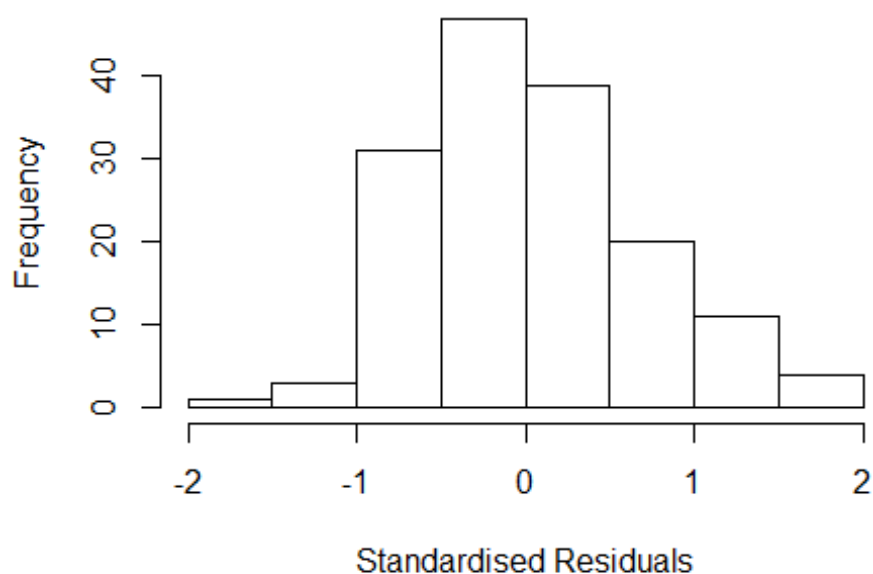
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.550 -0.442 -0.104 0.420 1.824
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.858 0.222 8.38 6.7e-14 ***
## FirstAuthorFemale1 0.196 0.137 1.44 0.15361
## LastAuthorFemale1 -0.213 0.131 -1.63 0.10638
## UniqueAuthors2 0.126 0.142 0.88 0.37827
## UniqueAuthors3 0.228 0.204 1.12 0.26611
## UniqueAuthors4 -0.017 0.408 -0.04 0.96683
## UniqueAuthors5 0.320 0.273 1.17 0.24405
## Year1997 -1.128 0.353 -3.19 0.00175 **
## Year1998 -0.934 0.527 -1.77 0.07886 .
## Year1999 -1.226 0.263 -4.67 7.4e-06 ***
```

```

## Year2000          -0.945      0.398    -2.38  0.01896 *
## Year2001          -1.239      0.328    -3.78  0.00024 ***
## Year2002          -0.595      0.379    -1.57  0.11871
## Year2003          -1.014      0.263    -3.86  0.00018 ***
## Year2004          -0.979      0.313    -3.13  0.00216 **
## Year2005          -0.783      0.303    -2.58  0.01086 *
## Year2006          -0.417      0.297    -1.41  0.16218
## Year2007          -0.642      0.285    -2.25  0.02583 *
## Year2008          -0.778      0.337    -2.31  0.02228 *
## Year2009          -0.917      0.300    -3.06  0.00266 **
## Year2010          -1.290      0.287    -4.50  1.5e-05 ***
## Year2011          -0.713      0.297    -2.40  0.01761 *
## Year2012          -1.246      0.264    -4.73  5.7e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.66
## Multiple R-squared:  0.183, Adjusted R-squared:  0.0479
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 145 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.426  0.877  0.952  0.912  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      6.41e-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 1.680 1      1.296
## LastAuthorFemale  1.528 1      1.236
## Year              1.737 16      1.017

```

Residuals from first and last author



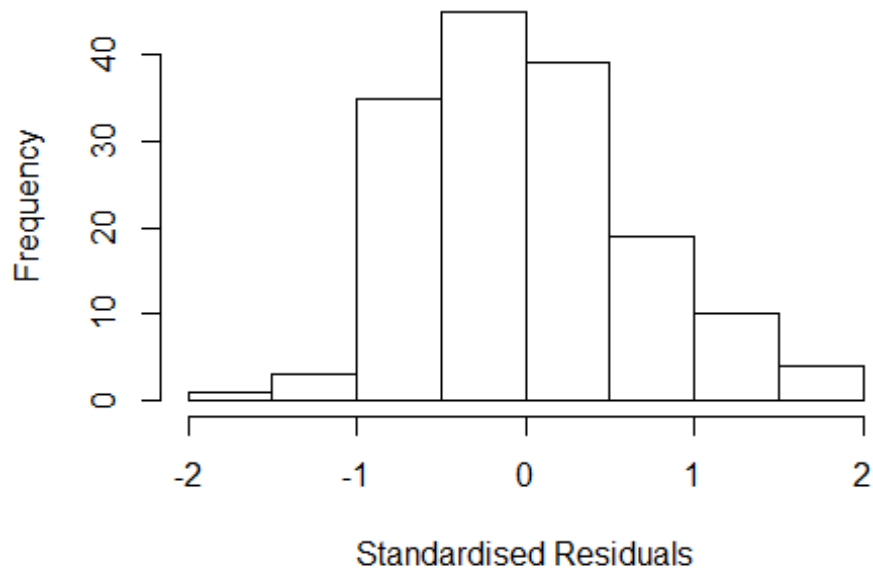
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5269 -0.4149 -0.0528 0.4131 1.7699
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.870 0.212 8.80 5.1e-15 ***
## FirstAuthorFemale1 0.207 0.125 1.66 0.09992 .
## LastAuthorFemale1 -0.247 0.119 -2.09 0.03874 *
## Year1997 -1.114 0.373 -2.98 0.00336 **
## Year1998 -0.861 0.494 -1.74 0.08355 .
## Year1999 -1.225 0.261 -4.70 6.2e-06 ***
## Year2000 -0.937 0.415 -2.26 0.02538 *
## Year2001 -1.182 0.318 -3.72 0.00029 ***
## Year2002 -0.482 0.352 -1.37 0.17405
## Year2003 -0.991 0.249 -3.98 0.00011 ***
## Year2004 -0.913 0.305 -2.99 0.00329 **
## Year2005 -0.650 0.258 -2.52 0.01291 *
```

```

## Year2006          -0.302      0.259   -1.17   0.24597
## Year2007          -0.547      0.248   -2.20   0.02928 *
## Year2008          -0.702      0.328   -2.14   0.03427 *
## Year2009          -0.876      0.290   -3.02   0.00299 **
## Year2010          -1.208      0.257   -4.70   6.3e-06 ***
## Year2011          -0.607      0.270   -2.25   0.02627 *
## Year2012          -1.184      0.252   -4.70   6.2e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.659
## Multiple R-squared:  0.174, Adjusted R-squared:  0.0654
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 142 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.451  0.881  0.951  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      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 1.388 1      1.178
## Year              1.388 16      1.010

```


Residuals from first author



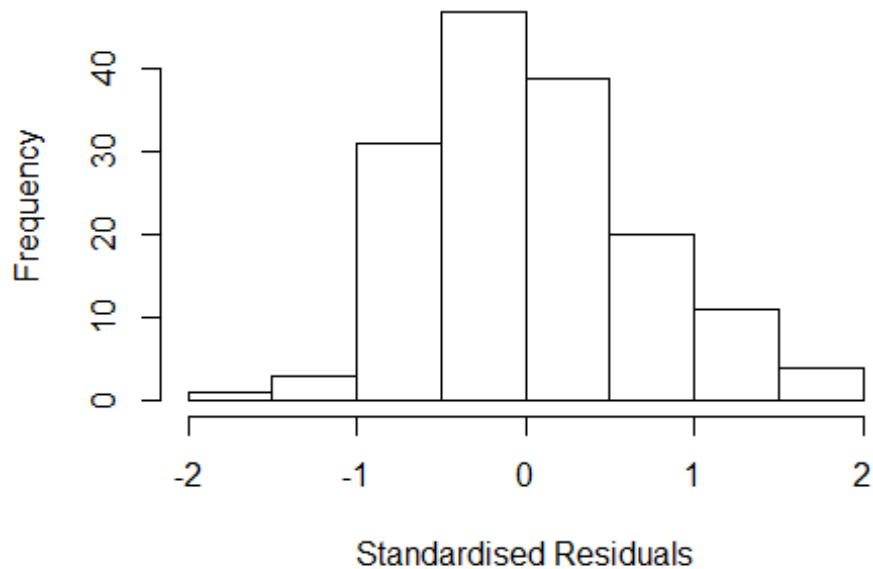
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5325 -0.4766 -0.0548 0.4646 1.8121
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.8208 0.2471 7.37 1.4e-11 ***
## FirstAuthorFemale1 0.0574 0.1165 0.49 0.6228
## Year1997 -1.0983 0.3696 -2.97 0.0035 **
## Year1998 -0.8102 0.5066 -1.60 0.1121
## Year1999 -1.2255 0.2724 -4.50 1.4e-05 ***
## Year2000 -0.9180 0.4366 -2.10 0.0373 *
## Year2001 -1.1193 0.3536 -3.17 0.0019 **
## Year2002 -0.5096 0.3687 -1.38 0.1691
## Year2003 -0.9361 0.2826 -3.31 0.0012 **
## Year2004 -0.8670 0.3284 -2.64 0.0092 **
## Year2005 -0.6628 0.2938 -2.26 0.0256 *
## Year2006 -0.3457 0.3034 -1.14 0.2564
```

```

## Year2007          -0.5297      0.2788   -1.90    0.0595 .
## Year2008          -0.6632      0.3808   -1.74    0.0838 .
## Year2009          -0.8507      0.3119   -2.73    0.0072 **
## Year2010          -1.1877      0.2918   -4.07    7.9e-05 ***
## Year2011          -0.6067      0.2983   -2.03    0.0439 *
## Year2012          -1.1959      0.2823   -4.24    4.1e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.677
## Multiple R-squared:  0.155, Adjusted R-squared:  0.0506
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 142 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.454  0.872  0.943  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      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 1.319 1          1.148
## Year            1.319 16          1.009

```

Residuals from last author



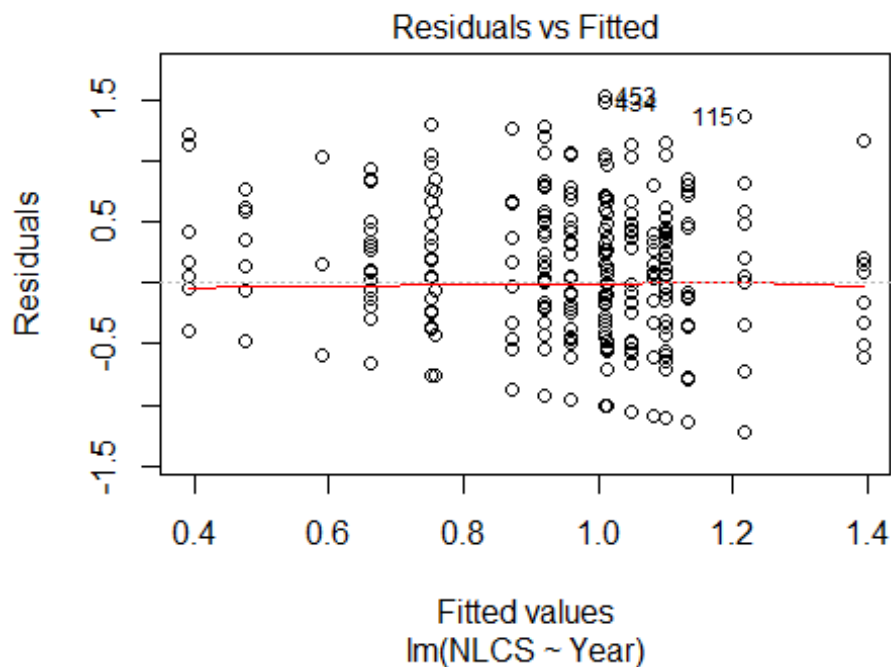
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4530 -0.4657 -0.0648 0.4306 1.7182
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.912 0.179 10.70 < 2e-16 ***
## LastAuthorFemale1 -0.125 0.114 -1.10 0.27528
## Year1997 -1.128 0.382 -2.95 0.00374 **
## Year1998 -0.904 0.481 -1.88 0.06251 .
## Year1999 -1.225 0.260 -4.71 6.0e-06 ***
## Year2000 -0.967 0.393 -2.46 0.01522 *
## Year2001 -1.141 0.305 -3.74 0.00027 ***
## Year2002 -0.459 0.341 -1.35 0.18046
## Year2003 -0.956 0.216 -4.42 2.0e-05 ***
## Year2004 -0.882 0.274 -3.22 0.00161 **
## Year2005 -0.639 0.240 -2.66 0.00869 **
## Year2006 -0.350 0.241 -1.45 0.14905
```

```

## Year2007          -0.524      0.229   -2.29  0.02330 *
## Year2008          -0.702      0.335   -2.10  0.03777 *
## Year2009          -0.854      0.268   -3.19  0.00176 **
## Year2010          -1.219      0.234   -5.20  6.9e-07 ***
## Year2011          -0.632      0.244   -2.59  0.01066 *
## Year2012          -1.193      0.225   -5.31  4.3e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.655
## Multiple R-squared:  0.161, Adjusted R-squared:  0.0579
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 142 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.471  0.865   0.948   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      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 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
##    9   27   12   12    8   16   13   18   18   29   25   41   34   33   28
## 2011 2012
##   50   35
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4   17   11   11    8    9   12   11   15   23   21   34   27   29   21
## 2011 2012

```

```
## 40 35
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4 17 10 11 8 8 10 10 14 20 21 33 26 29 20
## 2011 2012
## 40 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 = 0.8
```



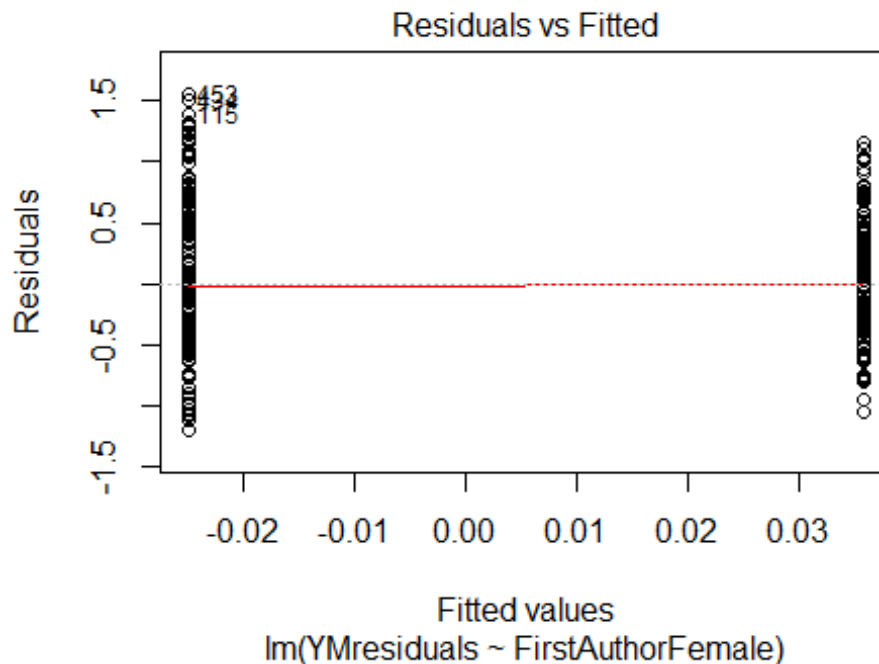
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 6.8, df = 1, p-value = 0.009

## [1] "Female first author team size 2018 geometric mean: 1.47087029034946"
## [1] "Male first author team size 2018 geometric mean: 1.53367468683856"

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

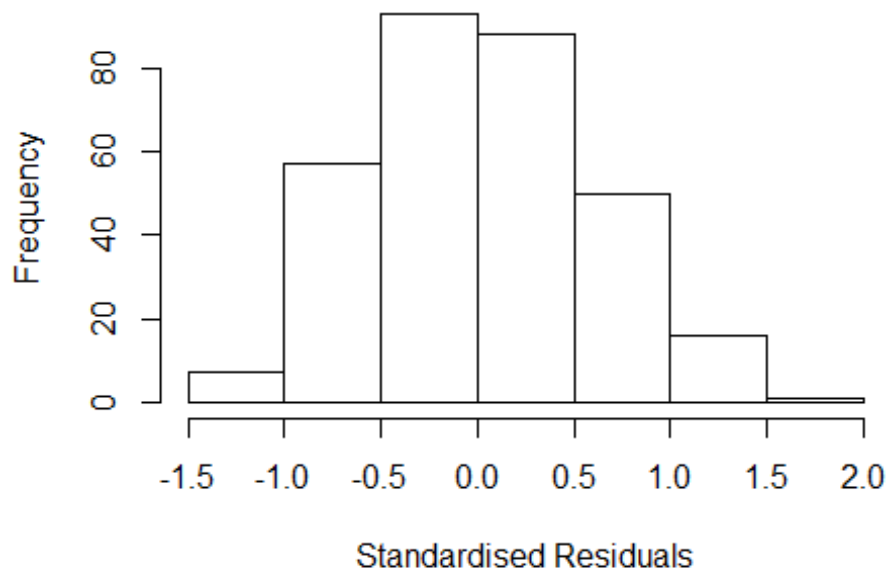
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 100, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.44950376037809"
## [1] "Male last author team size 2018 geometric mean: 1.57113958920376"

## Warning in wilcox.test.default(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 = 0.6
## 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 5.160  1      2.272
## LastAuthorFemale  5.288  1      2.300
## UniqueAuthors    2.189  4      1.103
## Year              2.418 16      1.028
```

Residuals from first and last author and team size



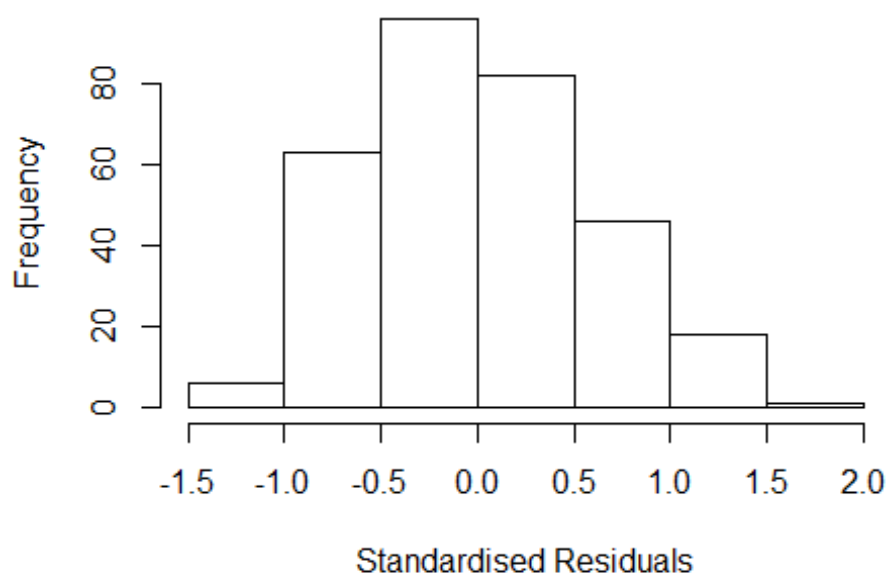
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.22142 -0.43170 -0.00078 0.44732 1.79703
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5380 0.3836 1.40 0.1619
## FirstAuthorFemale1 -0.1219 0.1499 -0.81 0.4167
## LastAuthorFemale1 0.2684 0.1528 1.76 0.0800 .
## UniqueAuthors2 0.2487 0.1064 2.34 0.0201 *
## UniqueAuthors3 0.3675 0.2011 1.83 0.0686 .
## UniqueAuthors4 0.7033 0.2238 3.14 0.0018 **
## UniqueAuthors5 0.0170 0.4194 0.04 0.9678
## Year1997 -0.2375 0.4035 -0.59 0.5567
## Year1998 0.1223 0.4301 0.28 0.7763
## Year1999 -0.1185 0.4071 -0.29 0.7712
```

```

## Year2000          0.6933      0.4179      1.66      0.0982 .
## Year2001          0.0114      0.4628      0.02      0.9803
## Year2002          0.4739      0.4225      1.12      0.2629
## Year2003          0.4777      0.4517      1.06      0.2912
## Year2004          0.4296      0.4222      1.02      0.3098
## Year2005          0.4200      0.4195      1.00      0.3175
## Year2006          0.3430      0.4017      0.85      0.3939
## Year2007          0.3968      0.4019      0.99      0.3244
## Year2008          0.0476      0.3999      0.12      0.9054
## Year2009          0.2676      0.4038      0.66      0.5081
## Year2010         -0.0158      0.3994     -0.04      0.9684
## Year2011          0.2161      0.4038      0.54      0.5929
## Year2012          0.1990      0.4061      0.49      0.6245
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.625
## Multiple R-squared:  0.165, Adjusted R-squared:  0.101
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 283 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.388  0.877  0.948  0.917  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      3.21e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 4.817 1      2.195
## LastAuthorFemale  4.793 1      2.189
## Year              1.445 16      1.012

```


Residuals from first and last author



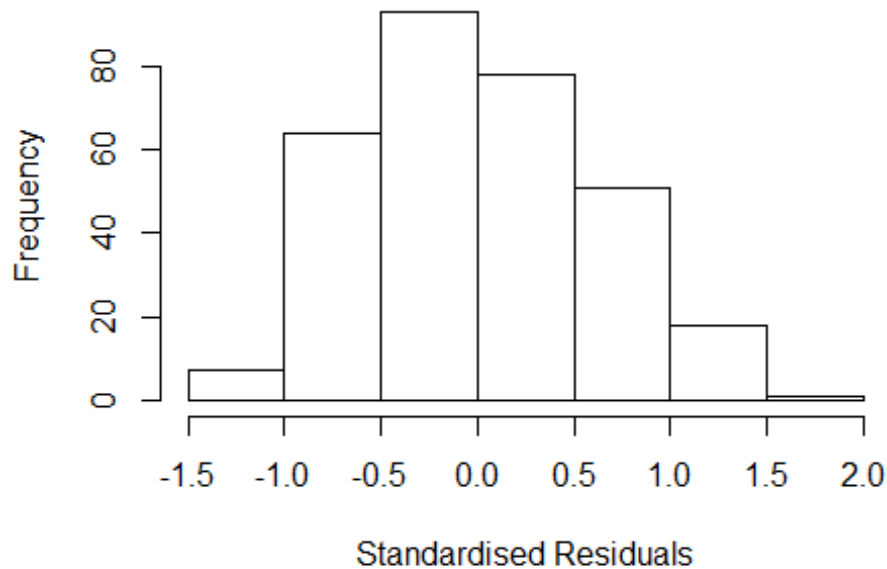
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0975 -0.4204 -0.0453 0.4315 1.7198
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5410 0.3808 1.42 0.156
## FirstAuthorFemale1 -0.0965 0.1483 -0.65 0.516
## LastAuthorFemale1 0.2528 0.1498 1.69 0.092 .
## Year1997 -0.2259 0.4005 -0.56 0.573
## Year1998 0.1527 0.4349 0.35 0.726
## Year1999 -0.1246 0.4039 -0.31 0.758
## Year2000 0.7241 0.4181 1.73 0.084 .
## Year2001 0.1020 0.4527 0.23 0.822
## Year2002 0.4658 0.4202 1.11 0.268
## Year2003 0.4893 0.4549 1.08 0.283
## Year2004 0.4401 0.4170 1.06 0.292
## Year2005 0.4383 0.4167 1.05 0.294
```

```

## Year2006          0.4002      0.3997      1.00      0.318
## Year2007          0.5069      0.3960      1.28      0.202
## Year2008          0.0802      0.3976      0.20      0.840
## Year2009          0.3259      0.3997      0.82      0.415
## Year2010          0.0977      0.3987      0.25      0.806
## Year2011          0.2704      0.4031      0.67      0.503
## Year2012          0.2732      0.4080      0.67      0.504
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.64
## Multiple R-squared:  0.129, Adjusted R-squared:  0.0759
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 290 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.450  0.881  0.951  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      3.21e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
##  nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
##  trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.242 1      1.114
## Year              1.242 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
## -1.0892 -0.4310 -0.0237 0.4433 1.7121
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5446 0.3763 1.45 0.149
## FirstAuthorFemale1 0.1165 0.0753 1.55 0.123
## Year1997 -0.2300 0.3980 -0.58 0.564
## Year1998 0.1929 0.4367 0.44 0.659
## Year1999 -0.1143 0.4018 -0.28 0.776
## Year2000 0.7468 0.4169 1.79 0.074 .
## Year2001 0.0543 0.4492 0.12 0.904
## Year2002 0.4787 0.4140 1.16 0.249
## Year2003 0.5048 0.4497 1.12 0.263
## Year2004 0.4601 0.4137 1.11 0.267
## Year2005 0.4267 0.4121 1.04 0.301
## Year2006 0.4282 0.3961 1.08 0.281
```

```

## Year2007          0.4953      0.3915      1.27      0.207
## Year2008          0.1083      0.3931      0.28      0.783
## Year2009          0.3454      0.3953      0.87      0.383
## Year2010          0.0997      0.3956      0.25      0.801
## Year2011          0.3008      0.3991      0.75      0.452
## Year2012          0.2774      0.4054      0.68      0.494
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.66
## Multiple R-squared:  0.119, Adjusted R-squared:  0.0677
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 289 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.481  0.883   0.949   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      3.21e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.2 1      1.095
## Year              1.2 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
## -1.112 -0.433 -0.038  0.427  1.731

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.5412    0.3806   1.42    0.156
## LastAuthorFemale1 0.1694    0.0753   2.25    0.025 *
## Year1997         -0.2359    0.4002  -0.59    0.556
## Year1998          0.1596    0.4361   0.37    0.715
## Year1999         -0.1293    0.4031  -0.32    0.749
## Year2000          0.7164    0.4172   1.72    0.087 .
## Year2001          0.0783    0.4510   0.17    0.862
## Year2002          0.4599    0.4204   1.09    0.275
## Year2003          0.4819    0.4542   1.06    0.290
## Year2004          0.4322    0.4162   1.04    0.300
## Year2005          0.4209    0.4147   1.01    0.311
## Year2006          0.4009    0.4002   1.00    0.317
## Year2007          0.4927    0.3948   1.25    0.213
## Year2008          0.0820    0.3973   0.21    0.837
## Year2009          0.3184    0.3993   0.80    0.426
## Year2010          0.0892    0.3983   0.22    0.823
## Year2011          0.2676    0.4030   0.66    0.507
## Year2012          0.2616    0.4067   0.64    0.521
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.641
## Multiple R-squared:  0.128, Adjusted R-squared:  0.0772
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 289 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.446  0.883   0.951   0.920   0.985   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00        5.00e-01      4.69e+00      1.00e-07
##      rel.tol        solve.tol      eps.outlier      eps.x
##      1.00e-07        1.00e-07      3.21e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01        5.00e-01
##   nResample    max.it    best.r.s    k.fast.s    k.max maxit.scale
##      500         50         2         1        1000         200
##   trace.lev    mts    compute.rd
##      0         1000         0
##           psi      subsampling      cov
##      "bisquare"    "nonsingular"    ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 312"
## [1] ""

```

```

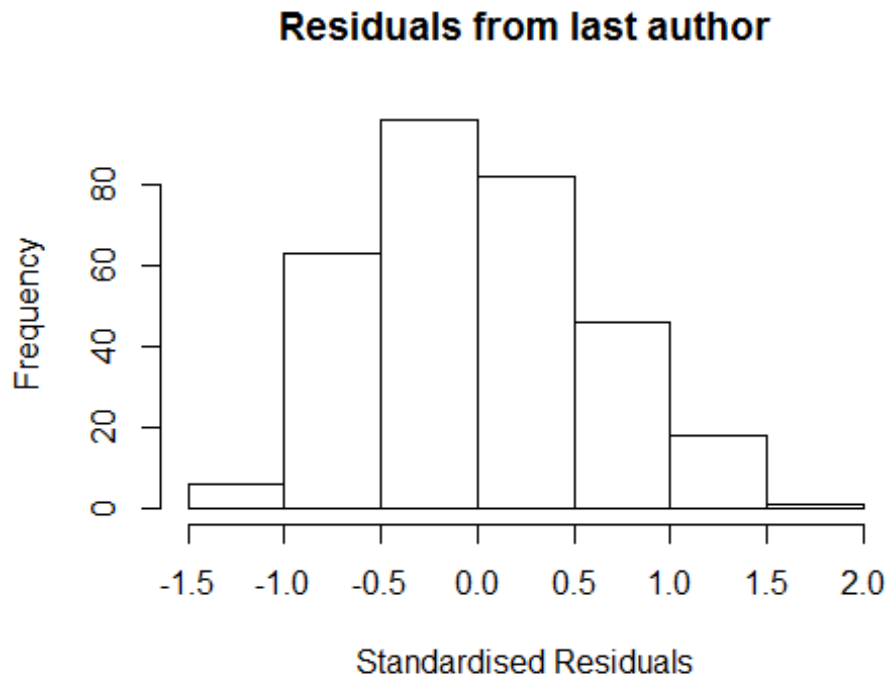
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 3311"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1998 2000 2002 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2    1    2    1    4    2    3    3    1    7    3    4
##
## 1998 2000 2002 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    0    1    1    4    0    2    3    1    3    2    3
##
## 1998 2000 2002 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    0    1    1    3    0    2    1    1    3    2    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: 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 = 0.5, p-value = NA
## 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"

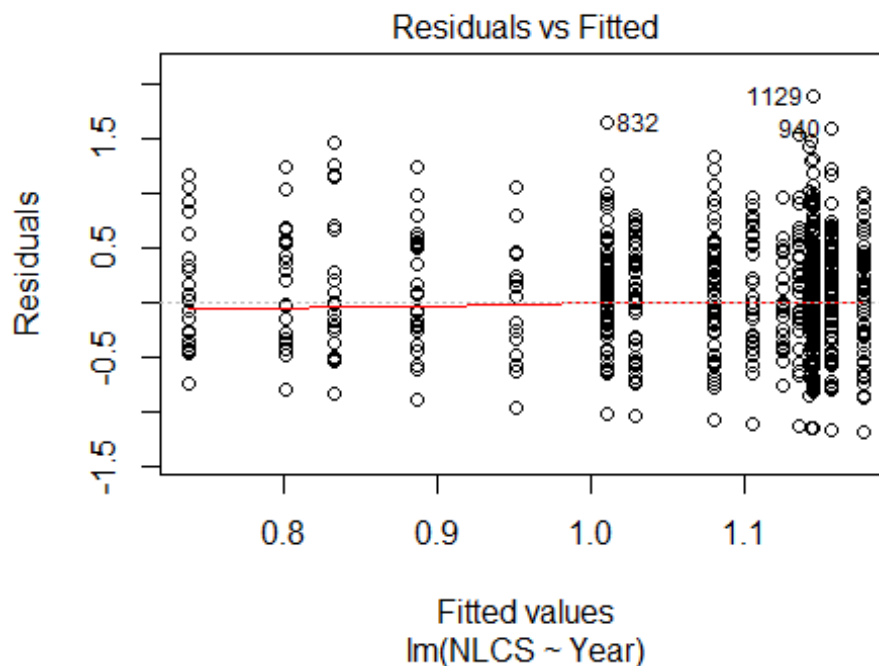
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```

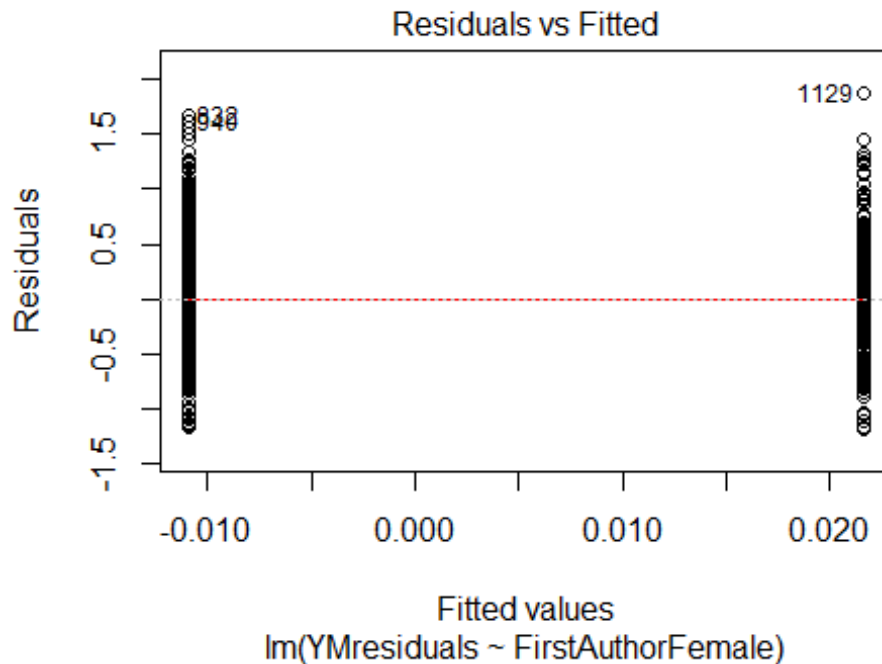


```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0.5, p-value = NA
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
## [1] "Regression 2: 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: 18"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 3312"
## [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   34   32   30   35   31   39   38   44   40   54   66   76   98  127
## 2011 2012
##    130  103
##
```

```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 15 29 26 29 20 25 33 34 35 35 46 56 68 91 112
## 2011 2012
## 107 92
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 14 29 26 29 19 23 32 34 34 31 44 51 65 87 103
## 2011 2012
## 102 86
## [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
```

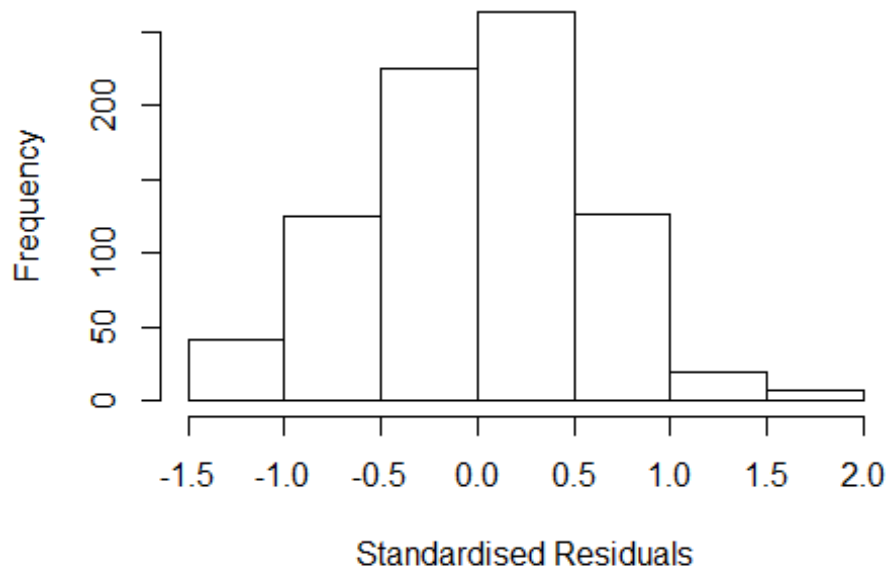


```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.4, df = 1, p-value = 0.2
```

```
## [1] "Female first author team size 2018 geometric mean: 1.53851265645705"
## [1] "Male first author team size 2018 geometric mean: 1.2908789596097"
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2000, p-value = 0.06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.49997062419598"
## [1] "Male last author team size 2018 geometric mean: 1.31809286908993"
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1900, 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 2.664 1      1.632
## LastAuthorFemale  2.544 1      1.595
## UniqueAuthors    1.507 4      1.053
## Year             1.725 16      1.017
```

Residuals from first and last author and team size



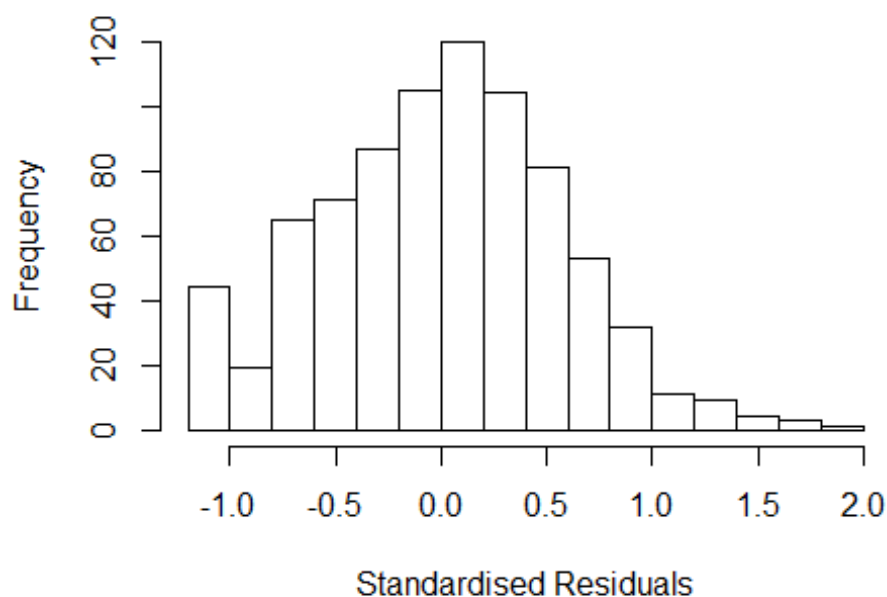
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.241 -0.395 0.020 0.382 1.885
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9017 0.1636 5.51 4.8e-08 ***
## FirstAuthorFemale1 0.0182 0.0677 0.27 0.788
## LastAuthorFemale1 0.0176 0.0664 0.27 0.791
## UniqueAuthors2 0.1072 0.0530 2.02 0.043 *
## UniqueAuthors3 0.0260 0.0738 0.35 0.725
## UniqueAuthors4 0.0650 0.1181 0.55 0.582
## UniqueAuthors5 -0.0359 0.7889 -0.05 0.964
## Year1997 -0.1687 0.2003 -0.84 0.400
## Year1998 -0.2194 0.2018 -1.09 0.277
## Year1999 -0.1416 0.2041 -0.69 0.488
```

```

## Year2000      0.1346      0.2150      0.63      0.532
## Year2001      0.1913      0.1875      1.02      0.308
## Year2002     -0.0770      0.1949     -0.39      0.693
## Year2003      0.1983      0.2022      0.98      0.327
## Year2004      0.1629      0.1980      0.82      0.411
## Year2005      0.1988      0.2080      0.96      0.339
## Year2006      0.1066      0.1827      0.58      0.560
## Year2007      0.2657      0.1850      1.44      0.151
## Year2008      0.1292      0.1825      0.71      0.479
## Year2009      0.2326      0.1743      1.33      0.182
## Year2010      0.0592      0.1713      0.35      0.730
## Year2011      0.1896      0.1727      1.10      0.273
## Year2012      0.2129      0.1771      1.20      0.230
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.575
## Multiple R-squared:  0.0545, Adjusted R-squared:  0.028
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 70 weights are ~= 1. The remaining 739 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.261  0.873  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      1.24e-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 2.576 1      1.605
## LastAuthorFemale 2.397 1      1.548
## Year      1.251 16      1.007

```

Residuals from first and last author



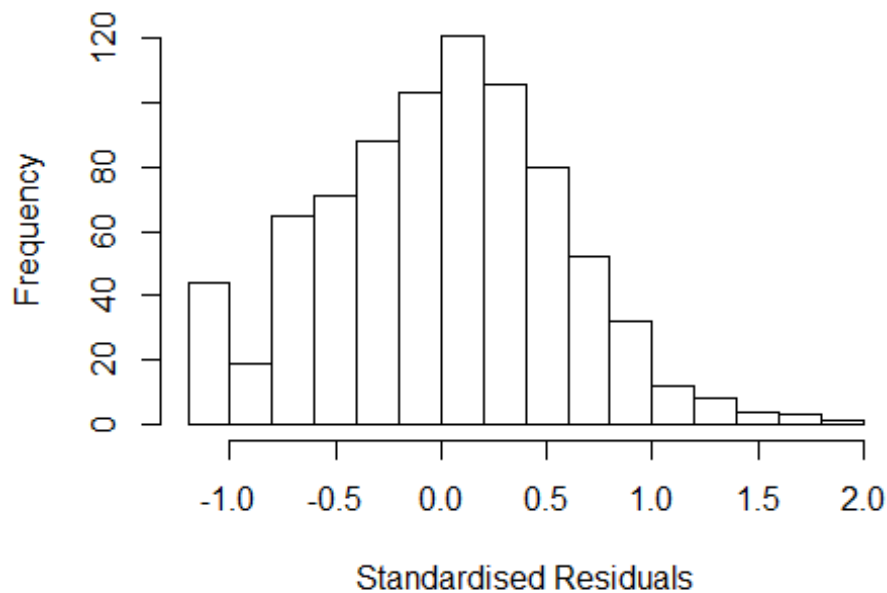
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1923 -0.3957 0.0233 0.3879 1.8601
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9197 0.1603 5.74 1.4e-08 ***
## FirstAuthorFemale1 0.0224 0.0666 0.34 0.74
## LastAuthorFemale1 0.0156 0.0646 0.24 0.81
## Year1997 -0.1637 0.1985 -0.82 0.41
## Year1998 -0.2205 0.1990 -1.11 0.27
## Year1999 -0.1457 0.2016 -0.72 0.47
## Year2000 0.1460 0.2113 0.69 0.49
## Year2001 0.2102 0.1839 1.14 0.25
## Year2002 -0.0703 0.1921 -0.37 0.71
## Year2003 0.1996 0.1975 1.01 0.31
## Year2004 0.1794 0.1943 0.92 0.36
## Year2005 0.2000 0.2055 0.97 0.33
```

```

## Year2006          0.1005      0.1805      0.56      0.58
## Year2007          0.2681      0.1830      1.47      0.14
## Year2008          0.1208      0.1802      0.67      0.50
## Year2009          0.2346      0.1714      1.37      0.17
## Year2010          0.0601      0.1684      0.36      0.72
## Year2011          0.1995      0.1701      1.17      0.24
## Year2012          0.2172      0.1741      1.25      0.21
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.573
## Multiple R-squared:  0.0502, Adjusted R-squared:  0.0286
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 70 weights are ~= 1. The remaining 739 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.271  0.873   0.950   0.908   0.985   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      1.24e-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 1.184 1      1.088
## Year              1.184 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
## -1.1893 -0.3968 0.0197 0.3859 1.8621
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9213 0.1601 5.76 1.2e-08 ***
## FirstAuthorFemale1 0.0346 0.0452 0.76 0.45
## Year1997 -0.1636 0.1989 -0.82 0.41
## Year1998 -0.2211 0.1990 -1.11 0.27
## Year1999 -0.1456 0.2019 -0.72 0.47
## Year2000 0.1449 0.2115 0.68 0.49
## Year2001 0.2081 0.1837 1.13 0.26
## Year2002 -0.0702 0.1924 -0.36 0.72
## Year2003 0.1994 0.1977 1.01 0.31
## Year2004 0.1791 0.1947 0.92 0.36
## Year2005 0.1995 0.2059 0.97 0.33
## Year2006 0.0997 0.1806 0.55 0.58
```

```

## Year2007          0.2671      0.1831      1.46      0.14
## Year2008          0.1206      0.1804      0.67      0.50
## Year2009          0.2334      0.1712      1.36      0.17
## Year2010          0.0589      0.1685      0.35      0.73
## Year2011          0.1993      0.1703      1.17      0.24
## Year2012          0.2170      0.1743      1.24      0.21
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.573
## Multiple R-squared:  0.0502, Adjusted R-squared:  0.0298
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 71 weights are ~= 1. The remaining 738 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.268  0.872  0.949  0.907  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      1.24e-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 1.111 1      1.054
## Year      1.111 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
## -1.1903 -0.3961  0.0249  0.3884  1.8628

```

```

##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9184    0.1602    5.73 1.4e-08 ***
## LastAuthorFemale1 0.0330    0.0441    0.75  0.45
## Year1997       -0.1635    0.1986   -0.82  0.41
## Year1998       -0.2193    0.1991   -1.10  0.27
## Year1999       -0.1437    0.2015   -0.71  0.48
## Year2000        0.1506    0.2101    0.72  0.47
## Year2001        0.2146    0.1830    1.17  0.24
## Year2002       -0.0698    0.1922   -0.36  0.72
## Year2003        0.2039    0.1967    1.04  0.30
## Year2004        0.1831    0.1931    0.95  0.34
## Year2005        0.2043    0.2045    1.00  0.32
## Year2006        0.1031    0.1802    0.57  0.57
## Year2007        0.2719    0.1823    1.49  0.14
## Year2008        0.1242    0.1794    0.69  0.49
## Year2009        0.2384    0.1705    1.40  0.16
## Year2010        0.0637    0.1681    0.38  0.71
## Year2011        0.2024    0.1696    1.19  0.23
## Year2012        0.2208    0.1731    1.28  0.20
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.573
## Multiple R-squared:  0.0503, Adjusted R-squared:  0.0299
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 72 weights are ~= 1. The remaining 737 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.268  0.872  0.950  0.907  0.984  0.999
## Algorithmic parameters:
##           tuning.chi           bb           tuning.psi           refine.tol
##           1.55e+00           5.00e-01           4.69e+00           1.00e-07
##           rel.tol           solve.tol           eps.outlier           eps.x
##           1.00e-07           1.00e-07           1.24e-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: 809"
## [1] ""

```



```

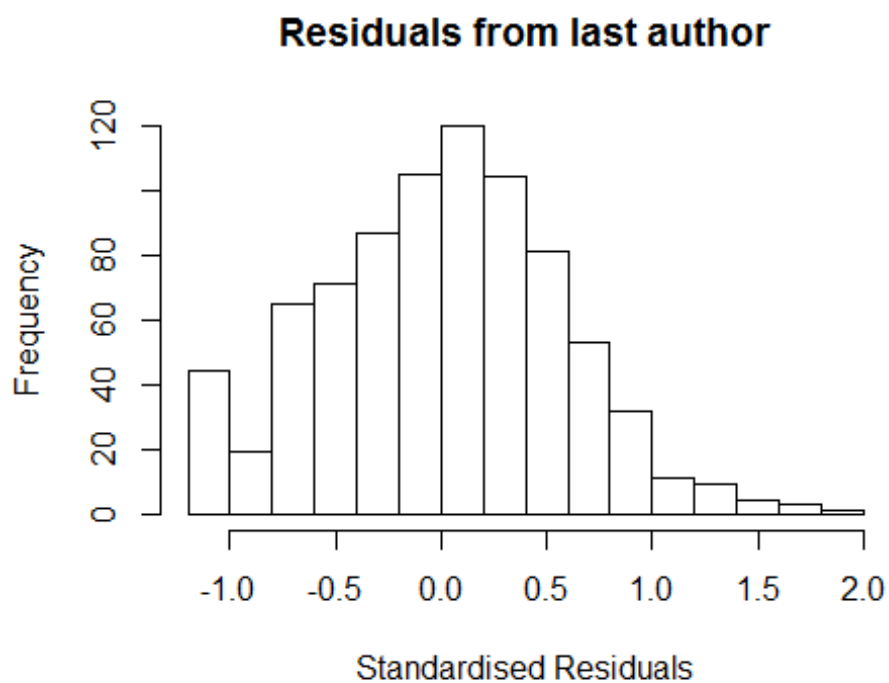
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 3313"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    3    3    1    8    5    2    4    3    2    9    7   10   13    7   11
## 2012
##    12
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    2    3    1    4    3    0    2    2    0    6    3    8    8    6    6
## 2012
##    8
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    2    2    1    4    3    0    2    2    0    6    2    7    7    6    6
## 2012
##    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: 1.64375182951723"

## Warning in wilcox.test.default(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: 1.81712059283214"
## [1] "Male last author team size 2018 geometric mean: 1.41421356237309"

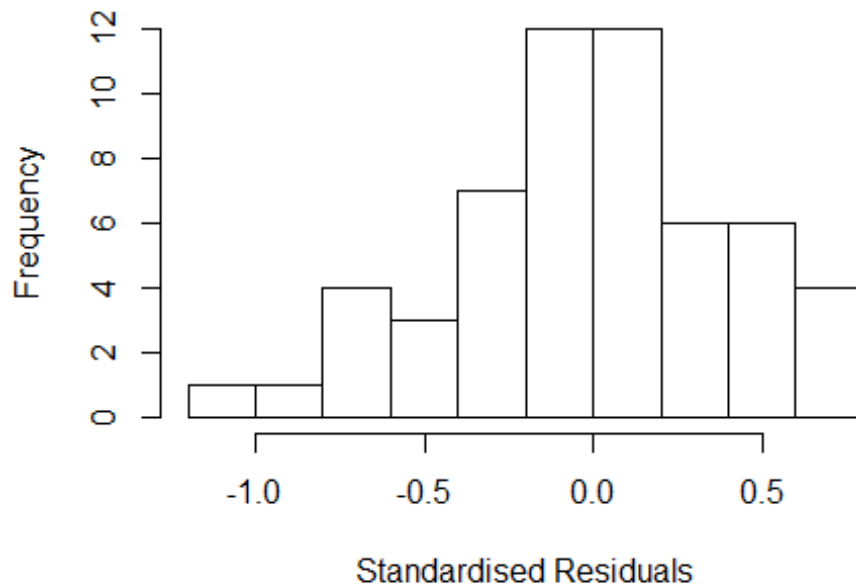
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 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"
##               GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -1.037e+01  1          NaN
## LastAuthorFemale  5.831e+00  1          2.415
## UniqueAuthors    1.625e+18  4          188.957
## Year              4.064e+18 13          5.197
```

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
## -1.05826 -0.23821 0.00286 0.24234 0.68398
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8537 0.7157 1.19 0.24
## FirstAuthorFemale1 -0.1271 0.2070 -0.61 0.54
## LastAuthorFemale1 0.3024 0.1914 1.58 0.12
## UniqueAuthors2 0.1596 0.2601 0.61 0.54
## UniqueAuthors3 0.0641 0.2186 0.29 0.77
## UniqueAuthors4 -0.0676 0.3318 -0.20 0.84
## UniqueAuthors5 0.6457 0.6024 1.07 0.29
## Year1998 0.5588 0.7377 0.76 0.45
## Year1999 0.5733 0.7157 0.80 0.43
## Year2000 0.5289 0.7713 0.69 0.50
```

```

## Year2001      0.5156      0.9454      0.55      0.59
## Year2003     -0.1697      1.1029     -0.15      0.88
## Year2004      0.9903      0.7252      1.37      0.18
## Year2006      0.4116      0.7139      0.58      0.57
## Year2007      0.9755      0.7164      1.36      0.18
## Year2008      0.0292      0.7958      0.04      0.97
## Year2009      0.2985      0.7273      0.41      0.68
## Year2010      0.7023      0.7635      0.92      0.36
## Year2011      0.2494      0.7309      0.34      0.73
## Year2012      0.3893      0.7256      0.54      0.59
##
## Robust residual standard error: 0.47
## Multiple R-squared: 0.316, Adjusted R-squared: -0.0453
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~1. The remaining 52 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.591 0.866 0.974 0.928 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 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 4.718 1      2.172
## LastAuthorFemale 1.773 1      1.332
## Year      6.287 13      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 + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))

```

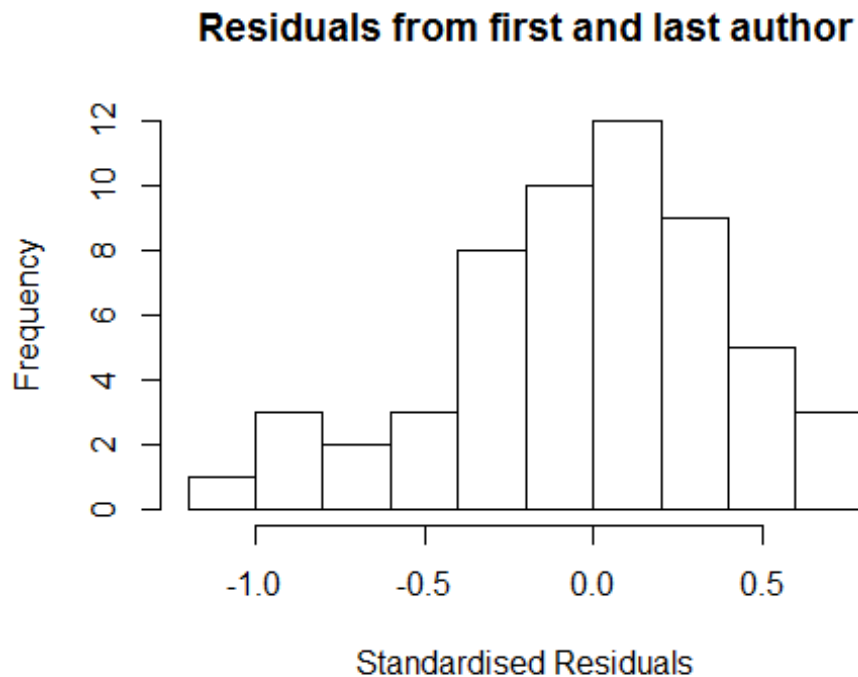
```

## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1116 -0.2311  0.0188  0.2814  0.7951
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9335     0.5238   1.78   0.082 .
## FirstAuthorFemale1 -0.0902     0.1491  -0.60   0.549
## LastAuthorFemale1  0.2397     0.1784   1.34   0.187
## Year1998         0.4790     0.5533   0.87   0.392
## Year1999         0.4935     0.5238   0.94   0.352
## Year2000         0.5104     0.5992   0.85   0.399
## Year2001         0.7296     0.6890   1.06   0.296
## Year2003        -0.1384     1.3682  -0.10   0.920
## Year2004         0.9105     0.5367   1.70   0.098 .
## Year2006         0.4124     0.5326   0.77   0.443
## Year2007         0.9571     0.5438   1.76   0.086 .
## Year2008         0.0285     0.6042   0.05   0.963
## Year2009         0.2770     0.5407   0.51   0.611
## Year2010         0.6831     0.5806   1.18   0.246
## Year2011         0.3324     0.5484   0.61   0.548
## Year2012         0.3156     0.5360   0.59   0.559
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.487
## Multiple R-squared:  0.288, Adjusted R-squared:  0.021
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
##  6 weights are ~= 1. The remaining 50 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.581  0.885  0.971   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      1.79e-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 lmrob.fit(x, y, control, init = init, mf = mf): M-step did NOT
## converge. Returning unconverged SM-estimate
```



```
## [1] "Regression 4: Last author gender, Year as factors"
##               GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.792 1      1.339
## Year             1.792 13      1.023

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1635 -0.2417  0.0332  0.2710  0.8032
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.93350    0.53544   1.74   0.089 .
```

```

## LastAuthorFemale1  0.22354    0.18246    1.23    0.228
## Year1998           0.47900    0.56444    0.85    0.401
## Year1999           0.49350    0.53544    0.92    0.362
## Year2000           0.48205    0.59774    0.81    0.425
## Year2001           0.70139    0.67626    1.04    0.306
## Year2003          -0.13027    1.54110   -0.08    0.933
## Year2004           0.91050    0.54809    1.66    0.104
## Year2006           0.39923    0.54402    0.73    0.467
## Year2007           0.91200    0.54408    1.68    0.101
## Year2008           0.00641    0.60897    0.01    0.992
## Year2009           0.24271    0.55221    0.44    0.663
## Year2010           0.62291    0.57313    1.09    0.283
## Year2011           0.31964    0.55723    0.57    0.569
## Year2012           0.27253    0.54712    0.50    0.621
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.473
## Multiple R-squared:  0.281, Adjusted R-squared:  0.0353
## Convergence in 29 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 50 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.524  0.884  0.969  0.922  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] "Sample size for the above analysis: 56"
## [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]"
##

```

```

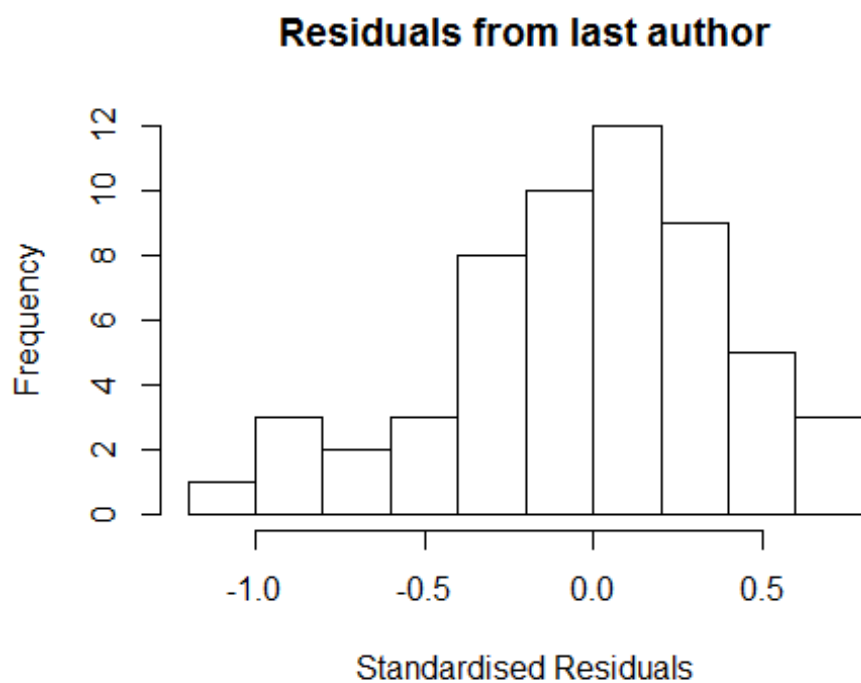
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      1      3      1      1      3      2      6      3      2      2      2      9      5      5
## 2011 2012
##      22     10
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      1      2      0      1      3      1      5      2      2      2      2      8      3      5
## 2011 2012
##      17      8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      1      2      0      1      3      1      5      2      2      2      2      8      3      5
## 2011 2012
##      16      7
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.08005973889231"
## [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 = 40, p-value = 0.4
## 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: 1.08005973889231"

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

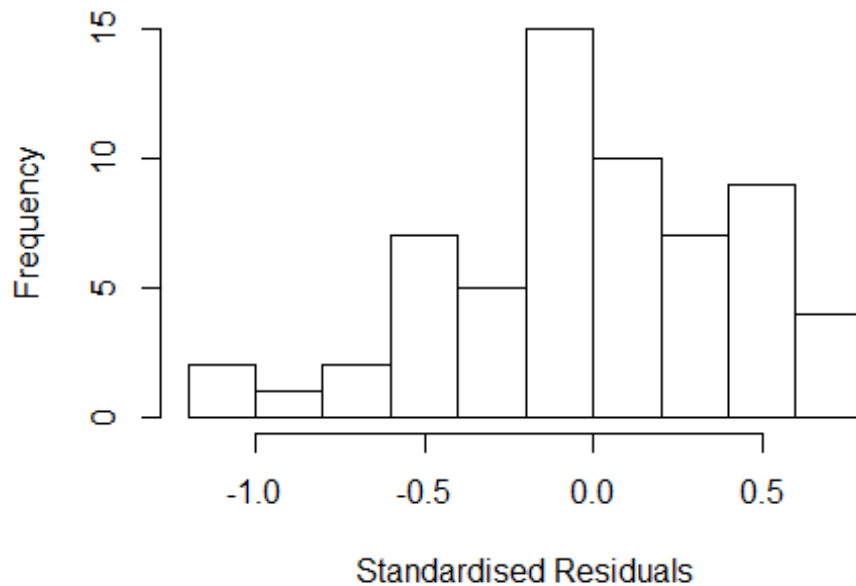
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 32, 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 | -6.808e+15 | 1 | NaN |
| ## LastAuthorFemale | -2.125e+15 | 1 | NaN |
| ## UniqueAuthors | -6.489e+18 | 3 | NaN |
| ## Year | 2.046e+30 | 15 | 10.24 |

Residuals from first and last author and team size



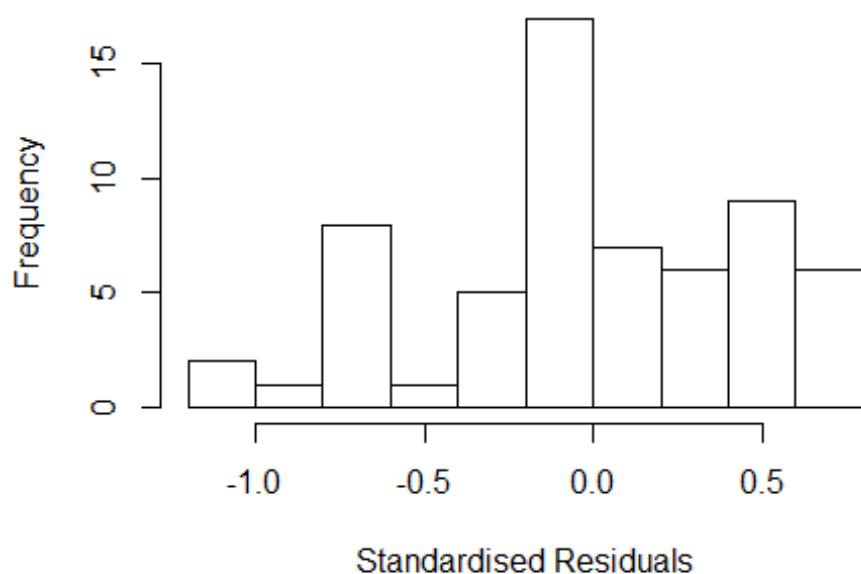
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.16e+00 -2.08e-01 4.16e-16 2.87e-01 7.84e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.33e+00 1.05e-01 1.26e+01 9.6e-16 ***
## FirstAuthorFemale1 1.09e+00 2.09e-01 5.21e+00 5.7e-06 ***
## LastAuthorFemale1 -9.27e-01 3.04e-01 -3.05e+00 0.00402 **
## UniqueAuthors2 4.47e-01 2.09e-01 2.14e+00 0.03837 *
## UniqueAuthors3 7.35e-01 1.57e-01 4.67e+00 3.2e-05 ***
## UniqueAuthors4 2.42e-01 1.26e-15 1.92e+14 < 2e-16 ***
## Year1997 1.53e-01 1.05e-01 1.46e+00 0.15270
## Year1998 1.97e-01 1.05e-01 1.88e+00 0.06752 .
## Year2000 2.06e-01 1.05e-01 1.96e+00 0.05637 .
## Year2001 9.14e-02 1.05e-01 8.70e-01 0.38955
```

```

## Year2002      -6.69e-02   1.05e-01 -6.40e-01   0.52725
## Year2003      -1.76e-01   2.50e-01 -7.00e-01   0.48616
## Year2004      -1.83e-01   4.28e-01 -4.30e-01   0.67095
## Year2005       1.13e-01   3.13e-01  3.60e-01   0.72018
## Year2006      -1.17e-01   1.80e-01 -6.50e-01   0.52014
## Year2007      -7.97e-01   1.97e-01 -4.03e+00   0.00023 ***
## Year2008      -6.13e-01   2.00e-01 -3.07e+00   0.00381 **
## Year2009      -1.75e-01   2.89e-01 -6.00e-01   0.54853
## Year2010      -3.24e-01   2.87e-01 -1.13e+00   0.26508
## Year2011      -8.20e-01   1.51e-01 -5.43e+00   2.8e-06 ***
## Year2012      -2.66e-01   2.23e-01 -1.20e+00   0.23824
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.609
## Multiple R-squared:  0.363, Adjusted R-squared:  0.053
## Convergence in 9 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 50 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.694  0.925  0.958  0.944  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 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 7.009e+11 1      837172
## LastAuthorFemale  6.820e+14 1      26114376
## Year              -2.451e+28 15      NaN

```

Residuals from first and last author



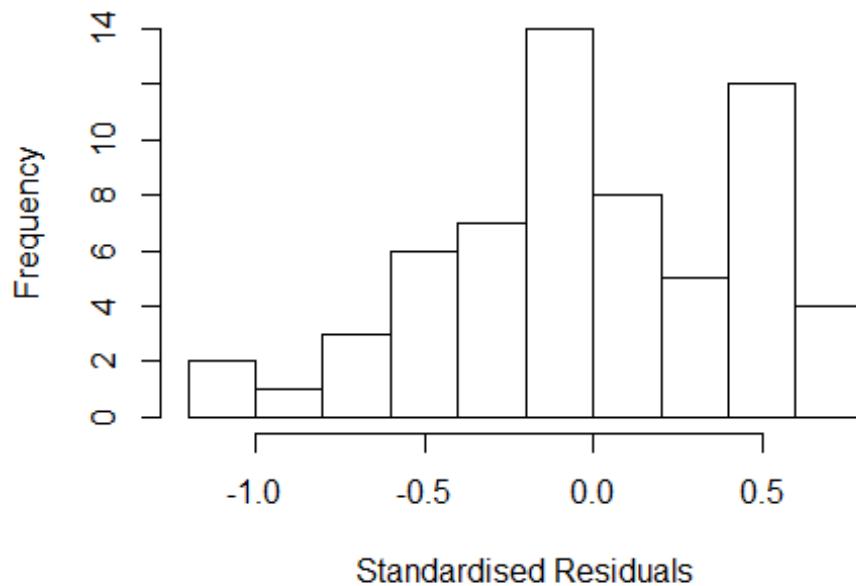
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.16e+00 -2.02e-01 -1.47e-15 3.77e-01 7.43e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.35538 0.09489 14.28 < 2e-16 ***
## FirstAuthorFemale1 0.64150 0.00601 106.73 < 2e-16 ***
## LastAuthorFemale1 -0.53527 0.16306 -3.28 0.002 **
## Year1997 0.18038 0.09489 1.90 0.064 .
## Year1998 0.29062 0.12824 2.27 0.028 *
## Year2000 0.17862 0.09489 1.88 0.066 .
## Year2001 0.11888 0.09508 1.25 0.218
## Year2002 -0.09438 0.09489 -0.99 0.325
## Year2003 -0.17880 0.25254 -0.71 0.483
## Year2004 -0.15562 0.43545 -0.36 0.723
## Year2005 0.11300 0.28827 0.39 0.697
## Year2006 -0.14438 0.17490 -0.83 0.414
```

```

## Year2007          -0.54562      0.09541    -5.72  8.7e-07 ***
## Year2008          -0.49404      0.20125    -2.45   0.018 *
## Year2009          -0.17965      0.29126    -0.62   0.541
## Year2010          -0.29999      0.27736    -1.08   0.285
## Year2011          -0.75204      0.15922    -4.72  2.4e-05 ***
## Year2012          -0.25487      0.21579    -1.18   0.244
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.573
## Multiple R-squared:  0.303, Adjusted R-squared:  0.0331
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 51 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.661  0.900   0.944   0.932   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 4.033e+13 1      6.350e+06
## Year              4.033e+13 15      2.841e+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
## -1.17e+00 -2.36e-01 -6.52e-16 4.08e-01 7.56e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3364 0.0987 13.54 < 2e-16 ***
## FirstAuthorFemale1 0.1443 0.1532 0.94 0.351
## Year1997 0.1614 0.0987 1.64 0.109
## Year1998 0.3096 0.1310 2.36 0.023 *
## Year2000 0.1976 0.0987 2.00 0.051 .
## Year2001 -0.0634 0.1570 -0.40 0.688
## Year2002 -0.0754 0.0987 -0.76 0.449
## Year2003 -0.1765 0.2511 -0.70 0.486
## Year2004 -0.1746 0.4305 -0.41 0.687
## Year2005 0.1130 0.3061 0.37 0.714
## Year2006 -0.1254 0.1768 -0.71 0.482
## Year2007 -0.5646 0.0992 -5.69 8.9e-07 ***
```

```

## Year2008          -0.5043      0.2059   -2.45    0.018 *
## Year2009          -0.1762      0.2904   -0.61    0.547
## Year2010          -0.3151      0.2841   -1.11    0.273
## Year2011          -0.7467      0.1611   -4.63    3.1e-05 ***
## Year2012          -0.2623      0.2199   -1.19    0.239
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.591
## Multiple R-squared:  0.291, Adjusted R-squared:  0.0389
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 54 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.677  0.911  0.953   0.939  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 1.321e+15  1      3.634e+07
## Year              1.321e+15 15      3.192e+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
## -1.16e+00 -2.15e-01  4.02e-16  3.80e-01  7.43e-01
##

```

```

## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.3550    0.0947   14.31 < 2e-16 ***
## LastAuthorFemale1 0.1070    0.1621    0.66  0.513
## Year1997        0.1800    0.0947    1.90  0.064 .
## Year1998        0.2910    0.1281    2.27  0.028 *
## Year2000        0.1790    0.0947    1.89  0.065 .
## Year2001       -0.0898    0.2056   -0.44  0.664
## Year2002       -0.0940    0.0947   -0.99  0.326
## Year2003       -0.1790    0.2513   -0.71  0.480
## Year2004       -0.1560    0.4323   -0.36  0.720
## Year2005        0.1130    0.2877    0.39  0.696
## Year2006       -0.1440    0.1747   -0.82  0.414
## Year2007       -0.5460    0.0952   -5.74 7.7e-07 ***
## Year2008       -0.4955    0.2004   -2.47  0.017 *
## Year2009       -0.1800    0.2900   -0.62  0.538
## Year2010       -0.3035    0.2772   -1.09  0.279
## Year2011       -0.7518    0.1584   -4.75 2.1e-05 ***
## Year2012       -0.2564    0.2157   -1.19  0.241
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.585
## Multiple R-squared:  0.287, Adjusted R-squared:  0.0329
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~ = 1. The remaining 54 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.674  0.905  0.953  0.937  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] "Sample size for the above analysis: 62"
## [1] ""
## [1] ""
## [1] "#####"

```



```

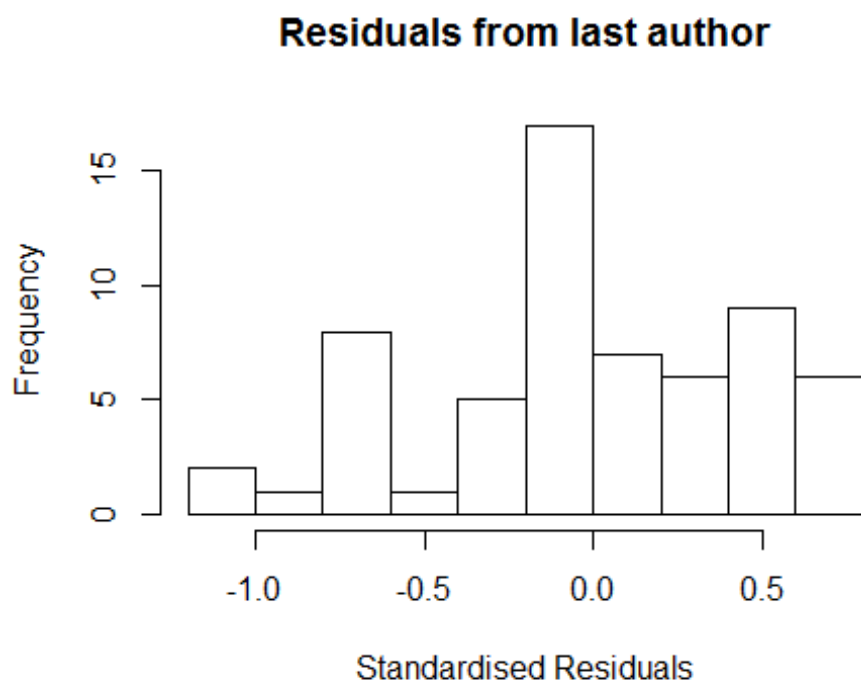
## [1] "Analysis of AJSC 3315"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    2    3    6    6    6    4    6   13    8    7    6   15   12   17
## 2012
##    21
##
## 1996 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    2    2    3    4    4    4    6    9    7    5    5   11    9   15
## 2012
##    21
##
## 1996 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    2    2    3    4    3    3    6    7    7    5    5    8    9   15
## 2012
##    18
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.46564193991492"
## [1] "Male first author team size 2018 geometric mean: 1.28208885398682"

## Warning in wilcox.test.default(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 = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.4985488060077"
## [1] "Male last author team size 2018 geometric mean: 1.22028493587281"

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

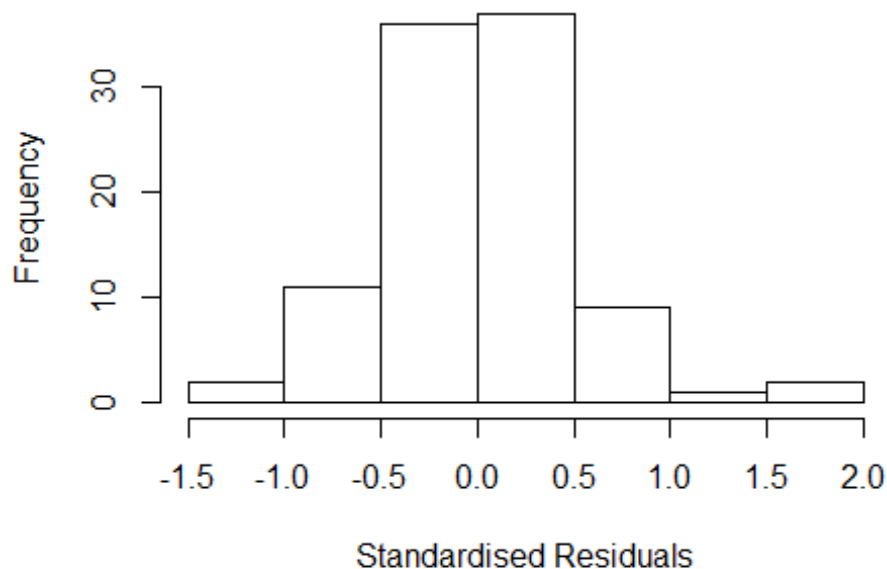
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 76, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
```

| | GVI | F | Df | $GVI^{\frac{1}{2 \cdot Df}}$ |
|-------------------|-----------|----|----|------------------------------|
| FirstAuthorFemale | 9.850e+12 | 1 | | 3.139e+06 |
| LastAuthorFemale | 9.712e+12 | 1 | | 3.116e+06 |
| UniqueAuthors | 1.286e+13 | 3 | | 1.531e+02 |
| Year | 3.241e+13 | 15 | | 2.821e+00 |

Residuals from first and last author and team size



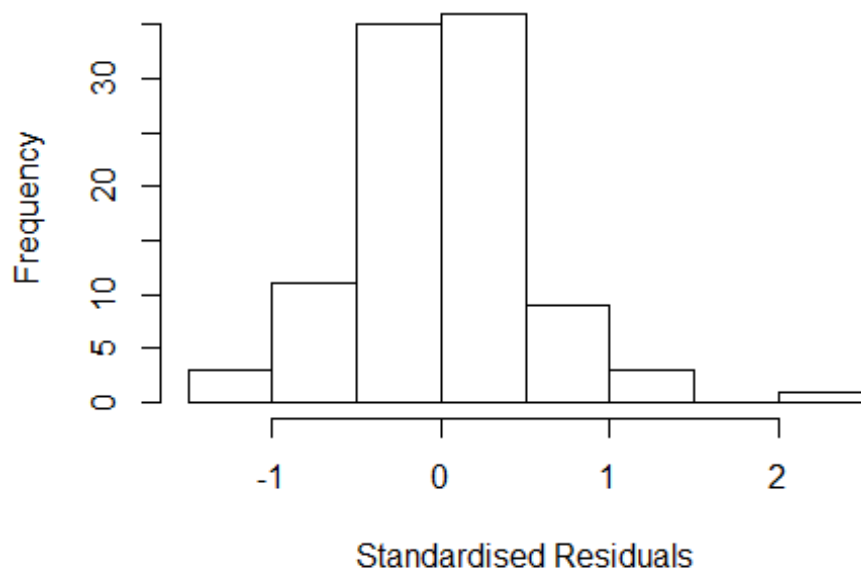
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.33814 -0.28123 0.00217 0.31065 1.96670
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6164 0.1093 14.79 < 2e-16 ***
## FirstAuthorFemale1 -0.4753 0.2874 -1.65 0.10229
## LastAuthorFemale1 0.4118 0.2854 1.44 0.15306
## UniqueAuthors2 0.1521 0.1647 0.92 0.35866
## UniqueAuthors3 0.4224 0.2303 1.83 0.07052 .
## UniqueAuthors4 0.1921 0.7376 0.26 0.79525
## Year1997 -0.6107 0.2464 -2.48 0.01537 *
## Year1999 -0.7872 0.2598 -3.03 0.00333 **
## Year2000 -0.6695 0.2915 -2.30 0.02434 *
## Year2001 -0.3283 0.0776 -4.23 6.4e-05 ***
```

```

## Year2002          -0.8394      0.3579    -2.35   0.02160 *
## Year2003          -0.1411      0.4205    -0.34   0.73810
## Year2004          -0.8019      0.2013    -3.98   0.00015 ***
## Year2005          -0.6168      0.1380    -4.47   2.7e-05 ***
## Year2006          -0.7578      0.1388    -5.46   5.6e-07 ***
## Year2007          -1.3352      0.1764    -7.57   6.9e-11 ***
## Year2008          -0.7589      0.3077    -2.47   0.01589 *
## Year2009          -0.7770      0.1700    -4.57   1.8e-05 ***
## Year2010          -0.8422      0.2120    -3.97   0.00016 ***
## Year2011          -1.0639      0.1335    -7.97   1.2e-11 ***
## Year2012          -0.6968      0.2157    -3.23   0.00182 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.511
## Multiple R-squared:  0.265, Adjusted R-squared:  0.0739
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 88 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.106  0.896  0.959  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      1.02e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 22.070 1      4.698
## LastAuthorFemale  21.626 1      4.650
## Year              5.601 15      1.059

```

Residuals from first and last author



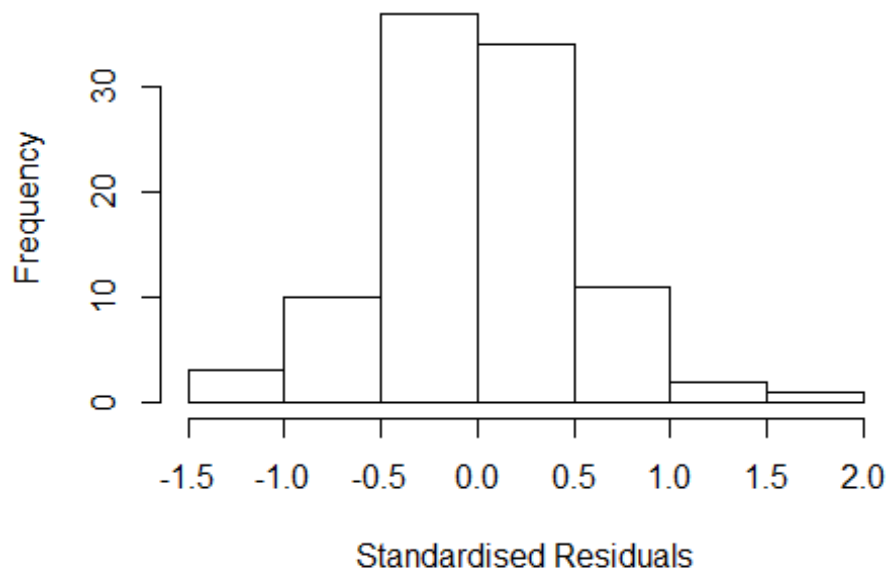
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.24353 -0.33155 0.00588 0.32757 2.01874
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6362 0.1095 14.94 < 2e-16 ***
## FirstAuthorFemale1 -0.5103 0.2654 -1.92 0.05810 .
## LastAuthorFemale1 0.4271 0.2621 1.63 0.10716
## Year1997 -0.6206 0.2540 -2.44 0.01675 *
## Year1999 -0.7971 0.2676 -2.98 0.00383 **
## Year2000 -0.6222 0.2427 -2.56 0.01221 *
## Year2001 -0.2955 0.0536 -5.52 4.2e-07 ***
## Year2002 -0.8553 0.3475 -2.46 0.01599 *
## Year2003 -0.1333 0.5195 -0.26 0.79822
## Year2004 -0.8178 0.2015 -4.06 0.00011 ***
## Year2005 -0.6250 0.1351 -4.63 1.4e-05 ***
## Year2006 -0.7288 0.1216 -5.99 5.6e-08 ***
```

```

## Year2007          -1.3275      0.1925    -6.90  1.1e-09 ***
## Year2008          -0.7224      0.3075    -2.35  0.02129 *
## Year2009          -0.7687      0.1736    -4.43  3.0e-05 ***
## Year2010          -0.8198      0.1950    -4.20  6.8e-05 ***
## Year2011          -1.0498      0.1348    -7.79  2.1e-11 ***
## Year2012          -0.5967      0.2145    -2.78  0.00675 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.516
## Multiple R-squared:  0.247, Adjusted R-squared:  0.0868
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 88 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0914 0.8940 0.9590 0.9060 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      1.02e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.57 1          1.889
## Year              3.57 15          1.043

```

Residuals from first author



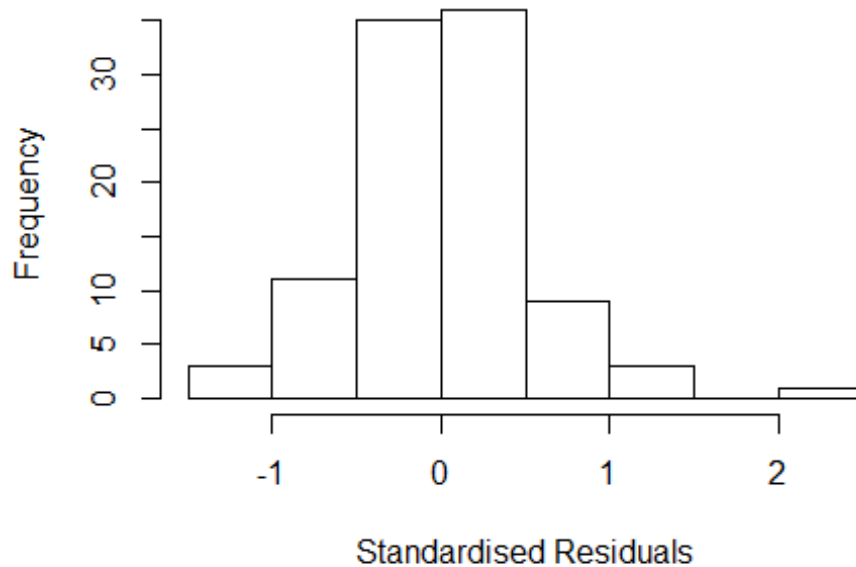
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.161177 -0.371884 -0.000824 0.333704 1.966342
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.7055 0.1096 15.56 < 2e-16 ***
## FirstAuthorFemale1 -0.1525 0.1096 -1.39 0.1679
## Year1997 -0.6553 0.2829 -2.32 0.0231 *
## Year1999 -0.8318 0.2989 -2.78 0.0067 **
## Year2000 -0.7601 0.3896 -1.95 0.0545 .
## Year2001 -0.3128 0.0618 -5.06 2.6e-06 ***
## Year2002 -0.9092 0.3163 -2.87 0.0052 **
## Year2003 -0.2285 0.3345 -0.68 0.4964
## Year2004 -0.8734 0.2034 -4.29 4.8e-05 ***
## Year2005 -0.6540 0.1256 -5.21 1.4e-06 ***
## Year2006 -0.7762 0.1291 -6.01 5.0e-08 ***
## Year2007 -1.3284 0.2485 -5.35 8.1e-07 ***
```

```

## Year2008          -0.6944      0.2547   -2.73    0.0078 **
## Year2009          -0.8132      0.1724   -4.72    9.9e-06 ***
## Year2010          -0.8064      0.1910   -4.22    6.3e-05 ***
## Year2011          -1.0878      0.1376   -7.91    1.1e-11 ***
## Year2012          -0.5443      0.2360   -2.31    0.0236 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.532
## Multiple R-squared:  0.203, Adjusted R-squared:  0.0458
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 91 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.143  0.887   0.958   0.910   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      1.02e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##  nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##  trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 3.768 1          1.941
## Year             3.768 15          1.045

```


Residuals from last author



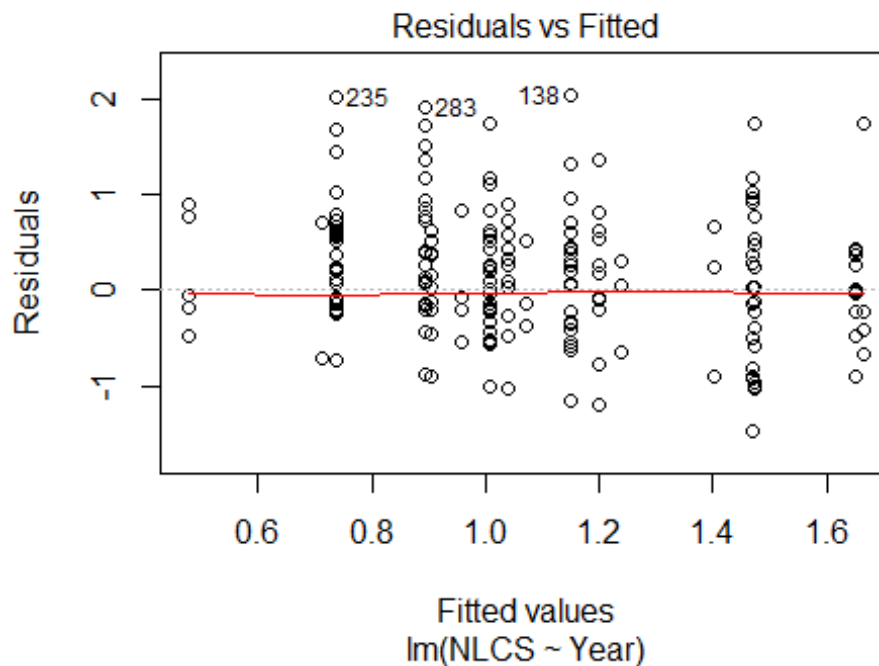
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1212 -0.3664 -0.0184 0.2997 1.8538
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5394 0.1176 13.09 < 2e-16 ***
## LastAuthorFemale1 0.0136 0.1176 0.12 0.90809
## Year1997 -0.5722 0.2132 -2.68 0.00882 **
## Year1999 -0.7487 0.2253 -3.32 0.00134 **
## Year2000 -0.6919 0.3867 -1.79 0.07733 .
## Year2001 -0.2712 0.0486 -5.58 3.1e-07 ***
## Year2002 -0.7877 0.3741 -2.11 0.03830 *
## Year2003 -0.1537 0.3238 -0.47 0.63629
## Year2004 -0.7419 0.2057 -3.61 0.00053 ***
## Year2005 -0.5842 0.1498 -3.90 0.00020 ***
## Year2006 -0.6544 0.1427 -4.59 1.6e-05 ***
## Year2007 -1.1960 0.2763 -4.33 4.2e-05 ***
```

```

## Year2008          -0.5945      0.2271    -2.62   0.01056 *
## Year2009          -0.7120      0.1765    -4.03   0.00012 ***
## Year2010          -0.7163      0.2084    -3.44   0.00093 ***
## Year2011          -0.9857      0.1364    -7.23   2.4e-10 ***
## Year2012          -0.4318      0.2380    -1.81   0.07338 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.57
## Multiple R-squared:  0.184, Adjusted R-squared:  0.0224
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 86 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.269  0.906   0.953   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      1.02e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##  nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##  trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 98"
## [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
##    3    3    5    7    6    4    8   16   14   18   14   19   27   16   35
## 2011 2012
##   66   45
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    3    4    7    4    3    5   13   12   13   12   14   26   13   33
## 2011 2012
##   51   36

```

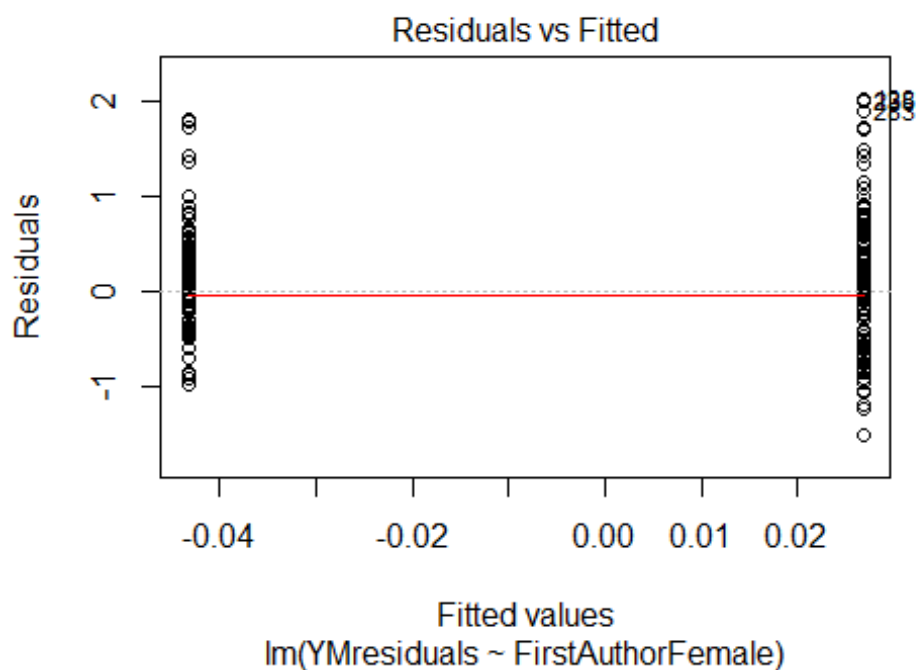
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    3    4    7    4    3    5   13   12   13   12   14   25   13   32
## 2011 2012
##   50   36
## [1] "Heteroscedasticity checks, confirming that there are problems with
##      these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 14, df = 16, p-value = 0.6
```



```
##
## 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: 1.05202627094242"
## [1] "Male first author team size 2018 geometric mean: 1.06757773192114"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
##
## Wilcoxon rank sum test with continuity correction
```

```
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 760, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.04467042246242"
## [1] "Male last author team size 2018 geometric mean: 1.07569058622018"

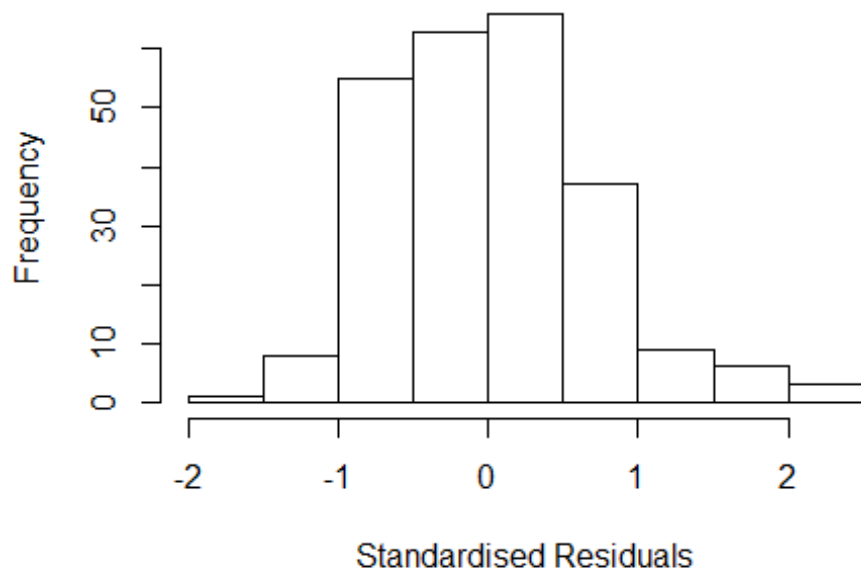
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 760, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
## Year as factors"
##
```

| | GVIF | Df | GVIF^(1/(2*Df)) |
|-------------------|-------|----|-----------------|
| FirstAuthorFemale | 16.42 | 1 | 4.053 |
| LastAuthorFemale | 10.70 | 1 | 3.271 |
| UniqueAuthors | 54.62 | 3 | 1.948 |
| Year | 31.88 | 16 | 1.114 |

Residuals from first and last author and team size



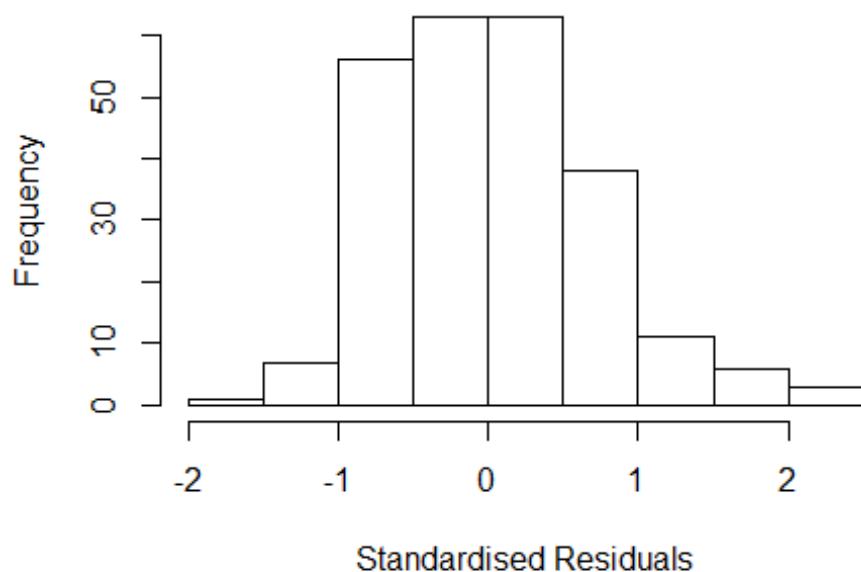
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5170 -0.5251 -0.0112 0.4582 2.1931
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7120 0.6138 1.16 0.247
## FirstAuthorFemale1 0.0300 0.2197 0.14 0.891
## LastAuthorFemale1 -0.1034 0.2167 -0.48 0.634
## UniqueAuthors2 0.3992 0.1935 2.06 0.040 *
## UniqueAuthors3 0.2739 0.3460 0.79 0.429
## UniqueAuthors4 0.5260 0.2423 2.17 0.031 *
## Year1997 0.6262 0.7058 0.89 0.376
## Year1998 0.2277 0.6677 0.34 0.733
## Year1999 -0.2864 0.6428 -0.45 0.656
## Year2000 0.5550 0.6483 0.86 0.393
```

```

## Year2001      0.2734      0.6906      0.40      0.693
## Year2002      0.5753      0.7139      0.81      0.421
## Year2003      0.5236      0.6539      0.80      0.424
## Year2004      0.9382      0.6261      1.50      0.135
## Year2005      0.8050      0.6785      1.19      0.237
## Year2006      0.2289      0.6307      0.36      0.717
## Year2007      0.6977      0.6501      1.07      0.284
## Year2008      0.3359      0.6327      0.53      0.596
## Year2009      0.2984      0.6428      0.46      0.643
## Year2010      0.2263      0.6261      0.36      0.718
## Year2011     -0.0396      0.6235     -0.06      0.949
## Year2012      0.1511      0.6369      0.24      0.813
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.733
## Multiple R-squared:  0.171, Adjusted R-squared:  0.0943
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 228 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.350  0.887  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      4.03e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 6.561 1      2.561
## LastAuthorFemale  6.330 1      2.516
## Year              1.917 16      1.021

```

Residuals from first and last author



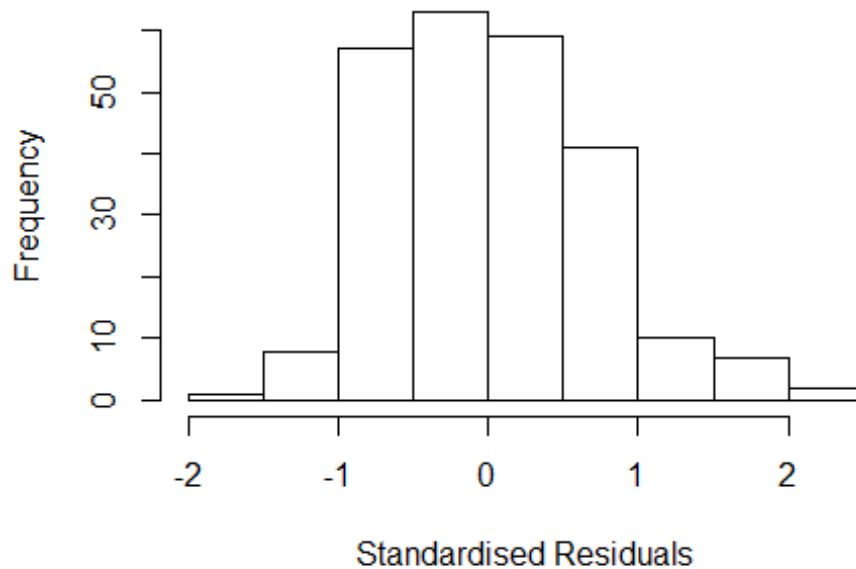
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5145 -0.5157 -0.0297 0.4610 2.0998
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.71200 0.61499 1.16 0.25
## FirstAuthorFemale1 0.13610 0.21945 0.62 0.54
## LastAuthorFemale1 -0.20347 0.21861 -0.93 0.35
## Year1997 0.76485 0.74662 1.02 0.31
## Year1998 0.22757 0.66949 0.34 0.73
## Year1999 -0.23991 0.65674 -0.37 0.72
## Year2000 0.55352 0.64897 0.85 0.39
## Year2001 0.45040 0.66908 0.67 0.50
## Year2002 0.73539 0.70291 1.05 0.30
## Year2003 0.52142 0.65469 0.80 0.43
## Year2004 0.98144 0.62631 1.57 0.12
## Year2005 0.80246 0.67928 1.18 0.24
```

```

## Year2006          0.22724    0.63167    0.36    0.72
## Year2007          0.72250    0.65159    1.11    0.27
## Year2008          0.37623    0.63702    0.59    0.56
## Year2009          0.33554    0.64304    0.52    0.60
## Year2010          0.27613    0.62900    0.44    0.66
## Year2011          0.00344    0.62639    0.01    1.00
## Year2012          0.14364    0.64211    0.22    0.82
##
## Robust residual standard error: 0.73
## Multiple R-squared:  0.155, Adjusted R-squared:  0.0889
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 230 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.388  0.885   0.952   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      4.03e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.329 1          1.153
## Year              1.329 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
## -1.5045 -0.5257 -0.0386 0.4825 2.0935
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7120 0.6146 1.16 0.25
## FirstAuthorFemale1 -0.0457 0.0992 -0.46 0.65
## Year1997 0.7505 0.7501 1.00 0.32
## Year1998 0.2276 0.6684 0.34 0.73
## Year1999 -0.2459 0.6546 -0.38 0.71
## Year2000 0.5479 0.6479 0.85 0.40
## Year2001 0.3703 0.6599 0.56 0.58
## Year2002 0.7784 0.6906 1.13 0.26
## Year2003 0.5130 0.6543 0.78 0.43
## Year2004 0.9583 0.6252 1.53 0.13
## Year2005 0.7925 0.6790 1.17 0.24
## Year2006 0.2207 0.6321 0.35 0.73
```

```

## Year2007          0.7357      0.6488      1.13      0.26
## Year2008          0.3825      0.6353      0.60      0.55
## Year2009          0.3608      0.6421      0.56      0.57
## Year2010          0.2681      0.6268      0.43      0.67
## Year2011         -0.0119      0.6238     -0.02      0.98
## Year2012          0.1316      0.6383      0.21      0.84
##
## Robust residual standard error: 0.731
## Multiple R-squared: 0.153, Adjusted R-squared: 0.0902
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 228 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.392  0.886  0.948  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      4.03e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## 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
## -1.5206 -0.5212 -0.0309  0.4753  2.0797
##
## Coefficients:

```

```

##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.71200    0.61314    1.16    0.25
## LastAuthorFemale1 -0.08162    0.09818   -0.83    0.41
## Year1997          0.77372    0.74246    1.04    0.30
## Year1998          0.22784    0.66772    0.34    0.73
## Year1999         -0.23558    0.65527   -0.36    0.72
## Year2000          0.55701    0.64772    0.86    0.39
## Year2001          0.41083    0.66426    0.62    0.54
## Year2002          0.78411    0.69476    1.13    0.26
## Year2003          0.52669    0.65274    0.81    0.42
## Year2004          0.97356    0.62416    1.56    0.12
## Year2005          0.80862    0.67688    1.19    0.23
## Year2006          0.23133    0.62948    0.37    0.71
## Year2007          0.73877    0.64798    1.14    0.26
## Year2008          0.39633    0.63485    0.62    0.53
## Year2009          0.35970    0.63943    0.56    0.57
## Year2010          0.28411    0.62710    0.45    0.65
## Year2011          0.00433    0.62444    0.01    0.99
## Year2012          0.15280    0.63964    0.24    0.81
##
## Robust residual standard error: 0.735
## Multiple R-squared:  0.154, Adjusted R-squared:  0.0912
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 229 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.403  0.887  0.950  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      4.03e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 248"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 3317"
## [1] "#####"

```

```

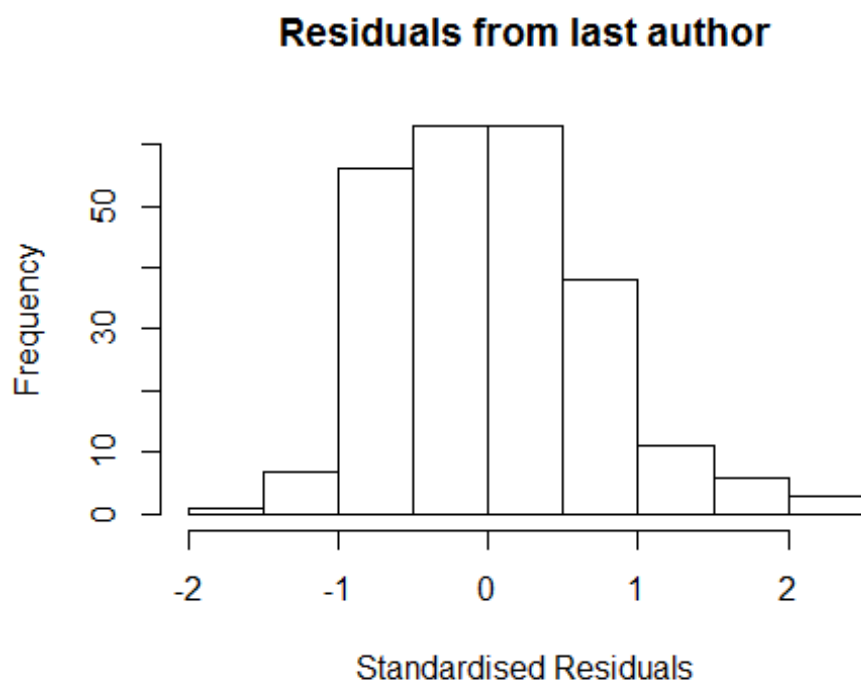
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1998 1999 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    3    1    1    2    1    1    3    7    6    5    5    6   16    9
##
## 1998 1999 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    3    1    0    1    0    1    1    5    6    5    4    5   12    8
##
## 1998 1999 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    3    1    0    1    0    1    0    5    5    5    3    5   11    7
## [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

##
## 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: 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

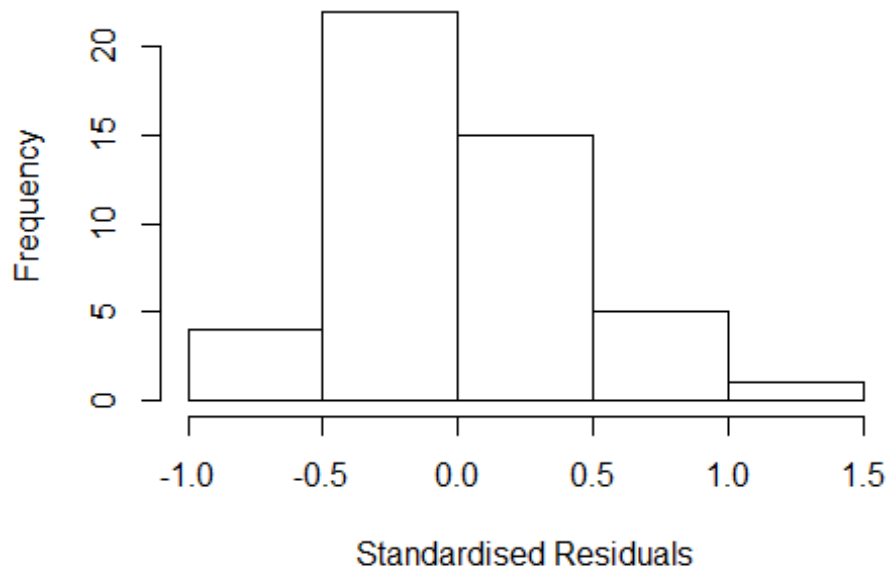
```



```
##
## 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] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
```

| | GVIF | Df | GVIF ^{1/(2*Df)} |
|----------------------|-----------|----|--------------------------|
| ## FirstAuthorFemale | 7.982e+00 | 1 | 2.825e+00 |
| ## LastAuthorFemale | 5.442e+13 | 1 | 7.377e+06 |
| ## UniqueAuthors | 1.862e+29 | 2 | 2.077e+07 |
| ## Year | 1.842e+30 | 10 | 3.260e+01 |

Residuals from first and last author and team size



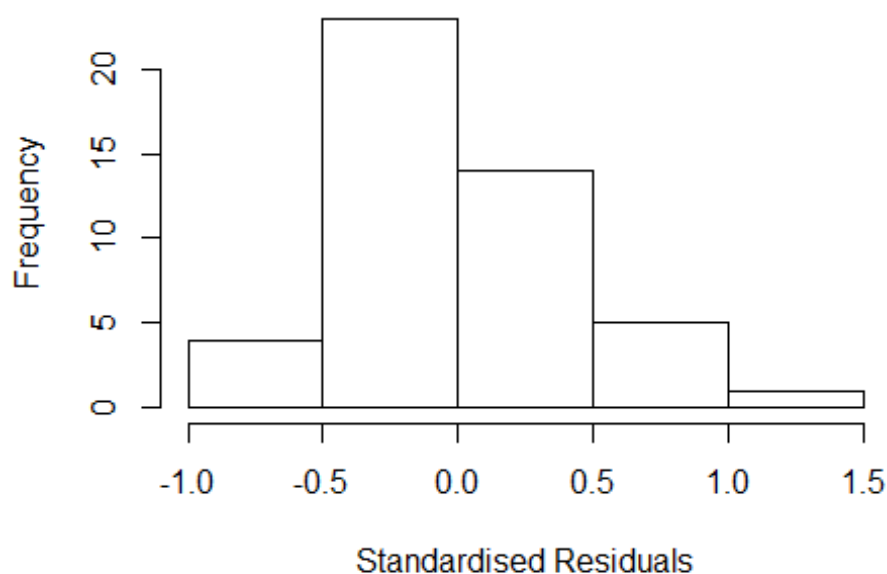
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -8.82e-01 -2.85e-01 -4.44e-16 2.70e-01 1.16e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.37809 0.23870 1.58 0.12303
## FirstAuthorFemale1 0.18818 0.18956 0.99 0.32830
## LastAuthorFemale1 -0.27466 0.15418 -1.78 0.08434 .
## UniqueAuthors2 0.18236 0.17615 1.04 0.30831
## UniqueAuthors3 -0.00556 0.20827 -0.03 0.97885
## Year1999 0.41447 0.23016 1.80 0.08117 .
## Year2002 0.83591 0.23870 3.50 0.00139 **
## Year2004 1.34221 0.23475 5.72 2.5e-06 ***
## Year2006 0.59016 0.43178 1.37 0.18121
## Year2007 1.00515 0.24798 4.05 0.00030 ***
```

```

## Year2008          1.40839    0.27329    5.15  1.3e-05 ***
## Year2009          1.23161    0.32383    3.80  0.00061 ***
## Year2010          0.76182    0.23294    3.27  0.00257 **
## Year2011          0.87297    0.24852    3.51  0.00135 **
## Year2012          0.32431    0.30664    1.06  0.29814
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.514
## Multiple R-squared:  0.504, Adjusted R-squared:  0.287
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 41 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.592  0.941  0.964  0.938  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      2.13e-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 7.416e+00 1      2.723e+00
## LastAuthorFemale  2.791e+14 1      1.671e+07
## Year              1.647e+15 10      5.766e+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
## -0.91228 -0.24416 -0.00679 0.28525 1.12572
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.357 0.224 1.59 0.12080
## FirstAuthorFemale1 0.141 0.212 0.67 0.51007
## LastAuthorFemale1 -0.226 0.173 -1.31 0.20002
## Year1999 0.430 0.224 1.92 0.06339 .
## Year2002 0.857 0.224 3.82 0.00053 ***
## Year2004 1.497 0.201 7.45 1.2e-08 ***
## Year2006 0.641 0.428 1.50 0.14396
## Year2007 1.041 0.257 4.04 0.00029 ***
## Year2008 1.510 0.247 6.12 6.1e-07 ***
## Year2009 1.246 0.320 3.90 0.00043 ***
## Year2010 0.811 0.213 3.80 0.00057 ***
## Year2011 0.947 0.239 3.96 0.00036 ***
```

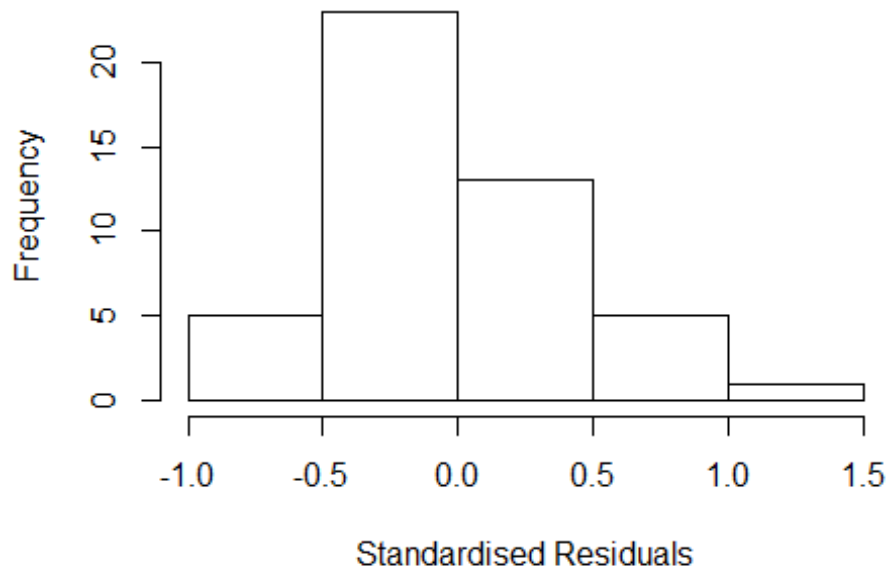


```

## Year2012          0.399      0.315      1.27  0.21430
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.471
## Multiple R-squared:  0.5,   Adjusted R-squared:  0.323
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 40 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.547  0.906  0.962  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      2.13e-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 4.012  1      2.003
## Year              4.012 10      1.072

```

Residuals from first author



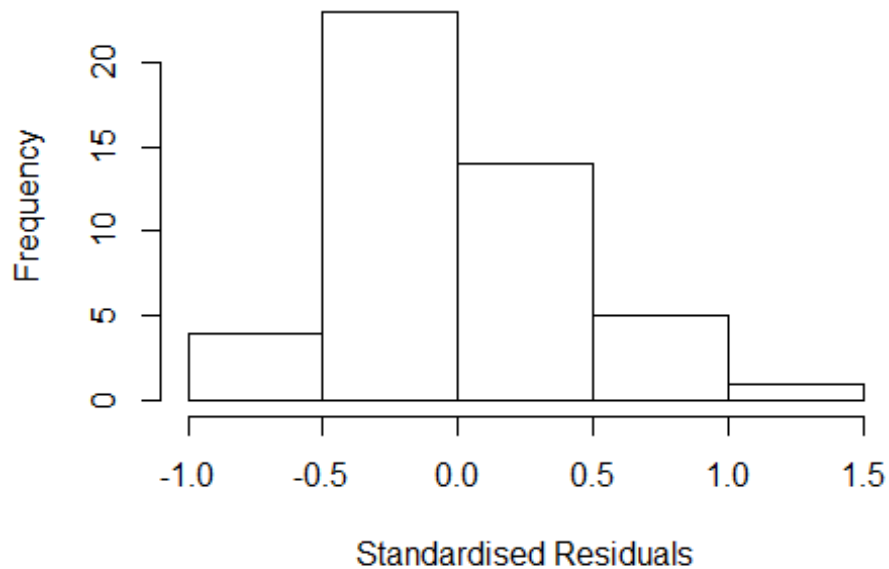
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9368 -0.2530 -0.0339 0.2601 1.1012
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.24601 0.24228 1.02 0.31689
## FirstAuthorFemale1 0.00221 0.17298 0.01 0.98990
## Year1999 0.54099 0.24228 2.23 0.03205 *
## Year2002 0.96799 0.24228 4.00 0.00032 ***
## Year2004 1.38199 0.24228 5.70 1.9e-06 ***
## Year2006 0.68863 0.45501 1.51 0.13915
## Year2007 1.08487 0.25859 4.20 0.00018 ***
## Year2008 1.61523 0.25467 6.34 2.7e-07 ***
## Year2009 1.33391 0.30805 4.33 0.00012 ***
## Year2010 0.83784 0.25374 3.30 0.00222 **
## Year2011 1.00809 0.26506 3.80 0.00055 ***
## Year2012 0.45343 0.35384 1.28 0.20846
```

```

## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.516
## Multiple R-squared:  0.477, Adjusted R-squared:  0.313
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 41 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.629  0.914  0.974  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] "Regression 4: Last author gender, Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 4.984e+14 1           2.233e+07
## Year           4.984e+14 10           5.431e+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
## -8.19e-01 -2.04e-01 -2.22e-16 2.43e-01 1.22e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.348 0.228 1.53 0.13609
## LastAuthorFemale1 -0.145 0.162 -0.90 0.37576
## Year1999 0.439 0.228 1.92 0.06277 .
## Year2002 0.866 0.228 3.79 0.00057 ***
## Year2004 1.425 0.184 7.76 4.1e-09 ***
## Year2006 0.616 0.426 1.45 0.15644
## Year2007 1.066 0.234 4.56 6.0e-05 ***
## Year2008 1.571 0.227 6.92 4.8e-08 ***
## Year2009 1.270 0.342 3.71 0.00071 ***
## Year2010 0.827 0.214 3.87 0.00045 ***
## Year2011 0.990 0.223 4.44 8.5e-05 ***
## Year2012 0.411 0.333 1.23 0.22557
```

```

## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.464
## Multiple R-squared:  0.502, Adjusted R-squared:  0.345
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 40 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.470  0.906  0.966  0.919  0.985  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 3318"
## [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
##    1    2    2    2    2    6    5    4    4    6    3    6    7    13    12
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    2    2    1    2    6    4    3    4    5    3    5    6    9    9
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    2    2    1    2    6    3    3    4    5    3    5    6    8    9
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.23007550557797"
## [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 6.573e+00 1      2.564e+00
## LastAuthorFemale  2.649e+14 1      1.628e+07
## UniqueAuthors     1.776e+16 3      5.108e+02
## Year              4.739e+16 14     3.941e+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
## -1.1660 -0.2778  0.0399  0.2842  0.9262
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    3.12e-01   2.70e-01   1.16e+00  0.25470
## FirstAuthorFemale1 3.52e-01   1.91e-01   1.84e+00  0.07332 .
## LastAuthorFemale1 -2.16e-01   1.38e-01  -1.57e+00  0.12514
## UniqueAuthors2     5.36e-01   1.60e-01   3.35e+00  0.00176 **
## UniqueAuthors3     6.61e-01   7.76e-01   8.50e-01  0.39982
## UniqueAuthors5    -1.45e-01   3.12e-01  -4.70e-01  0.64435
## Year1999          6.75e-01   1.01e-01   6.66e+00  5.6e-08 ***
## Year2000          2.66e-01   3.59e-01   7.40e-01  0.46331
## Year2001          9.64e-01   5.08e-08   1.90e+07  < 2e-16 ***
## Year2002          2.51e-01   7.96e-01   3.20e-01  0.75402
## Year2003          2.82e-01   2.28e-01   1.24e+00  0.22386
## Year2004          8.75e-01   2.46e-01   3.56e+00  0.00096 ***
## Year2005          7.57e-01   1.85e-01   4.10e+00  0.00019 ***
## Year2006          7.86e-01   8.39e-02   9.37e+00  1.2e-11 ***
## Year2007          7.27e-02   1.67e-01   4.40e-01  0.66509
## Year2008          3.20e-01   2.10e-01   1.52e+00  0.13519
## Year2009          6.21e-01   2.92e-01   2.13e+00  0.03967 *
## Year2010          4.33e-01   2.44e-01   1.77e+00  0.08380 .
## Year2011          3.57e-01   2.97e-01   1.20e+00  0.23567
## Year2012          2.73e-01   1.89e-01   1.45e+00  0.15515
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.513
## Multiple R-squared:  0.428, Adjusted R-squared:  0.157
## Convergence in 21 IRWLS iterations
##

```

```

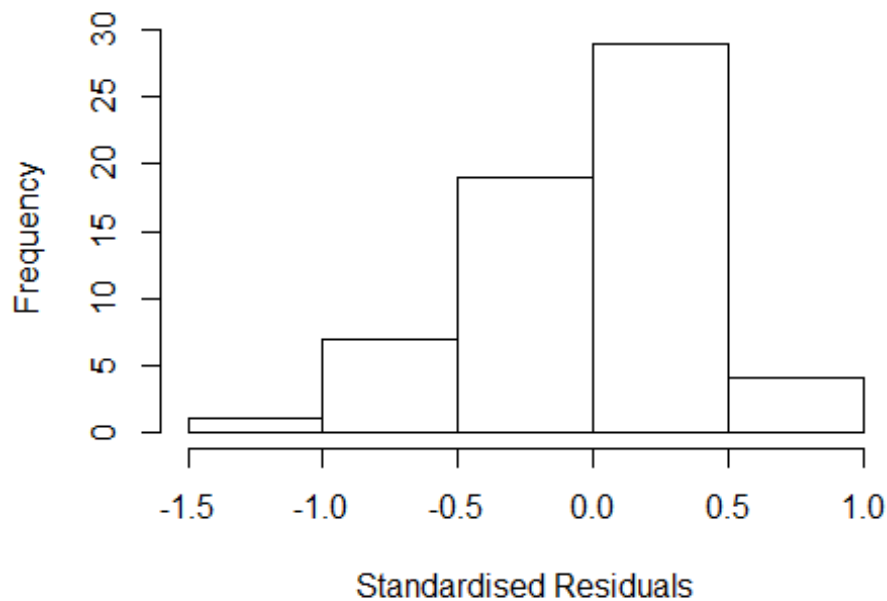
## Robustness weights:
## 7 weights are ~= 1. The remaining 53 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.584  0.927  0.961  0.936  0.989  0.998
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00      5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      1.67e-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"

## 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
## -0.9811 -0.3658  0.0431  0.3446  1.1053
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.469     0.261    1.80  0.07965 .
## FirstAuthorFemale1  0.303     0.274    1.11  0.27463
## LastAuthorFemale1 -0.325     0.186   -1.75  0.08710 .
## Year1999         0.889     0.264    3.37  0.00160 **
## Year2000         0.265     0.343    0.77  0.44342
## Year2001         0.964     0.000    Inf < 2e-16 ***
## Year2002         0.172     0.565    0.31  0.76154
```

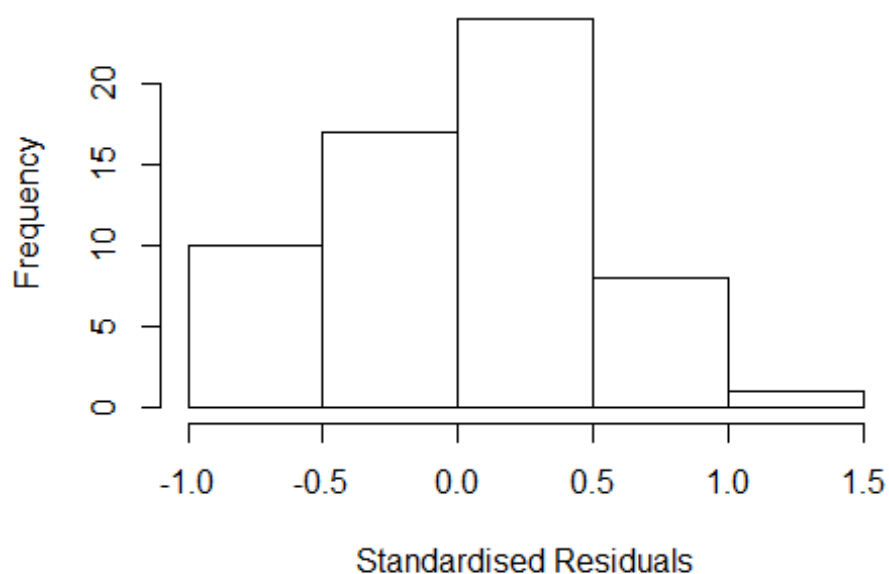


```

## Year2003          0.319      0.216      1.48  0.14611
## Year2004          0.981      0.326      3.01  0.00432 **
## Year2005          0.874      0.282      3.10  0.00343 **
## Year2006          0.915      0.160      5.71  9.8e-07 ***
## Year2007          0.152      0.182      0.83  0.40925
## Year2008          0.461      0.288      1.60  0.11705
## Year2009          0.610      0.361      1.69  0.09883 .
## Year2010          0.658      0.171      3.85  0.00038 ***
## Year2011          0.512      0.264      1.94  0.05879 .
## Year2012          0.366      0.155      2.37  0.02226 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.602
## Multiple R-squared:  0.24, Adjusted R-squared:  -0.0423
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 53 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.716  0.912  0.963   0.940  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      1.67e-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"
## 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"
## 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))
## 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
## -1.1058 -0.3706  0.0373  0.3489  1.1965
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.597     0.205    2.91  0.00559 **
## LastAuthorFemale1 -0.150     0.205   -0.73  0.46790
## Year1999        0.976     0.341    2.87  0.00635 **
## Year2000        0.265     0.344    0.77  0.44436
## Year2001        0.964     0.000    Inf < 2e-16 ***
## Year2002        0.108     0.481    0.23  0.82272
## Year2003        0.271     0.208    1.30  0.20088
## Year2004        0.837     0.202    4.15  0.00015 ***
```

```

## Year2005          0.836      0.245      3.41  0.00139 **
## Year2006          0.915      0.161      5.70  9.3e-07 ***
## Year2007          0.125      0.170      0.74  0.46534
## Year2008          0.515      0.340      1.52  0.13666
## Year2009          0.518      0.359      1.44  0.15629
## Year2010          0.658      0.171      3.85  0.00038 ***
## Year2011          0.509      0.263      1.94  0.05911 .
## Year2012          0.367      0.161      2.28  0.02720 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.596
## Multiple R-squared:  0.225, Adjusted R-squared:  -0.039
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~ = 1. The remaining 53 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.666  0.922  0.962   0.937   0.985   0.999
## Algorithmic parameters:
##           tuning.chi          bb          tuning.psi          refine.tol
##           1.55e+00          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 3319"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 2000 2001 2002 2004 2006 2009 2010 2011 2012
##    2    1    1    1    1    2    8    2    6
##
## 2000 2001 2002 2004 2006 2009 2010 2011 2012
##    0    1    0    1    0    2    7    2    4
##

```

```

## 2000 2001 2002 2004 2006 2009 2010 2011 2012
##    0    1    0    1    0    2    7    2    4
## [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"

## Warning in wilcox.test.default(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 = 0.4
## 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: 1.5874010519682"

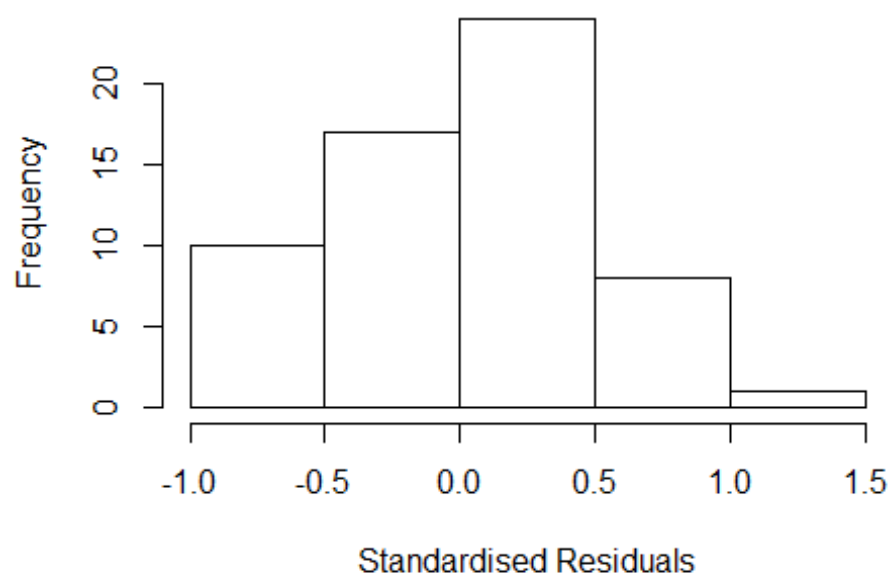
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 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"
## [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

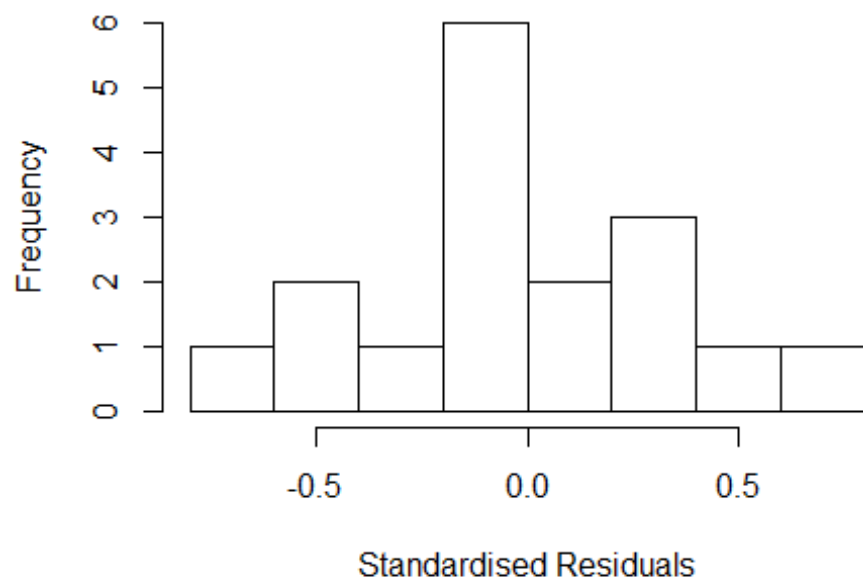
```

Residuals from last author



| | | | | |
|----|-------------------|------------|----|---------------------|
| ## | | GVIF | Df | $GVIF^{(1/(2*Df))}$ |
| ## | FirstAuthorFemale | -1.501e+15 | 1 | NaN |
| ## | Year | -1.501e+15 | 5 | NaN |

Residuals from first author



```

## [1] "Regression 4: Last author gender, Year as factors"
##              GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 2.812  1          1.677
## Year              2.812  5          1.109

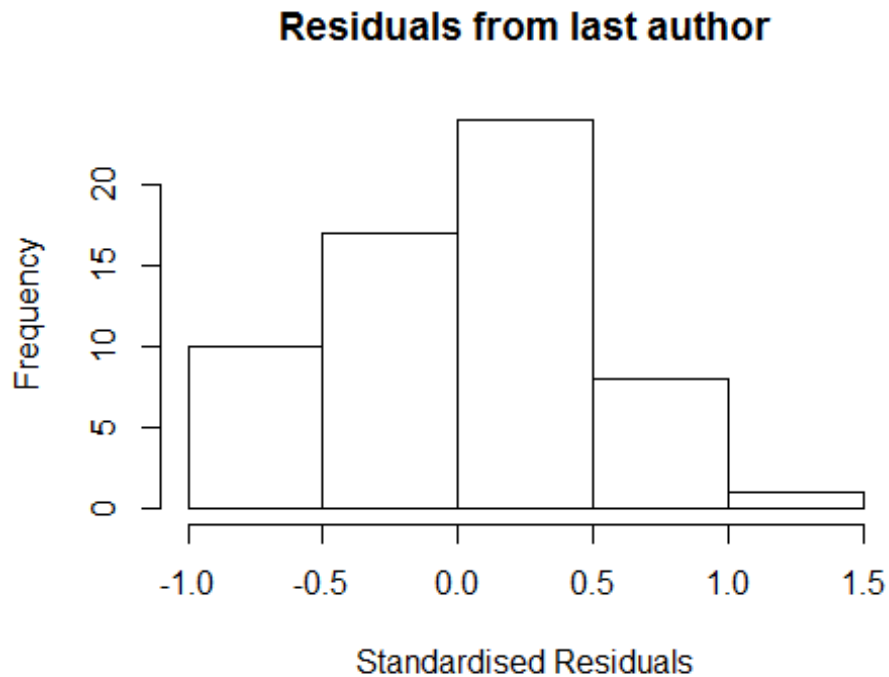
## [1] "Sample size for the above analysis: 17"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 3320"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    2    9   12    5   10   14   16   20   13   26   31   26   46   55
## 2011 2012
##   62   57
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    1    7   10    3    6   11   14   15   13   23   26   22   46   42
## 2011 2012
##   46   51
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    1    7   10    3    6   11   14   15   13   23   25   21   44   41
## 2011 2012
##   46   47
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.33497689498765"
## [1] "Male first author team size 2018 geometric mean: 1.16722859511375"

## Warning in wilcox.test.default(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 = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.28208885398682"
## [1] "Male last author team size 2018 geometric mean: 1.18532661211153"

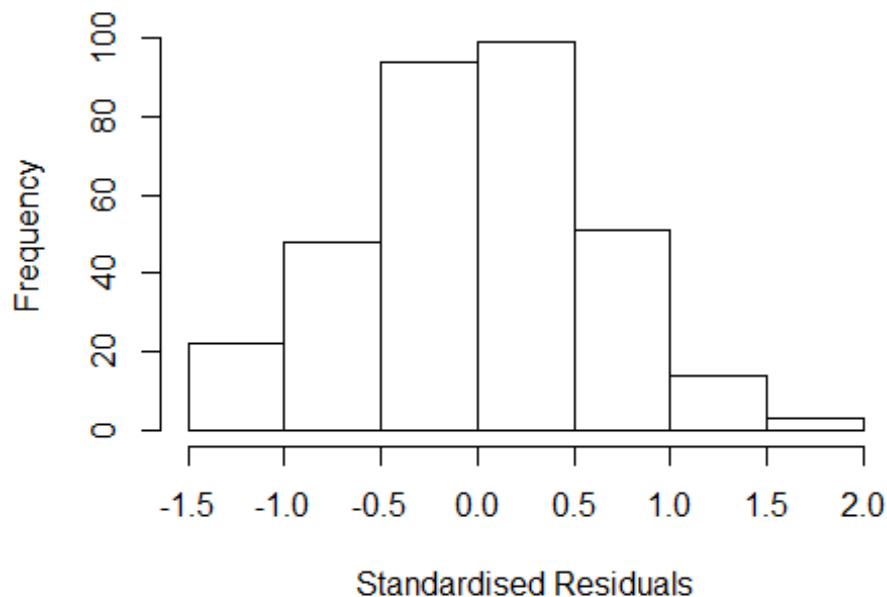
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 190, p-value = 0.6
## 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 7.563  1      2.750
## LastAuthorFemale  8.132  1      2.852
## UniqueAuthors    2.964  3      1.199
## Year              4.755 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
## -1.3770 -0.4223 0.0142 0.4118 1.8874
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8343 0.1591 5.24 2.9e-07 ***
## FirstAuthorFemale1 -0.2185 0.1474 -1.48 0.1393
## LastAuthorFemale1 0.2664 0.1432 1.86 0.0638 .
## UniqueAuthors2 0.1195 0.1077 1.11 0.2680
## UniqueAuthors3 0.0961 0.2735 0.35 0.7257
## UniqueAuthors4 0.2645 0.4270 0.62 0.5360
## Year1997 -0.4053 0.1591 -2.55 0.0114 *
## Year1998 -0.1533 0.2074 -0.74 0.4604
## Year1999 -0.1843 0.2589 -0.71 0.4771
## Year2000 -0.1361 0.9172 -0.15 0.8821
```

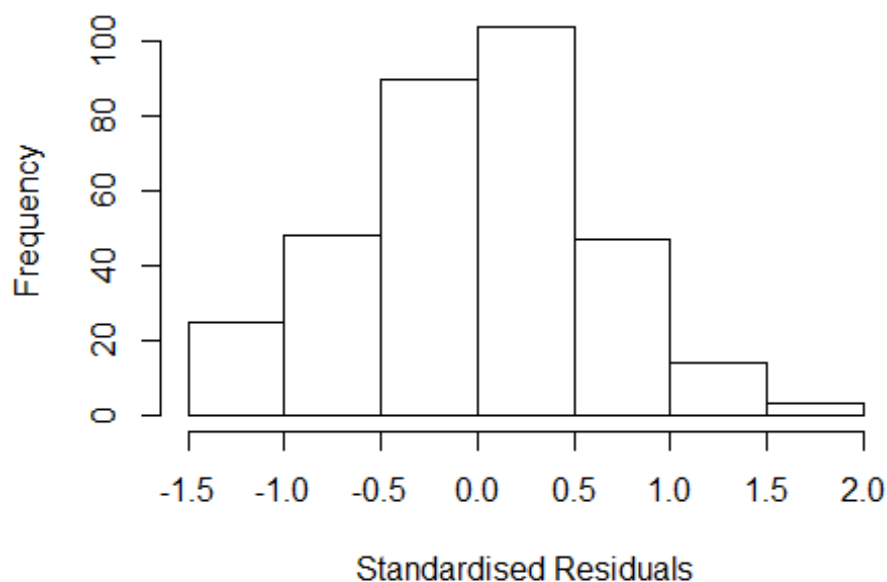


```

## Year2001          0.5854      0.1919      3.05      0.0025 **
## Year2002          0.2981      0.3251      0.92      0.3599
## Year2003          0.0665      0.2303      0.29      0.7730
## Year2004          0.2017      0.2546      0.79      0.4289
## Year2005          0.4182      0.2740      1.53      0.1279
## Year2006          0.3302      0.1974      1.67      0.0953 .
## Year2007          0.3168      0.2096      1.51      0.1318
## Year2008          0.4949      0.2222      2.23      0.0267 *
## Year2009          0.2438      0.1800      1.35      0.1766
## Year2010          0.2098      0.1849      1.13      0.2573
## Year2011          0.2930      0.1866      1.57      0.1175
## Year2012          0.1613      0.1839      0.88      0.3811
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.628
## Multiple R-squared:  0.0678, Adjusted R-squared:  0.00448
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 303 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.346  0.875   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      3.02e-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 7.419 1      2.724
## LastAuthorFemale  7.884 1      2.808
## Year              2.482 16      1.029

```

Residuals from first and last author



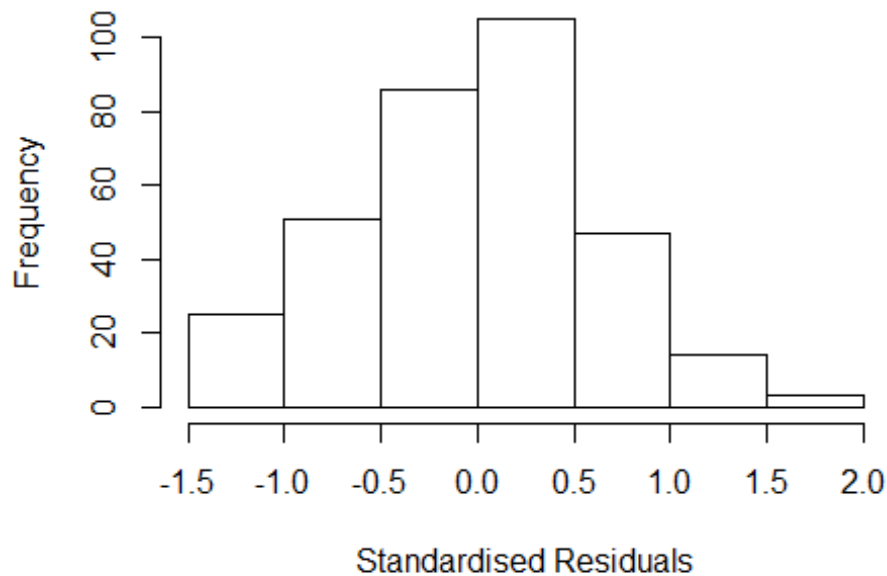
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.40431 -0.42457 0.00303 0.39835 1.95444
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8329 0.1594 5.22 3.2e-07 ***
## FirstAuthorFemale1 -0.2376 0.1489 -1.60 0.1116
## LastAuthorFemale1 0.2905 0.1438 2.02 0.0442 *
## Year1997 -0.4039 0.1594 -2.53 0.0118 *
## Year1998 -0.1527 0.2075 -0.74 0.4624
## Year1999 -0.1721 0.2595 -0.66 0.5078
## Year2000 -0.2094 0.8543 -0.25 0.8066
## Year2001 0.5859 0.1924 3.04 0.0025 **
## Year2002 0.3157 0.3270 0.97 0.3350
## Year2003 0.0779 0.2322 0.34 0.7375
## Year2004 0.2198 0.2501 0.88 0.3802
## Year2005 0.4306 0.2705 1.59 0.1124
```

```

## Year2006          0.3517      0.2011      1.75      0.0813 .
## Year2007          0.3261      0.2105      1.55      0.1223
## Year2008          0.5184      0.2241      2.31      0.0214 *
## Year2009          0.2579      0.1789      1.44      0.1505
## Year2010          0.2218      0.1851      1.20      0.2316
## Year2011          0.3180      0.1827      1.74      0.0828 .
## Year2012          0.1861      0.1812      1.03      0.3054
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.626
## Multiple R-squared:  0.0652, Adjusted R-squared:  0.0113
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 302 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.309  0.868  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      3.02e-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 1.775 1      1.332
## Year              1.775 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
## -1.3906 -0.4473 0.0129 0.4066 1.9186
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8374 0.1583 5.29 2.3e-07 ***
## FirstAuthorFemale1 0.0358 0.0725 0.49 0.6216
## Year1997 -0.4084 0.1583 -2.58 0.0103 *
## Year1998 -0.1549 0.2072 -0.75 0.4554
## Year1999 -0.1387 0.2648 -0.52 0.6008
## Year2000 -0.1780 0.8848 -0.20 0.8407
## Year2001 0.5842 0.1905 3.07 0.0024 **
## Year2002 0.3133 0.3260 0.96 0.3373
## Year2003 0.0823 0.2311 0.36 0.7219
## Year2004 0.2475 0.2479 1.00 0.3187
## Year2005 0.4087 0.2770 1.48 0.1411
## Year2006 0.3699 0.1964 1.88 0.0605 .
```

```

## Year2007          0.3270      0.2093      1.56      0.1193
## Year2008          0.5174      0.2228      2.32      0.0209 *
## Year2009          0.2514      0.1780      1.41      0.1589
## Year2010          0.2296      0.1849      1.24      0.2151
## Year2011          0.3205      0.1816      1.77      0.0785 .
## Year2012          0.1942      0.1799      1.08      0.2812
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.635
## Multiple R-squared:  0.0583, Adjusted R-squared:  0.00718
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 31 weights are ~= 1. The remaining 300 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.341  0.871  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      3.02e-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 1.834 1      1.354
## Year      1.834 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.419 -0.437  0.014  0.399  1.958

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.8274    0.1606    5.15 4.6e-07 ***
## LastAuthorFemale1 0.0737    0.0711    1.04 0.3005
## Year1997       -0.3984    0.1606   -2.48 0.0136 *
## Year1998       -0.1503    0.2077   -0.72 0.4698
## Year1999       -0.1444    0.2612   -0.55 0.5808
## Year2000       -0.2073    0.8923   -0.23 0.8165
## Year2001        0.5880    0.1945    3.02 0.0027 **
## Year2002        0.3202    0.3255    0.98 0.3259
## Year2003        0.0738    0.2332    0.32 0.7520
## Year2004        0.2431    0.2478    0.98 0.3274
## Year2005        0.4072    0.2765    1.47 0.1419
## Year2006        0.3666    0.1978    1.85 0.0648 .
## Year2007        0.3246    0.2115    1.53 0.1258
## Year2008        0.5184    0.2251    2.30 0.0219 *
## Year2009        0.2498    0.1801    1.39 0.1664
## Year2010        0.2249    0.1868    1.20 0.2294
## Year2011        0.3143    0.1838    1.71 0.0882 .
## Year2012        0.1928    0.1820    1.06 0.2903
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.63
## Multiple R-squared:  0.0612, Adjusted R-squared:  0.0102
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 303 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.313  0.875   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      3.02e-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: 331"
## [1] ""

```

```

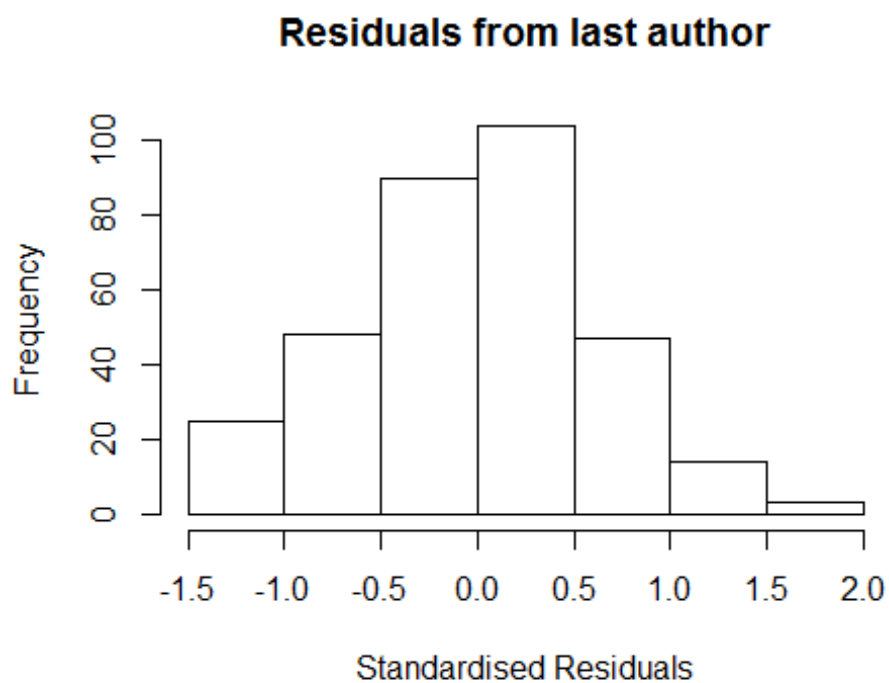
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 3321"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    2    6    2    6    5    6    4    4    5    5    7    7   12   12
## 2011 2012
##   14   12
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    2    4    1    4    2    6    4    3    5    4    7    7   11   12
## 2011 2012
##   12    8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    2    4    1    4    2    5    3    2    5    4    6    6   11   12
## 2011 2012
##   12    8
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.5874010519682"
## [1] "Male first 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 = 24, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.73851050644476"
## [1] "Male last author team size 2018 geometric mean: 1.62238960361098"

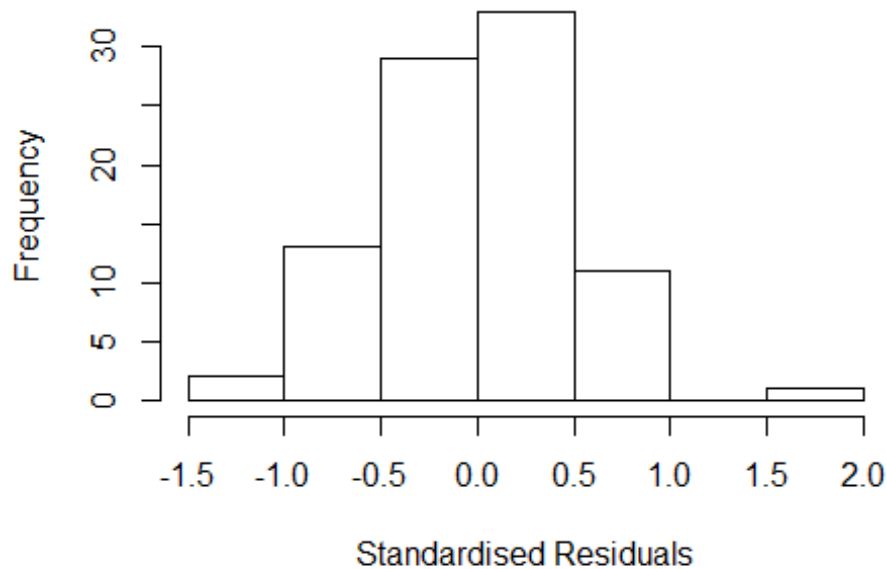
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 30, 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 12.056 1          3.472
## LastAuthorFemale  12.734 1          3.569
## UniqueAuthors     5.462 3          1.327
## Year              63.875 16          1.139
```


Residuals from first and last author and team size



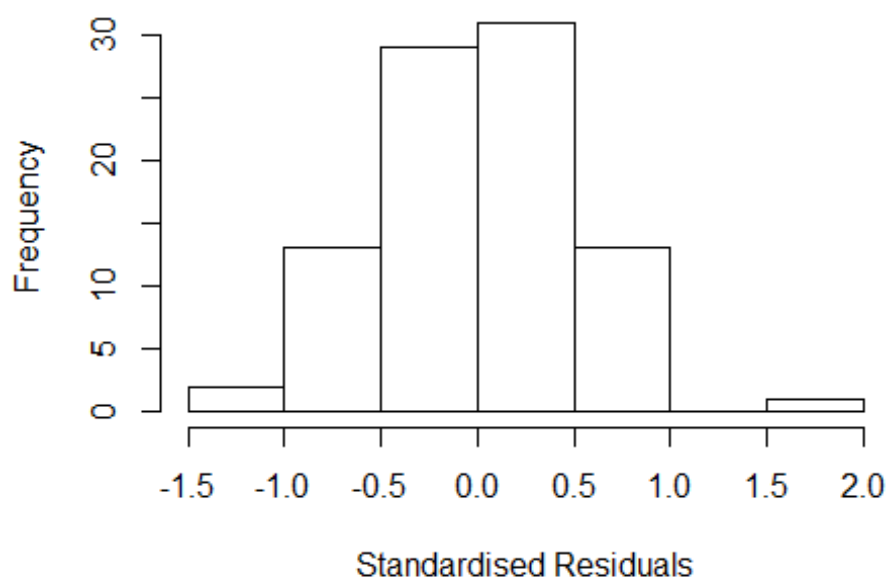
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2824 -0.2915 0.0114 0.3594 1.5256
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6995 0.0931 7.51 1.8e-10 ***
## FirstAuthorFemale1 0.1730 0.2024 0.85 0.3957
## LastAuthorFemale1 -0.0290 0.2083 -0.14 0.8895
## UniqueAuthors2 0.0702 0.1731 0.41 0.6864
## UniqueAuthors3 -0.0600 0.1871 -0.32 0.7495
## UniqueAuthors5 0.2938 0.3504 0.84 0.4048
## Year1997 0.3850 0.8243 0.47 0.6420
## Year1998 0.3753 0.3003 1.25 0.2158
## Year1999 0.8095 0.0931 8.69 1.4e-12 ***
## Year2000 0.0963 0.1975 0.49 0.6275
```

```

## Year2001          0.5815      0.2200      2.64      0.0102 *
## Year2002          0.2480      0.2756      0.90      0.3713
## Year2003          0.1907      0.5953      0.32      0.7497
## Year2004          0.2654      0.3318      0.80      0.4266
## Year2005          0.6308      0.1086      5.81      1.9e-07 ***
## Year2006          0.5309      0.2862      1.86      0.0680 .
## Year2007          0.6551      0.2463      2.66      0.0098 **
## Year2008          0.5642      0.3705      1.52      0.1325
## Year2009          0.4584      0.1838      2.49      0.0151 *
## Year2010          0.3687      0.1575      2.34      0.0222 *
## Year2011          0.4389      0.2064      2.13      0.0371 *
## Year2012          0.8106      0.1404      5.77      2.2e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.565
## Multiple R-squared:  0.149, Adjusted R-squared:  -0.118
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 80 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.446  0.882  0.963  0.924  0.985  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          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 10.88 1      3.299
## LastAuthorFemale  12.29 1      3.506
## Year              14.94 16      1.088

```

Residuals from first and last author



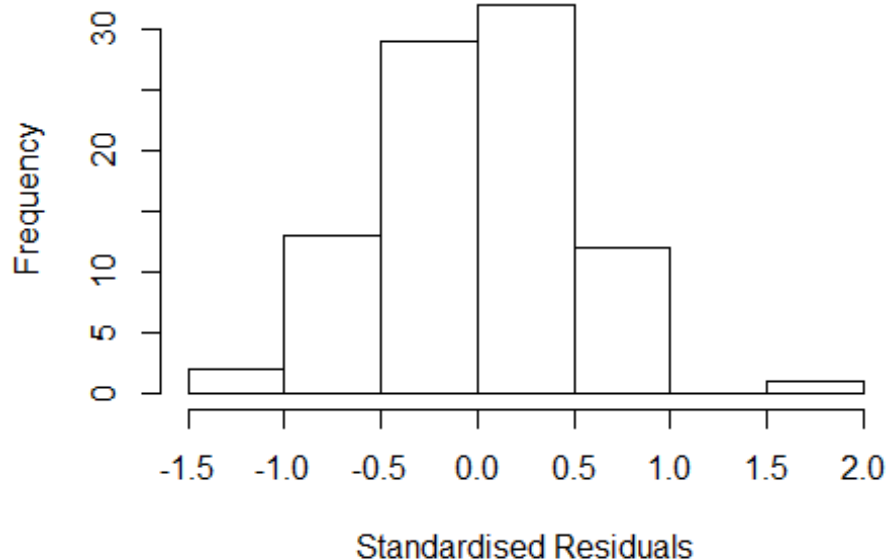
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.281 -0.284 0.022 0.346 1.591
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6981 0.0895 7.80 4.2e-11 ***
## FirstAuthorFemale1 0.1850 0.1942 0.95 0.3441
## LastAuthorFemale1 -0.0381 0.2054 -0.19 0.8533
## Year1997 0.3850 0.8464 0.45 0.6506
## Year1998 0.4075 0.2761 1.48 0.1445
## Year1999 0.8109 0.0895 9.07 2.0e-13 ***
## Year2000 0.0945 0.1769 0.53 0.5951
## Year2001 0.5501 0.2108 2.61 0.0111 *
## Year2002 0.2450 0.2645 0.93 0.3576
## Year2003 0.1965 0.6078 0.32 0.7474
## Year2004 0.3065 0.2847 1.08 0.2854
## Year2005 0.6867 0.0837 8.21 7.5e-12 ***
```

```

## Year2006          0.5353      0.2851      1.88      0.0646 .
## Year2007          0.6430      0.2091      3.08      0.0030 **
## Year2008          0.5568      0.3647      1.53      0.1314
## Year2009          0.4792      0.1703      2.81      0.0063 **
## Year2010          0.3746      0.1606      2.33      0.0226 *
## Year2011          0.4366      0.1977      2.21      0.0305 *
## Year2012          0.8602      0.1569      5.48      6.3e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.551
## Multiple R-squared:  0.142, Adjusted R-squared:  -0.0792
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 79 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.385  0.868  0.964  0.919  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      1.12e-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 4.766 1      2.183
## Year              4.766 16      1.050

```

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
## -1.2844 -0.2842 0.0101 0.3625 1.5863
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6949 0.0856 8.12 9.9e-12 ***
## FirstAuthorFemale1 0.1532 0.1304 1.17 0.2440
## Year1997 0.3850 0.8089 0.48 0.6356
## Year1998 0.4113 0.2709 1.52 0.1334
## Year1999 0.8141 0.0856 9.52 2.6e-14 ***
## Year2000 0.1032 0.1679 0.61 0.5408
## Year2001 0.5469 0.2088 2.62 0.0107 *
## Year2002 0.2291 0.2203 1.04 0.3019
## Year2003 0.1933 0.5945 0.33 0.7459
## Year2004 0.2906 0.2902 1.00 0.3200
## Year2005 0.6873 0.0852 8.07 1.3e-11 ***
## Year2006 0.5349 0.2821 1.90 0.0620 .
```

```

## Year2007          0.6519      0.2007      3.25      0.0018 **
## Year2008          0.5566      0.3559      1.56      0.1223
## Year2009          0.4761      0.1689      2.82      0.0062 **
## Year2010          0.3828      0.1633      2.34      0.0219 *
## Year2011          0.4363      0.1957      2.23      0.0289 *
## Year2012          0.8579      0.1555      5.52      5.3e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.568
## Multiple R-squared:  0.14,   Adjusted R-squared:  -0.0663
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 79 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.416  0.873  0.963   0.923  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      1.12e-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 8.037 1      2.835
## Year      8.037 16      1.067

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q    Median      3Q      Max
## -1.24e+00 -2.59e-01  2.22e-16  3.45e-01  1.56e+00

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.7226    0.0780   9.26 7.8e-14 ***
## LastAuthorFemale1 0.0978    0.1388   0.70 0.4834
## Year1997        0.3850    0.9071   0.42 0.6725
## Year1998        0.4332    0.3245   1.34 0.1861
## Year1999        0.7864    0.0780  10.08 2.5e-15 ***
## Year2000        0.1416    0.1570   0.90 0.3702
## Year2001        0.5746    0.2062   2.79 0.0068 **
## Year2002        0.1667    0.2195   0.76 0.4502
## Year2003        0.1768    0.5720   0.31 0.7582
## Year2004        0.2140    0.3264   0.66 0.5141
## Year2005        0.6820    0.0724   9.42 4.0e-14 ***
## Year2006        0.5103    0.2817   1.81 0.0743 .
## Year2007        0.6652    0.1939   3.43 0.0010 **
## Year2008        0.5405    0.3699   1.46 0.1483
## Year2009        0.4552    0.1609   2.83 0.0061 **
## Year2010        0.3784    0.1625   2.33 0.0227 *
## Year2011        0.4214    0.1909   2.21 0.0305 *
## Year2012        0.8333    0.1405   5.93 9.9e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.553
## Multiple R-squared:  0.135, Adjusted R-squared:  -0.0727
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 77 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.405  0.855  0.962  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      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 3322"
## [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
##    1    2    1    4    1    3    3    3    1    2    4    6    4    8    13
## 2012
##    6
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    1    1    2    0    2    3    3    1    1    3    4    4    5    11
## 2012
##    6
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    1    1    2    0    2    3    3    1    1    3    4    4    5    11
## 2012
##    6
## [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: 1.8346295092848"

## Warning in wilcox.test.default(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: 1.25992104989487"
## [1] "Male last author team size 2018 geometric mean: 1.8346295092848"

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

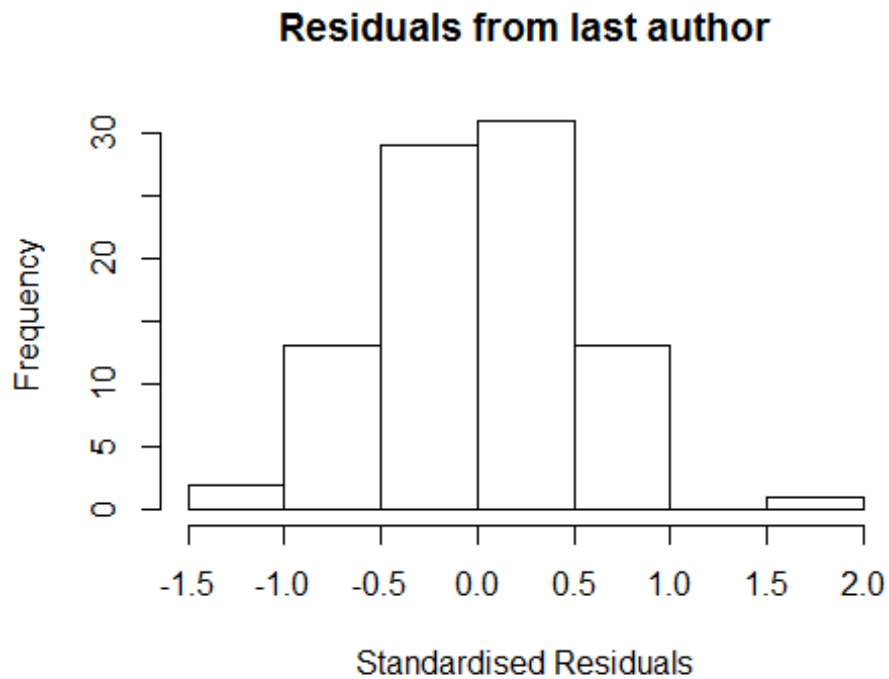
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 9, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"

```



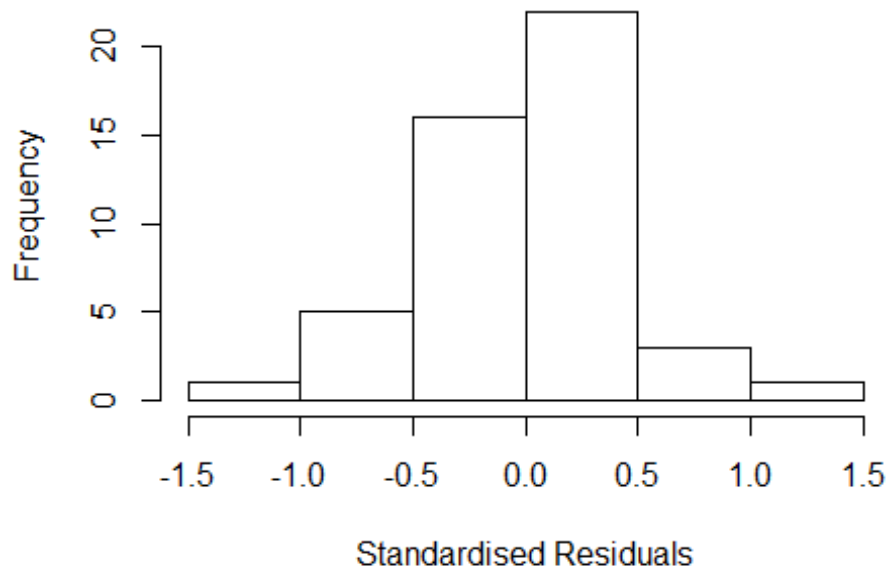
```
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

## 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



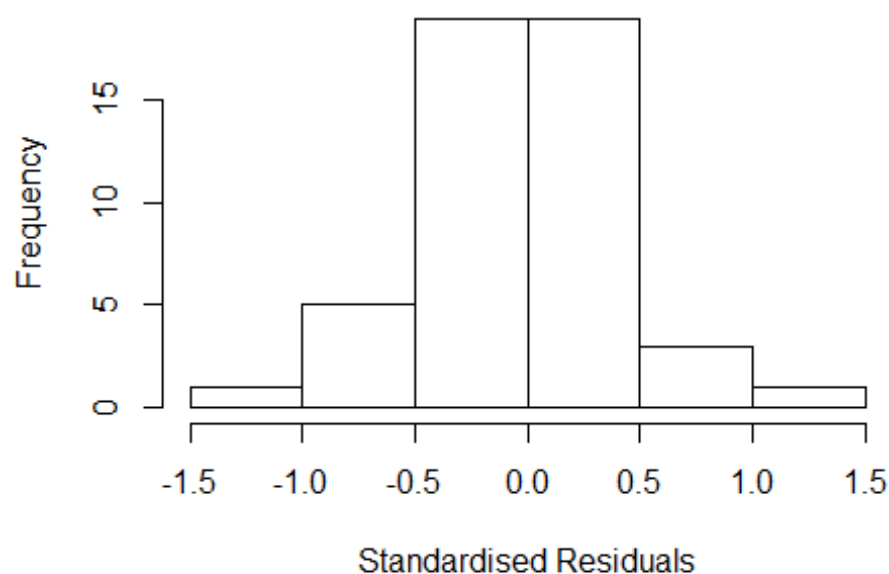
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1987 -0.2711 0.0199 0.2371 1.1267
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.4484 0.2417 1.86 0.07453 .
## FirstAuthorFemale1 0.1923 0.1684 1.14 0.26348
## LastAuthorFemale1 -0.3017 0.2105 -1.43 0.16331
## UniqueAuthors2 -0.1359 0.1775 -0.77 0.45058
## UniqueAuthors3 0.2228 0.2481 0.90 0.37698
## UniqueAuthors4 0.4158 0.1411 2.95 0.00653 **
## UniqueAuthors5 0.1301 0.2204 0.59 0.55989
## Year1998 1.1516 0.2417 4.76 5.7e-05 ***
## Year1999 0.6186 0.2417 2.56 0.01640 *
## Year2000 0.2108 0.6824 0.31 0.75976
```

```

## Year2002          0.7880      0.4720      1.67  0.10658
## Year2003          0.9540      0.2887      3.30  0.00269 **
## Year2004          0.8743      0.2585      3.38  0.00221 **
## Year2005          0.0896      0.2417      0.37  0.71373
## Year2006          0.8630      0.0000      Inf < 2e-16 ***
## Year2007          1.0892      0.3715      2.93  0.00679 **
## Year2008          0.6760      0.2544      2.66  0.01307 *
## Year2009          0.8613      0.7361      1.17  0.25219
## Year2010          0.7377      0.2469      2.99  0.00592 **
## Year2011          0.9688      0.2496      3.88  0.00061 ***
## Year2012          0.8597      0.4488      1.92  0.06608 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.528
## Multiple R-squared:  0.331, Adjusted R-squared:  -0.165
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 37 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.586  0.908  0.974  0.934  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      2.08e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.227e+16 1      1.108e+08
## LastAuthorFemale -2.720e+14 1          NaN
## Year              1.059e+17 14      4.055e+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
## -1.27e+00 -1.60e-01 2.36e-15 2.37e-01 1.11e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 4.43e-01 2.59e-01 1.71e+00 0.09705 .
## FirstAuthorFemale1 1.82e-01 1.91e-01 9.50e-01 0.34780
## LastAuthorFemale1 -2.85e-01 2.34e-01 -1.22e+00 0.23240
## Year1998 1.16e+00 2.59e-01 4.48e+00 9.6e-05 ***
## Year1999 6.24e-01 2.59e-01 2.41e+00 0.02184 *
## Year2000 2.14e-01 8.15e-01 2.60e-01 0.79485
## Year2002 7.88e-01 5.22e-01 1.51e+00 0.14096
## Year2003 9.83e-01 2.16e-01 4.54e+00 8.0e-05 ***
## Year2004 8.81e-01 2.78e-01 3.17e+00 0.00342 **
## Year2005 9.55e-02 2.59e-01 3.70e-01 0.71453
## Year2006 8.63e-01 3.86e-08 2.23e+07 < 2e-16 ***
## Year2007 1.00e+00 2.87e-01 3.49e+00 0.00146 **
```

```

## Year2008          7.40e-01  1.72e-01  4.29e+00  0.00016 ***
## Year2009          8.74e-01  9.53e-01  9.20e-01  0.36657
## Year2010          6.67e-01  2.23e-01  2.98e+00  0.00551 **
## Year2011          9.96e-01  2.22e-01  4.49e+00  9.3e-05 ***
## Year2012          9.34e-01  4.42e-01  2.11e+00  0.04296 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.452
## Multiple R-squared:  0.293, Adjusted R-squared:  -0.0712
## Convergence in 27 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.409  0.886  0.964  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          2.08e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
##  nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##  trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.198e+15  1      3.461e+07
## Year              1.198e+15 14      3.456e+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
## -1.379152 -0.294902  0.000823  0.258759  0.900429
##

```

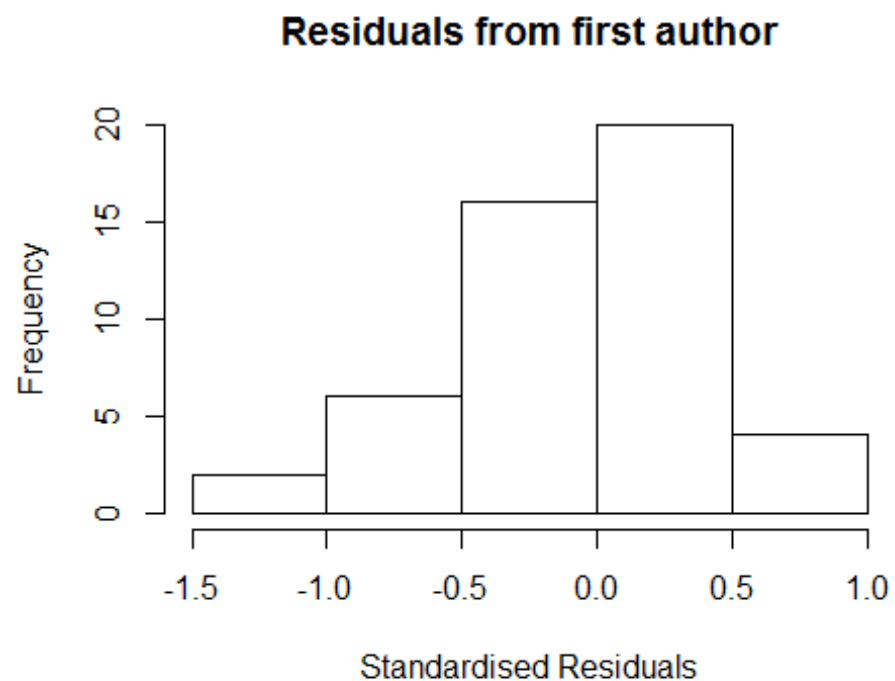
```

## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   3.10e-01   1.82e-01  1.70e+00  0.09851 .
## FirstAuthorFemale1 2.90e-02   1.82e-01  1.60e-01  0.87468
## Year1998       1.29e+00   1.82e-01  7.08e+00  5.0e-08 ***
## Year1999       7.57e-01   1.82e-01  4.15e+00  0.00023 ***
## Year2000       2.80e-01   6.04e-01  4.60e-01  0.64625
## Year2002       7.88e-01   5.04e-01  1.56e+00  0.12810
## Year2003       1.02e+00   2.21e-01  4.64e+00  5.7e-05 ***
## Year2004       9.75e-01   2.85e-01  3.42e+00  0.00175 **
## Year2005       2.28e-01   1.82e-01  1.25e+00  0.21991
## Year2006       8.63e-01   1.68e-08  5.14e+07  < 2e-16 ***
## Year2007       1.15e+00   3.23e-01  3.57e+00  0.00116 **
## Year2008       9.15e-01   6.43e-02  1.42e+01  2.2e-15 ***
## Year2009       1.09e+00   5.51e-01  1.97e+00  0.05725 .
## Year2010       7.97e-01   1.60e-01  4.97e+00  2.1e-05 ***
## Year2011       1.09e+00   1.67e-01  6.49e+00  2.7e-07 ***
## Year2012       1.04e+00   3.74e-01  2.78e+00  0.00898 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.473
## Multiple R-squared:  0.265, Adjusted R-squared:  -0.0802
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.375  0.892  0.956  0.912  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           2.08e-03           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01           5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##           500           50           2           1           1000           200
## trace.lev      mts      compute.rd
##           0           1000           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

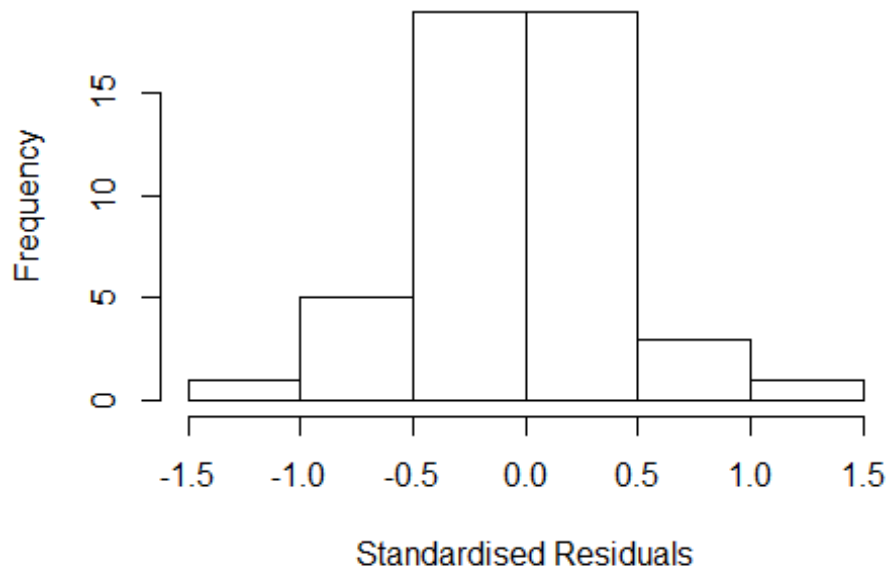
```

```
## 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 14          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
## -1.18732 -0.22687 0.00664 0.21025 1.10974
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5278 0.2352 2.24 0.03190 *
## LastAuthorFemale1 -0.1888 0.2352 -0.80 0.42809
## Year1998 1.0722 0.2352 4.56 7.1e-05 ***
## Year1999 0.5392 0.2352 2.29 0.02863 *
## Year2000 0.1711 0.9080 0.19 0.85173
## Year2002 0.7880 0.5077 1.55 0.13049
## Year2003 0.8659 0.1751 4.95 2.3e-05 ***
## Year2004 0.8203 0.2344 3.50 0.00139 **
## Year2005 0.0102 0.2352 0.04 0.96574
## Year2006 0.8630 0.0000 Inf < 2e-16 ***
## Year2007 1.0020 0.2954 3.39 0.00186 **
## Year2008 0.7667 0.1795 4.27 0.00016 ***
```

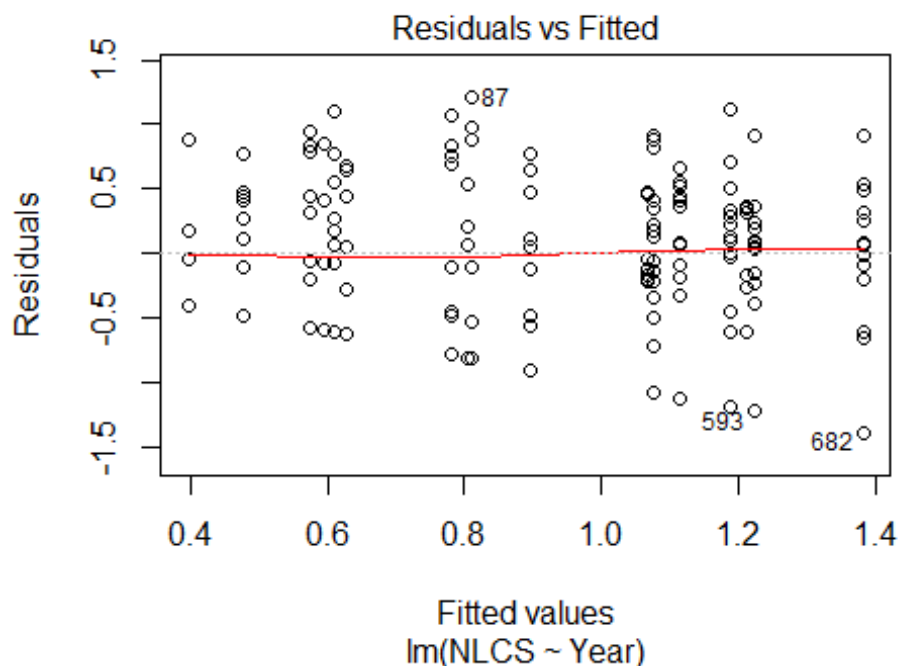


```

## Year2009          0.8783      0.8108      1.08  0.28682
## Year2010          0.6296      0.1998      3.15  0.00352 **
## Year2011          0.9318      0.2340      3.98  0.00037 ***
## Year2012          0.8483      0.4173      2.03  0.05046 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.469
## Multiple R-squared:  0.275, Adjusted R-squared:  -0.0646
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 37 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.501  0.887  0.948  0.909  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      2.08e-03      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 48"
## [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
##   19   24   21   19   26   33   34   16   18   15   38   31   44   46   54
## 2011 2012
##   39   26
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   13   12    7    8    8   13   10    6    5    4    8    9   17   17   15
## 2011 2012
##   12   15
##

```

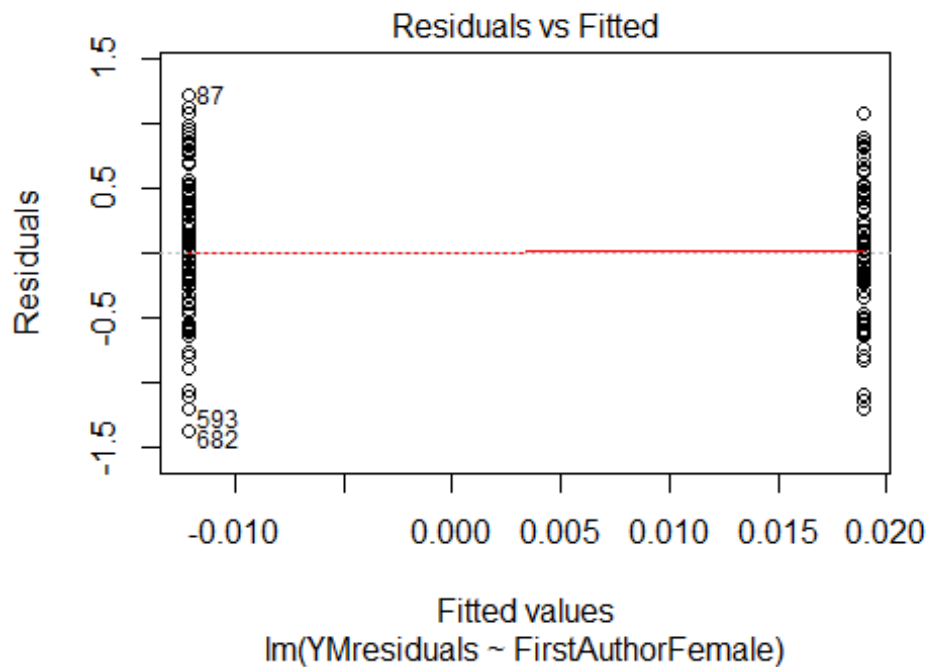
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 13 12 5 8 6 12 9 4 4 3 5 8 17 15 14
## 2011 2012
## 11 15
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
##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



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.05, df = 1, p-value = 0.8
## [1] "Female first author team size 2018 geometric mean: 4.30886938006377"
## [1] "Male first author team size 2018 geometric mean: 3.72791927319135"
## Warning in wilcox.test.default(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 = 0.4
## 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: 3.46410161513775"

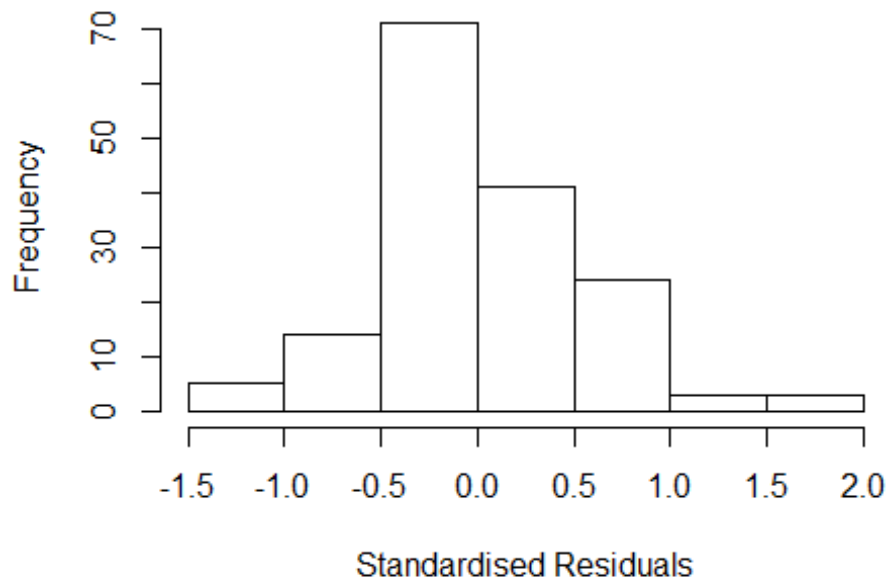
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 9, 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 | 2.751 | 1 | 1.659 |
| LastAuthorFemale | 2.509 | 1 | 1.584 |
| UniqueAuthors | 24.316 | 4 | 1.490 |
| Year | 71.698 | 16 | 1.143 |

Residuals from first and last author and team size



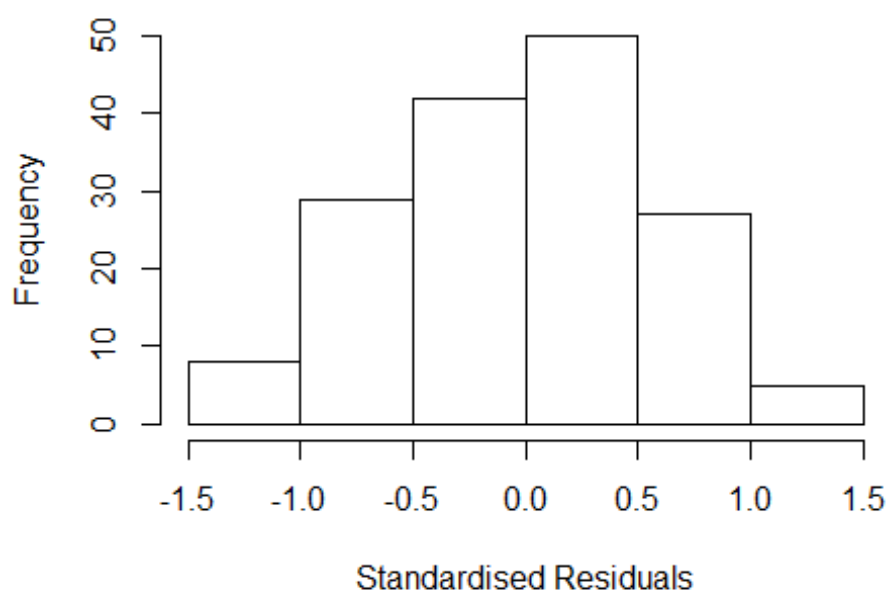
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3595 -0.2824 -0.0401 0.3913 1.7817
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.27594 0.13657 2.02 0.04527 *
## FirstAuthorFemale1 0.02517 0.10618 0.24 0.81295
## LastAuthorFemale1 -0.03495 0.11208 -0.31 0.75565
## UniqueAuthors2 0.53598 0.27442 1.95 0.05283 .
## UniqueAuthors3 0.72388 0.17195 4.21 4.6e-05 ***
## UniqueAuthors4 0.65241 0.17836 3.66 0.00036 ***
## UniqueAuthors5 0.77183 0.15222 5.07 1.3e-06 ***
## Year1997 0.04335 0.16802 0.26 0.79679
## Year1998 -0.23582 0.14484 -1.63 0.10577
## Year1999 -0.03866 0.44340 -0.09 0.93065
```

```

## Year2000          0.11824    0.30903    0.38  0.70259
## Year2001          0.03061    0.16428    0.19  0.85246
## Year2002          0.00229    0.38415    0.01  0.99526
## Year2003          0.20054    0.27982    0.72  0.47478
## Year2004         -0.08382    0.16584   -0.51  0.61405
## Year2005          0.00337    0.19383    0.02  0.98616
## Year2006          0.11583    0.28772    0.40  0.68787
## Year2007          0.01383    0.27420    0.05  0.95985
## Year2008          0.33374    0.19620    1.70  0.09118 .
## Year2009          0.36945    0.24983    1.48  0.14146
## Year2010          0.29031    0.18202    1.59  0.11301
## Year2011          0.28997    0.19356    1.50  0.13640
## Year2012          0.40299    0.22096    1.82  0.07034 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.469
## Multiple R-squared:  0.454, Adjusted R-squared:  0.367
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 145 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.118  0.855  0.958  0.889  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      6.21e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.697 1      1.303
## LastAuthorFemale  1.774 1      1.332
## Year              2.655 16      1.031

```

Residuals from first and last author



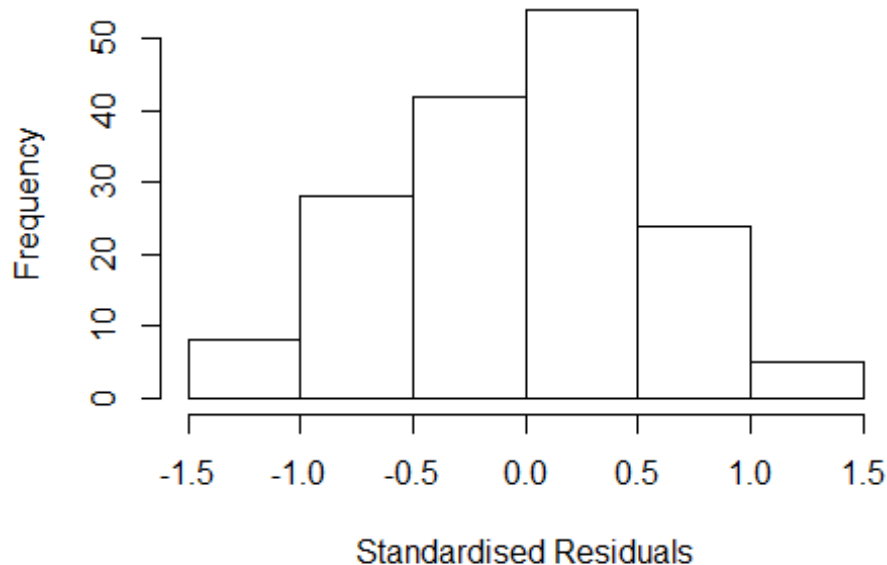
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2796 -0.4811 0.0285 0.4246 1.2518
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5661 0.1936 2.92 0.00402 **
## FirstAuthorFemale1 0.0689 0.1111 0.62 0.53621
## LastAuthorFemale1 -0.1233 0.1326 -0.93 0.35410
## Year1997 -0.0851 0.2340 -0.36 0.71669
## Year1998 -0.3814 0.2217 -1.72 0.08762 .
## Year1999 0.2010 0.4459 0.45 0.65275
## Year2000 0.0336 0.2965 0.11 0.90984
## Year2001 0.0498 0.2686 0.19 0.85304
## Year2002 0.2737 0.3879 0.71 0.48165
## Year2003 0.5935 0.2881 2.06 0.04119 *
## Year2004 -0.1670 0.2912 -0.57 0.56737
## Year2005 0.2154 0.4034 0.53 0.59414
```

```

## Year2006          0.5136      0.2184      2.35  0.02006 *
## Year2007          0.3360      0.3086      1.09  0.27815
## Year2008          0.5133      0.2435      2.11  0.03678 *
## Year2009          0.5752      0.2588      2.22  0.02781 *
## Year2010          0.6014      0.2458      2.45  0.01565 *
## Year2011          0.6952      0.2303      3.02  0.00301 **
## Year2012          0.8367      0.2372      3.53  0.00057 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.608
## Multiple R-squared:  0.249, Adjusted R-squared:  0.154
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 144 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.637  0.887   0.943   0.917   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      6.21e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
##      nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
##      trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.588 1      1.260
## Year              1.588 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
## -1.3991 -0.4665 0.0342 0.3903 1.3097
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5391 0.1869 2.88 0.00452 **
## FirstAuthorFemale1 0.0327 0.1102 0.30 0.76738
## Year1997 -0.0726 0.2303 -0.32 0.75316
## Year1998 -0.3544 0.2160 -1.64 0.10302
## Year1999 0.1702 0.4609 0.37 0.71246
## Year2000 0.0346 0.3093 0.11 0.91107
## Year2001 0.0521 0.2642 0.20 0.84409
## Year2002 0.2788 0.3855 0.72 0.47078
## Year2003 0.6301 0.2870 2.20 0.02973 *
## Year2004 -0.1805 0.2846 -0.63 0.52704
## Year2005 0.2046 0.4007 0.51 0.61049
## Year2006 0.4898 0.2118 2.31 0.02219 *
```



```

## Year2007          0.3318      0.2926      1.13  0.25874
## Year2008          0.5371      0.2400      2.24  0.02677 *
## Year2009          0.5973      0.2525      2.37  0.01937 *
## Year2010          0.6118      0.2389      2.56  0.01149 *
## Year2011          0.6890      0.2302      2.99  0.00326 **
## Year2012          0.8600      0.2317      3.71  0.00029 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.598
## Multiple R-squared:  0.249, Adjusted R-squared:  0.16
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 147 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.564  0.876  0.943  0.915  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      6.21e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.659 1      1.288
## Year      1.659 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
## -1.3423 -0.4814  0.0206  0.4004  1.2391

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.5799    0.1905   3.04 0.00279 **
## LastAuthorFemale1 -0.0985    0.1273  -0.77 0.44016
## Year1997          -0.0918    0.2332  -0.39 0.69454
## Year1998          -0.3952    0.2191  -1.80 0.07333 .
## Year1999           0.2000    0.4539   0.44 0.66013
## Year2000           0.0458    0.2985   0.15 0.87839
## Year2001           0.0410    0.2669   0.15 0.87823
## Year2002           0.2615    0.3834   0.68 0.49641
## Year2003           0.5977    0.2964   2.02 0.04563 *
## Year2004          -0.1609    0.2949  -0.55 0.58622
## Year2005           0.2162    0.4196   0.52 0.60713
## Year2006           0.5254    0.2200   2.39 0.01826 *
## Year2007           0.3423    0.3064   1.12 0.26592
## Year2008           0.5285    0.2417   2.19 0.03043 *
## Year2009           0.5923    0.2548   2.32 0.02149 *
## Year2010           0.6264    0.2410   2.60 0.01032 *
## Year2011           0.7054    0.2327   3.03 0.00289 **
## Year2012           0.8609    0.2326   3.70 0.00031 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.597
## Multiple R-squared:  0.25, Adjusted R-squared:  0.161
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 148 ones are summarized as
##   Min. 1st Qu. Median   Mean 3rd Qu.    Max.
## 0.593 0.885 0.941 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      6.21e-04      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 161"
## [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]"
##
## 1997 1999 2004 2006 2007 2009
##    1    1    1    1    2    1
##
## 1997 1999 2004 2006 2007 2009
##    0    0    0    0    0    0
##
## 1997 1999 2004 2006 2007 2009
##    0    0    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: 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 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 2004 2005 2006 2007 2008 2009 2010 2011
##    1    2    2    3    3    3    1    1    6    3    4    5    4    10    8
## 2012
##    4
##
## 1996 1997 1998 1999 2000 2001 2002 2004 2005 2006 2007 2008 2009 2010 2011
##    0    0    0    0    0    0    0    0    1    0    0    0    0    0    1
## 2012
##    0
##
## 1996 1997 1998 1999 2000 2001 2002 2004 2005 2006 2007 2008 2009 2010 2011
##    0    0    0    0    0    0    0    0    1    0    0    0    0    0    0
## 2012
##    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: 5"

```

[illegible]

```

## 2011 2012
##      3      2
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 5.95789213552895"
## [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 = 7.5, 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: 4.78938895318699"

## Warning in wilcox.test.default(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 = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
## [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 3404"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      1      7      5      3      2      4      3      2      2      6      6      4      4      5      3
## 2011 2012
##      4      6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010

```

```

##      0      1      2      2      0      0      0      0      0      1      0      0      0      1      0
## 2011 2012
##      1      0
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      0      1      2      2      0      0      0      0      0      1      0      0      0      1      0
## 2011 2012
##      0      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: 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 = 0.5, p-value = NA
## 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"

## Warning in wilcox.test.default(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: 7"
## [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]"
##

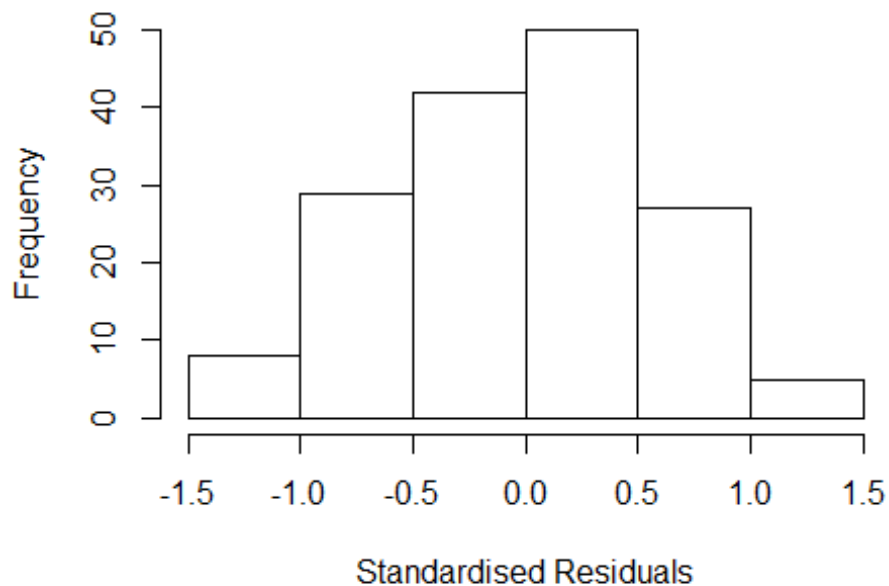
```

```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      4      8      4      4      8      5      5      6     10     10     16     12     12     12     10
## 2011 2012
##     16     10
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      1      1      1      1      2      1      2      2      4      5      9      8      6      5      2
## 2011 2012
##     11      7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      1      1      1      0      2      1      2      1      2      4      4      8      4      5      2
## 2011 2012
##      7      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: NaN"
## [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

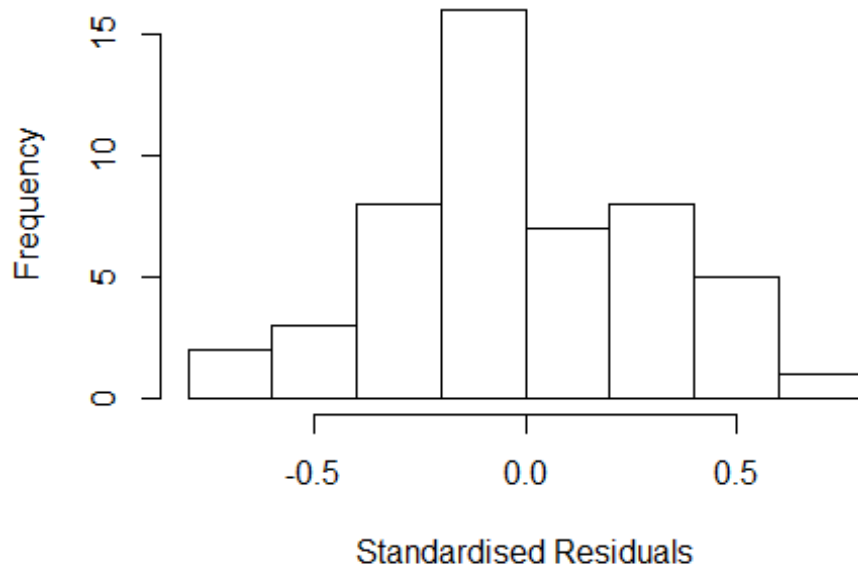
## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

Residuals from last author



| ## | | GVI | Df | $GVI^{1/(2 \cdot Df)}$ |
|----|-------------------|-----|----|------------------------|
| ## | FirstAuthorFemale | NaN | 1 | NaN |
| ## | LastAuthorFemale | NaN | 1 | NaN |
| ## | UniqueAuthors | NaN | 4 | NaN |
| ## | Year | NaN | 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
## -7.46e-01 -2.01e-01 -1.73e-15 2.10e-01 6.67e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3805 0.2388 5.78 3.3e-06 ***
## FirstAuthorFemale1 0.0471 0.1126 0.42 0.67866
## LastAuthorFemale1 -0.3401 0.1179 -2.88 0.00747 **
## UniqueAuthors2 0.6132 0.2054 2.99 0.00582 **
```



```

## UniqueAuthors3      0.6480      0.2241      2.89  0.00733 **
## UniqueAuthors4      0.3088      0.2087      1.48  0.15004
## UniqueAuthors5     -0.1114      0.1974     -0.56  0.57691
## Year1997             0.1709      0.1179      1.45  0.15847
## Year1998            -1.6137      0.2639     -6.11  1.3e-06 ***
## Year2000             0.0189      0.2165      0.09  0.93107
## Year2001             0.3680      0.0000      Inf < 2e-16 ***
## Year2002            -1.1932      0.2500     -4.77  5.2e-05 ***
## Year2003            -0.2254      0.2216     -1.02  0.31785
## Year2004            -0.9340      0.3155     -2.96  0.00620 **
## Year2005            -0.7880      0.3262     -2.42  0.02249 *
## Year2006            -0.9676      0.3702     -2.61  0.01425 *
## Year2007            -0.6306      0.1142     -5.52  6.7e-06 ***
## Year2008            -0.4965      0.1597     -3.11  0.00428 **
## Year2009            -1.0062      0.2411     -4.17  0.00026 ***
## Year2010            -0.7619      0.2298     -3.32  0.00254 **
## Year2011            -0.7538      0.2529     -2.98  0.00589 **
## Year2012            -0.8538      0.2145     -3.98  0.00044 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.405
## Multiple R-squared:  0.595, Adjusted R-squared:  0.292
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 43 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.714  0.920  0.973  0.942  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      2.00e-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 -4.402e+15  1          NaN

```

```

## LastAuthorFemale -4.628e+15 1 NaN
## Year -6.939e+29 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
## -1.12e+00 -2.75e-01 -3.05e-16 2.85e-01 7.90e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.29e+00 1.78e-01 7.29e+00 2.7e-08 ***
## FirstAuthorFemale1 -1.25e-02 1.30e-01 -1.00e-01 0.92399
## LastAuthorFemale1 -3.66e-01 1.78e-01 -2.06e+00 0.04750 *
## Year1997 1.45e-01 1.78e-01 8.20e-01 0.41983
## Year1998 -8.20e-01 2.13e-01 -3.85e+00 0.00054 ***
## Year2000 4.24e-01 4.69e-01 9.00e-01 0.37285
## Year2001 3.68e-01 2.19e-08 1.68e+07 < 2e-16 ***
## Year2002 -4.77e-01 2.61e-01 -1.83e+00 0.07721 .
## Year2003 1.69e-01 1.78e-01 9.50e-01 0.34802
## Year2004 -1.87e-01 2.49e-01 -7.50e-01 0.45718
## Year2005 -5.15e-01 2.43e-01 -2.12e+00 0.04191 *
## Year2006 -5.29e-01 3.81e-01 -1.39e+00 0.17542
## Year2007 -3.81e-01 1.76e-01 -2.16e+00 0.03833 *
## Year2008 -2.75e-01 2.17e-01 -1.27e+00 0.21370
## Year2009 -6.42e-01 2.13e-01 -3.01e+00 0.00503 **
## Year2010 -3.78e-01 2.10e-01 -1.80e+00 0.08154 .
## Year2011 -1.75e-01 3.69e-01 -4.70e-01 0.63890
## Year2012 -4.35e-01 1.86e-01 -2.34e+00 0.02567 *
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.53
## Multiple R-squared: 0.345, Adjusted R-squared: -0.00335
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~1. The remaining 40 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.634 0.922 0.968 0.939 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

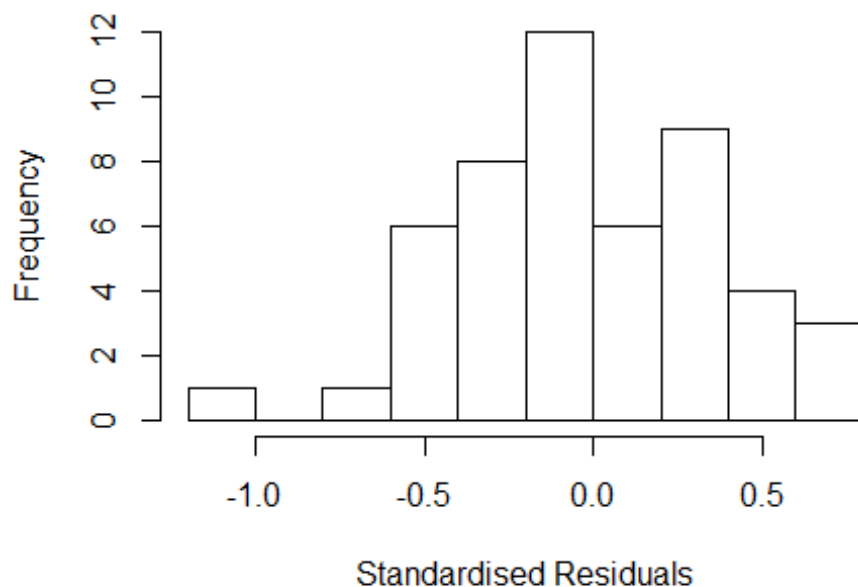
```

```
##          rel.tol          solve.tol          eps.outlier          eps.x
##          1.00e-07          1.00e-07          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

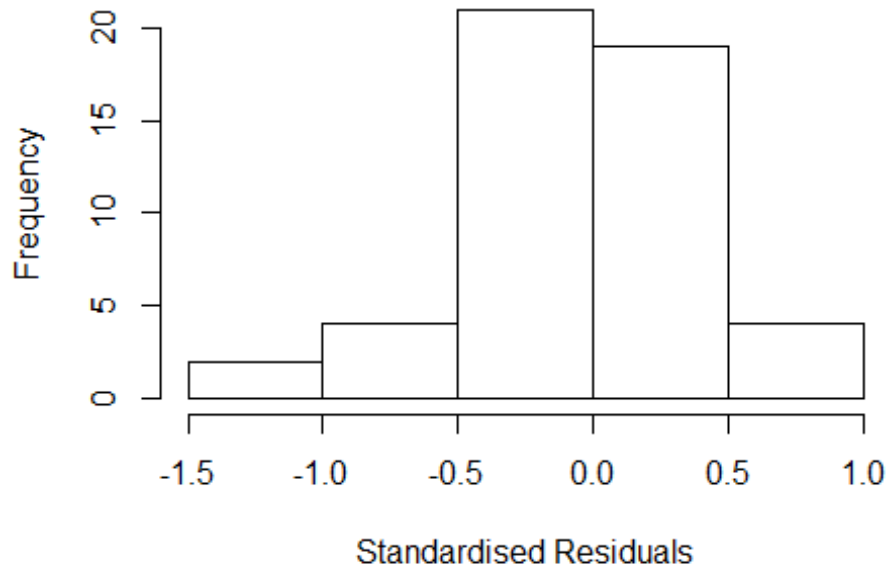
## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

Residuals from first and last author



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

Residuals from first author



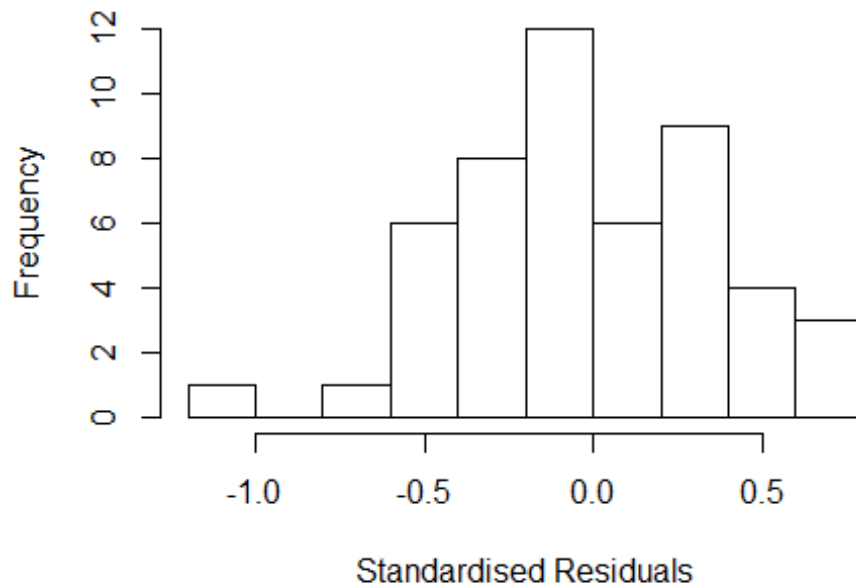
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.11e+00 -3.23e-01 -1.39e-16 3.01e-01 9.84e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 9.29e-01 0.00e+00 Inf <2e-16 ***
## FirstAuthorFemale1 -9.85e-02 1.54e-01 -6.40e-01 0.527
## Year1997 5.11e-01 0.00e+00 Inf <2e-16 ***
## Year1998 -3.68e-01 1.54e-01 -2.39e+00 0.023 *
## Year2000 6.07e-01 2.66e-01 2.28e+00 0.029 *
## Year2001 3.68e-01 1.05e-08 3.50e+07 <2e-16 ***
## Year2002 -1.11e-01 1.92e-01 -5.80e-01 0.567
## Year2003 5.35e-01 7.06e-09 7.57e+07 <2e-16 ***
## Year2004 -4.50e-03 8.97e-02 -5.00e-02 0.960
## Year2005 -1.25e-01 1.65e-01 -7.60e-01 0.452
## Year2006 -1.37e-01 3.52e-01 -3.90e-01 0.700
## Year2007 -8.52e-02 1.90e-01 -4.50e-01 0.657
```

```

## Year2008          3.82e-02   2.11e-01   1.80e-01   0.858
## Year2009          -4.70e-01   2.81e-01  -1.67e+00   0.104
## Year2010          3.08e-02   1.03e-01   3.00e-01   0.767
## Year2011          1.76e-01   3.03e-01   5.80e-01   0.564
## Year2012          -2.62e-01   1.75e-01  -1.50e+00   0.144
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.514
## Multiple R-squared:  0.292, Adjusted R-squared:  -0.0519
## 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.
##  0.623  0.926  0.956  0.932  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 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale -9.041e+13  1      NaN
## Year             -9.041e+13 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
## -1.12e+00 -2.75e-01 -5.41e-16 2.80e-01 7.98e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.30e+00 1.87e-01 6.96e+00 5.9e-08 ***
## LastAuthorFemale1 -3.70e-01 1.87e-01 -1.98e+00 0.05597 .
## Year1997 1.41e-01 1.87e-01 7.60e-01 0.45471
## Year1998 -8.37e-01 1.87e-01 -4.48e+00 8.4e-05 ***
## Year2000 4.22e-01 4.83e-01 8.70e-01 0.38848
## Year2001 3.68e-01 4.25e-08 8.65e+06 < 2e-16 ***
## Year2002 -4.81e-01 2.68e-01 -1.79e+00 0.08205 .
## Year2003 1.65e-01 1.87e-01 8.90e-01 0.38251
## Year2004 -1.89e-01 2.54e-01 -7.50e-01 0.46053
## Year2005 -5.22e-01 2.41e-01 -2.17e+00 0.03770 *
## Year2006 -5.38e-01 3.73e-01 -1.44e+00 0.15853
## Year2007 -3.87e-01 1.74e-01 -2.22e+00 0.03321 *
```

```

## Year2008          -2.85e-01    1.89e-01 -1.51e+00    0.14183
## Year2009          -6.54e-01    1.73e-01 -3.78e+00    0.00062 ***
## Year2010          -3.88e-01    2.14e-01 -1.82e+00    0.07858 .
## Year2011          -1.79e-01    3.63e-01 -4.90e-01    0.62503
## Year2012          -4.39e-01    1.80e-01 -2.44e+00    0.02046 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.507
## Multiple R-squared:  0.349, Adjusted R-squared:  0.0327
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 41 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.605  0.917  0.966  0.936  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      2.00e-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 3501"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 2007 2008 2009 2011 2012
##    1    1    1    2    2
##
## 2007 2008 2009 2011 2012
##    0    0    0    0    0
##
## 2007 2008 2009 2011 2012
##    0    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 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 3503"
## [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 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    5    3    1    3    4    6    7    3    1    4    1    6    4    3
## 2011 2012
##   10    4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    1    1    0    1    2    5    5    2    0    3    1    5    0    3
## 2011 2012
##    6    2
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    0    1    0    1    2    5    2    0    0    2    1    3    0    2
## 2011 2012
##    5    1
## [1] "Heteroscedasticity checks, confirming that there are problems with

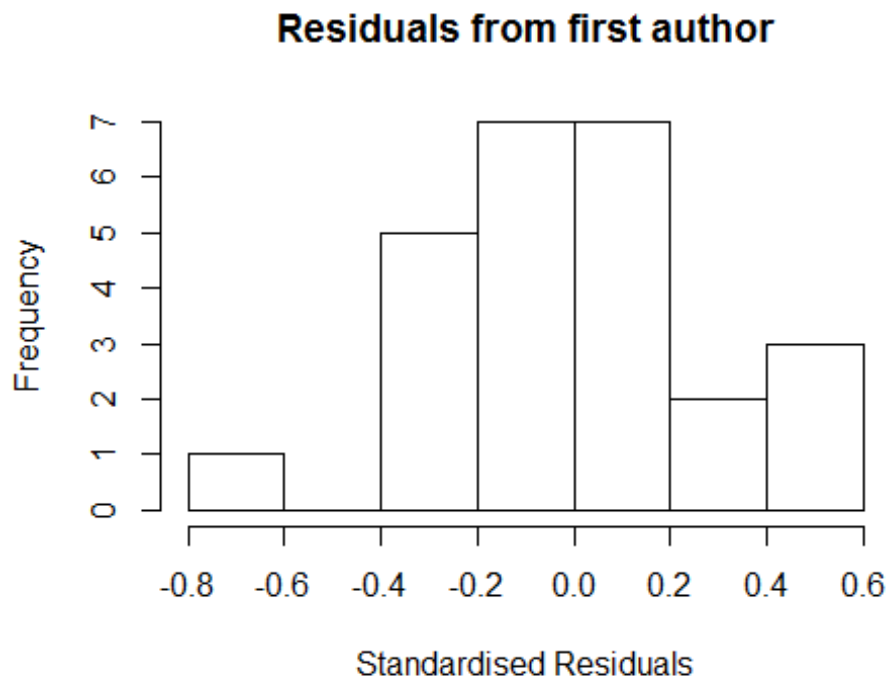
```



```

these"
## [1] "Female first author team size 2018 geometric mean: 4.93242414866094"
## [1] "Male first author team size 2018 geometric mean: 1"
##
## 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: NaN"
## [1] "Male last author team size 2018 geometric mean: 3.30975091964687"
## [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 34.11 1 5.840
## Year 34.11 10 1.193

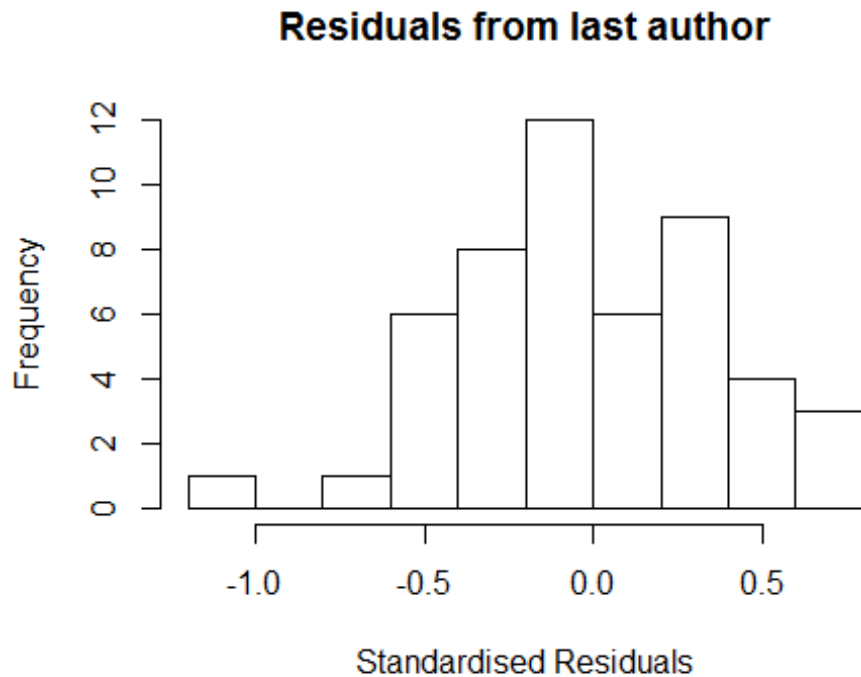
```



```

## [1] "Regression 4: Last author gender, Year as factors"
##
## GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 49.75 1 7.053
## Year 49.75 10 1.216

```



```
## [1] "Sample size for the above analysis: 25"
## [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]"
##
## 1997 1999 2000 2001 2002 2003 2008 2012
##    1    1    1    3    1    1    1    2
##
## 1997 1999 2000 2001 2002 2003 2008 2012
##    0    0    0    2    1    0    1    1
##
## 1997 1999 2000 2001 2002 2003 2008 2012
##    0    0    0    1    1    0    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: 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 3506"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 2011 2012
##    1    1    2    1
##
## 1996 1997 2011 2012
##    0    0    1    1
##
## 1996 1997 2011 2012
##    0    0    1    0
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 6.32455532033676"
## [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]"
##
## 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009
##    2    1    1    3    3    2    2    3    2    4    2
##
## 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009
##    1    1    1    2    2    2    2    2    2    2    2
##
## 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009
##    1    0    1    1    2    1    2    2    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: 16"
## [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]"
##
## 2000 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    2    5    2    5    2    5    1    5    7
##
## 2000 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    0    2    4    1    5    1    4    1    4    7
##
## 2000 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    0    1    2    1    5    1    3    1    4    7
## [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: 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 2.940e+00 1          1.715
## LastAuthorFemale -1.821e+15 1          NaN
## Year              -3.503e+15 8          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
## -1.36e+00 -2.09e-01 -2.78e-17  2.44e-01  1.01e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    5.86e-01  1.51e-01  3.87e+00  0.0017 **

```

```

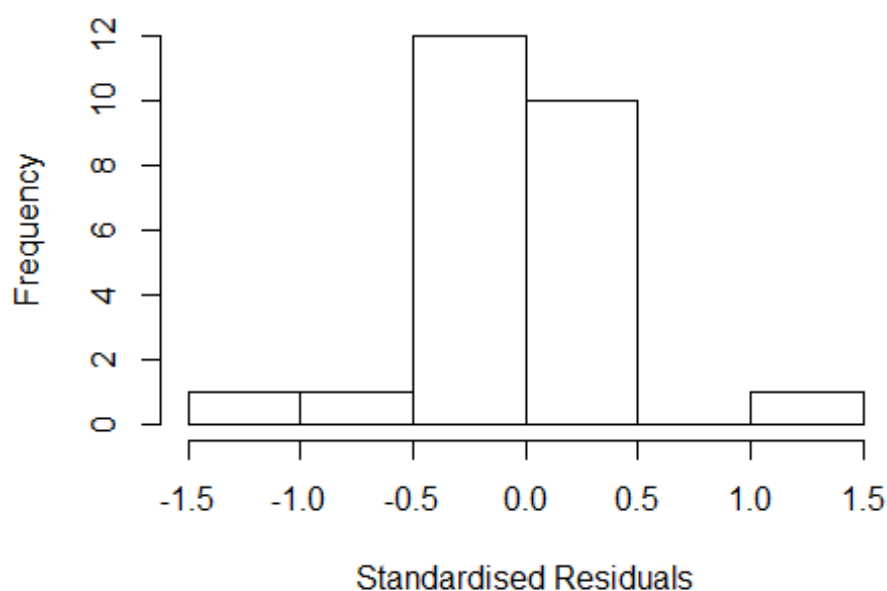
## FirstAuthorFemale1  2.45e-01  2.04e-01  1.20e+00  0.2496
## LastAuthorFemale1  2.81e-01  1.35e-01  2.08e+00  0.0569 .
## Year2005           5.40e-01  1.83e-01  2.95e+00  0.0105 *
## Year2006           5.19e-01  1.35e-01  3.83e+00  0.0018 **
## Year2007           2.47e-01  3.78e-01  6.50e-01  0.5242
## Year2008          -2.27e-01  1.35e-01 -1.68e+00  0.1153
## Year2009           6.52e-02  3.52e-01  1.90e-01  0.8557
## Year2010          -7.94e-01  1.27e-08 -6.25e+07  <2e-16 ***
## Year2011           3.23e-01  1.37e-01  2.36e+00  0.0333 *
## Year2012          -3.75e-01  2.50e-01 -1.50e+00  0.1552
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.484
## Multiple R-squared:  0.45, Adjusted R-squared:  0.0569
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 21 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.411  0.949  0.971  0.924  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      4.00e-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 8             NaN

## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.39e+00 -2.58e-01 -1.11e-16  2.27e-01  1.09e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.732      0.207   3.54  0.0029 **
## FirstAuthorFemale1 0.380      0.207   1.84  0.0859 .
## Year2005          0.467      0.152   3.08  0.0076 **
## Year2006          0.238      0.000    Inf <2e-16 ***
## Year2007          0.281      0.421   0.67  0.5144
## Year2008         -0.508      0.000   -Inf <2e-16 ***
## Year2009         -0.120      0.297  -0.40  0.6914
## Year2010         -0.794      0.000   -Inf <2e-16 ***
```

```

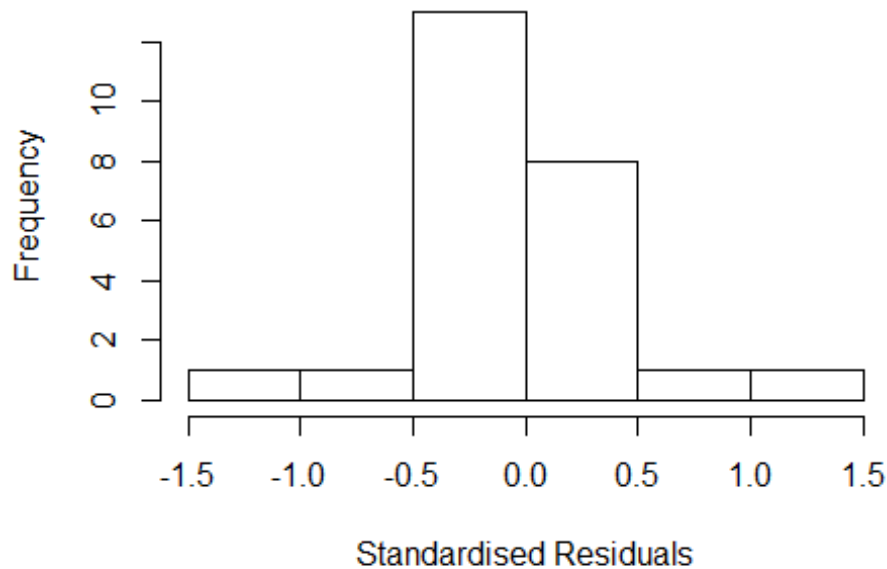
## Year2011          0.285      0.144      1.99      0.0654 .
## Year2012         -0.458      0.234     -1.96      0.0688 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.407
## Multiple R-squared:  0.498, Adjusted R-squared:  0.196
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 21 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.217  0.912  0.961  0.891  0.987  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      4.00e-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

```

Residuals from first author



```
##          GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale NaN 1          NaN
## Year            NaN 8          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
## -1.36e+00 -1.70e-01  9.99e-16  1.50e-01  1.01e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.7383    0.1586    4.65  0.00031 ***
## LastAuthorFemale1 0.3737    0.1586    2.36  0.03248 *
## Year2005        0.4639    0.1343    3.45  0.00355 **
## Year2006        0.6117    0.1586    3.86  0.00155 **
## Year2007        0.2521    0.2394    1.05  0.30896
## Year2008       -0.1343    0.1586   -0.85  0.41053
## Year2009        0.0162    0.4384    0.04  0.97101
## Year2010       -0.7940    0.0000   -Inf < 2e-16 ***
```



```

## Year2011          0.2843      0.1436      1.98  0.06639 .
## Year2012          -0.3736      0.2336     -1.60  0.13065
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.428
## Multiple R-squared:  0.468, Adjusted R-squared:  0.149
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 21 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.289  0.921  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           4.00e-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: 25"
## [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]"
## < 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: 4"
## [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 3603"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 2002 2010 2012
##    1    1    1
##
## 2002 2010 2012
##    1    1    1
##
## 2002 2010 2012
##    0    1    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: 2"
##
## 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: 2"
##
## 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"
## [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 3604"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##

```

```

## 2011
## 1
##
## 2011
## 0
##
## 2011
## 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 3605"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 2001 2002 2004 2006 2008 2010 2011 2012
## 1 1 1 2 4 1 1 1
##
## 2001 2002 2004 2006 2008 2010 2011 2012
## 0 0 0 1 4 1 1 0
##
## 2001 2002 2004 2006 2008 2010 2011 2012
## 0 0 0 1 2 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: 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: 5"
## [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]"
##
## 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    3    3    3    4    3    2    3    3    1    1    3    3    1
##
## 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    0    1    0    0    1    3    2    1    3    1    0    3    3    1
##
## 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    0    1    0    0    0    2    2    1    3    1    0    3    3    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"
## [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 3608"
## [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 3609"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1997 1998 1999 2000 2001 2004 2005 2007 2008 2009 2010 2011 2012
##    1    1    3    1    1    1    1    3    3    12    11    11    7

```

```

##
## 1997 1998 1999 2000 2001 2004 2005 2007 2008 2009 2010 2011 2012
##    1    1    1    0    0    1    1    2    2    3    7    6    3
##
## 1997 1998 1999 2000 2001 2004 2005 2007 2008 2009 2010 2011 2012
##    1    1    1    0    0    1    1    2    2    2    3    6    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"
## [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 3610"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 2006 2007 2008 2011 2012
##    1    1    1    2    1
##
## 2006 2007 2008 2011 2012
##    1    1    1    2    0
##
## 2006 2007 2008 2011 2012
##    1    1    1    2    0
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.63214802590498"
## [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: 5"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 3611"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last

```

```

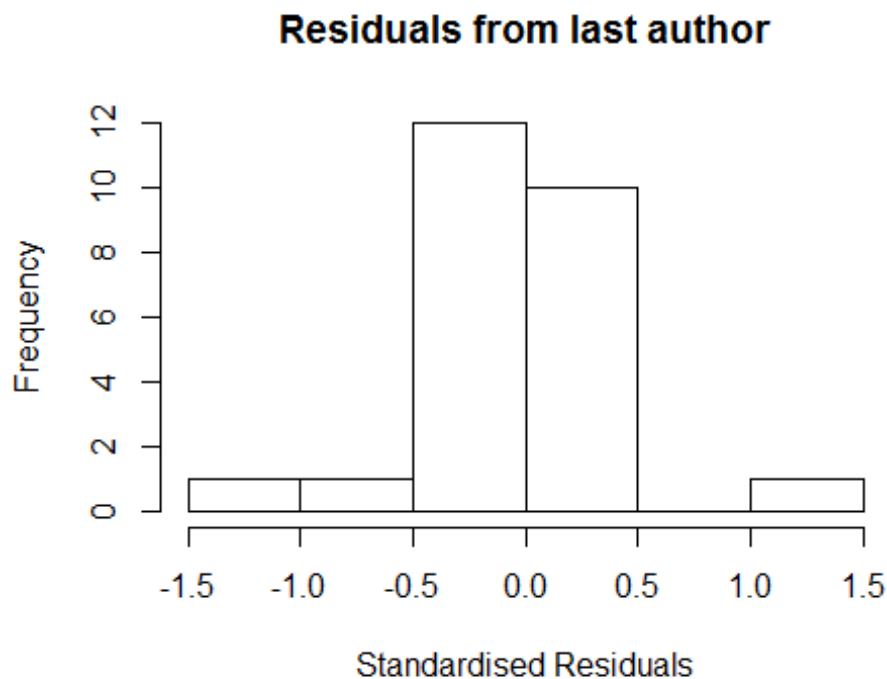
gendered] [check that these decrease]"
##
## 2008 2011 2012
##    1    5    2
##
## 2008 2011 2012
##    1    5    1
##
## 2008 2011 2012
##    1    5    1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: NaN"
## [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: 7"
## [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
##    3    5    2    6    6   11   12    9   14   16   10   17   16   25   28
## 2011 2012
##   39   39
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    1    2    2    8    7    2    9   12    7    9   13   12   19
## 2011 2012
##   30   22
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    1    2    2    8    6    2    7   11    7    8   13   11   15
## 2011 2012
##   29   21
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.30283813966941"
## [1] "Male first author team size 2018 geometric mean: 3.17041368686657"

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

```

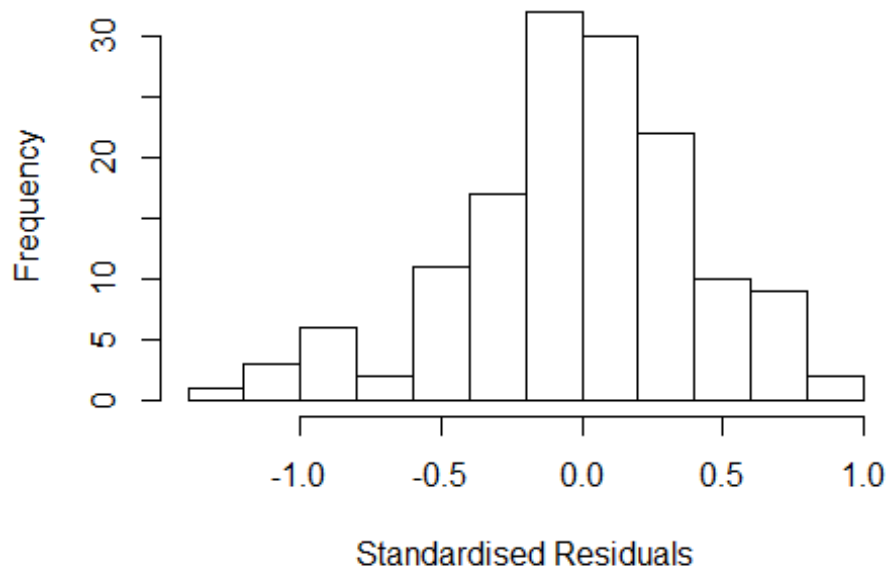
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 220, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.93369834225076"
## [1] "Male last author team size 2018 geometric mean: 3.43708657717475"

## Warning in wilcox.test.default(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 = 0.4
## 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 2.470e+12  1      1.572e+06
## LastAuthorFemale  2.409e+12  1      1.552e+06
## UniqueAuthors    3.716e+13  4      4.969e+01
## Year              1.842e+14 16      2.791e+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
## -1.29e+00 -2.40e-01 1.04e-05 2.42e-01 9.15e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.43e+00 1.13e-07 1.27e+07 < 2e-16 ***
## FirstAuthorFemale1 -2.37e-02 1.06e-01 -2.20e-01 0.82332
## LastAuthorFemale1 -1.23e-01 1.01e-01 -1.22e+00 0.22479
## UniqueAuthors2 4.35e-01 1.13e-01 3.86e+00 0.00018 ***
## UniqueAuthors3 4.04e-01 1.17e-01 3.45e+00 0.00076 ***
## UniqueAuthors4 5.51e-01 2.38e-01 2.31e+00 0.02229 *
## UniqueAuthors5 5.17e-01 1.62e-01 3.20e+00 0.00176 **
## Year1997 -9.23e-01 1.58e-01 -5.85e+00 4.2e-08 ***
## Year1998 -7.54e-01 1.58e-01 -4.78e+00 5.0e-06 ***
## Year1999 -1.06e+00 2.06e-01 -5.17e+00 9.2e-07 ***
```



```

## Year2000      -2.08e-01   1.08e-01 -1.94e+00   0.05526 .
## Year2001      -6.64e-01   2.13e-01 -3.12e+00   0.00224 **
## Year2002      -9.00e-01   2.07e-01 -4.34e+00   2.9e-05 ***
## Year2003      -1.41e+00   1.69e-01 -8.33e+00   1.4e-13 ***
## Year2004      -3.89e-01   1.94e-01 -2.00e+00   0.04754 *
## Year2005      -6.74e-01   1.18e-01 -5.70e+00   8.6e-08 ***
## Year2006      -3.48e-01   1.36e-01 -2.56e+00   0.01170 *
## Year2007      -5.30e-01   1.44e-01 -3.68e+00   0.00035 ***
## Year2008      -6.06e-01   1.92e-01 -3.15e+00   0.00203 **
## Year2009      -5.85e-01   2.51e-01 -2.34e+00   0.02115 *
## Year2010      -4.32e-01   1.51e-01 -2.86e+00   0.00503 **
## Year2011      -6.29e-01   1.47e-01 -4.29e+00   3.6e-05 ***
## Year2012      -8.12e-01   1.64e-01 -4.94e+00   2.5e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.398
## Multiple R-squared:  0.285, Adjusted R-squared:  0.156
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 128 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.275  0.845  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          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 2.484e+12  1      1.576e+06
## LastAuthorFemale  5.570e+12  1      2.360e+06
## Year              6.705e+12 16      2.517e+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)

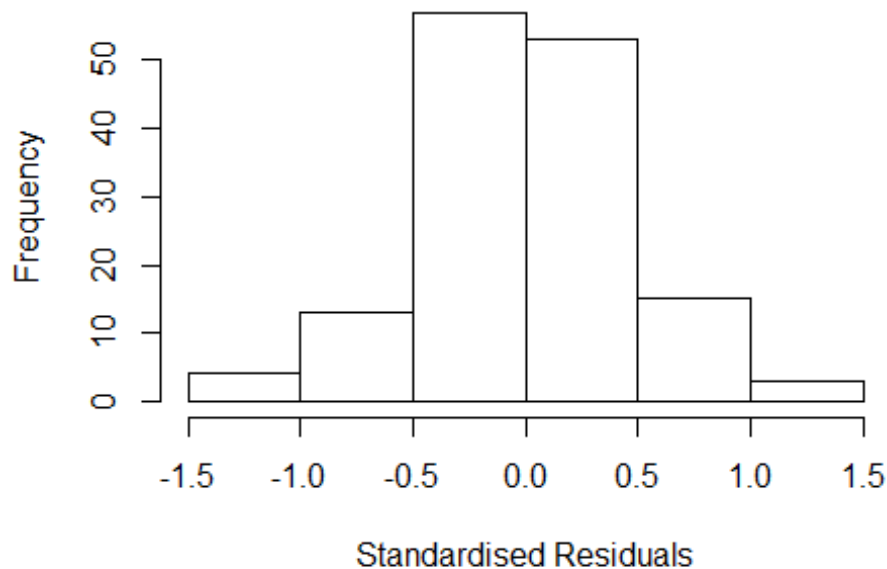
```

```
##
## 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
## -1.35e+00 -2.33e-01 -7.77e-16  2.71e-01  1.17e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.43e+00   7.77e-08  1.84e+07 < 2e-16 ***
## FirstAuthorFemale1 -3.69e-02   1.06e-01 -3.50e-01  0.72849
## LastAuthorFemale1 -1.61e-01   1.00e-01 -1.60e+00  0.11231
## Year1997        -4.38e-01   9.78e-02 -4.48e+00  1.7e-05 ***
## Year1998        -2.69e-01   9.78e-02 -2.75e+00  0.00690 **
## Year1999        -6.12e-01   1.71e-01 -3.59e+00  0.00048 ***
## Year2000         1.59e-02   2.67e-01  6.00e-02  0.95251
## Year2001        -2.86e-01   1.50e-01 -1.92e+00  0.05773 .
## Year2002        -4.94e-01   2.13e-01 -2.32e+00  0.02178 *
## Year2003        -9.55e-01   1.40e-01 -6.81e+00  3.6e-10 ***
## Year2004         1.60e-02   1.47e-01  1.10e-01  0.91353
## Year2005        -4.03e-01   1.40e-01 -2.87e+00  0.00482 **
## Year2006        -2.95e-02   1.20e-01 -2.50e-01  0.80601
## Year2007        -1.13e-01   1.00e-01 -1.13e+00  0.26042
## Year2008        -2.27e-01   2.22e-01 -1.02e+00  0.30879
## Year2009        -8.24e-02   1.82e-01 -4.50e-01  0.65200
## Year2010         1.71e-02   1.05e-01  1.60e-01  0.87073
## Year2011        -2.77e-01   1.03e-01 -2.70e+00  0.00793 **
## Year2012        -3.53e-01   1.26e-01 -2.79e+00  0.00602 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.438
## Multiple R-squared:  0.173, Adjusted R-squared:  0.0544
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~ = 1. The remaining 127 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.322  0.859  0.956  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      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"

## 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 -1.845e+15  1      NaN
## Year              -1.845e+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:
```

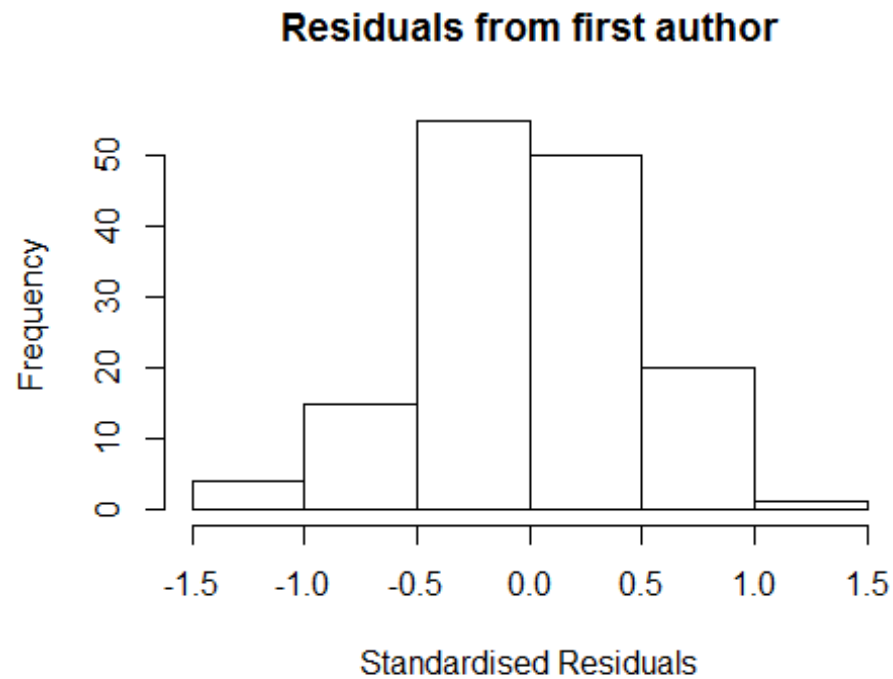
```

##           Min           1Q           Median           3Q           Max
## -1.29e+00 -2.43e-01  2.94e-15  2.65e-01  1.26e+00
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.4330     0.0000      Inf < 2e-16 ***
## FirstAuthorFemale1 -0.1147     0.0913     -1.26  0.21122
## Year1997          -0.5203     0.0913     -5.70  8.0e-08 ***
## Year1998          -0.3513     0.0913     -3.85  0.00019 ***
## Year1999          -0.6143     0.1244     -4.94  2.4e-06 ***
## Year2000           0.0549     0.3021      0.18  0.85620
## Year2001          -0.2640     0.1491     -1.77  0.07888 .
## Year2002          -0.4553     0.2442     -1.86  0.06457 .
## Year2003          -0.9573     0.1040     -9.21  8.7e-16 ***
## Year2004          -0.0167     0.1541     -0.11  0.91386
## Year2005          -0.4155     0.1418     -2.93  0.00402 **
## Year2006          -0.0761     0.1259     -0.60  0.54661
## Year2007          -0.1507     0.0935     -1.61  0.10969
## Year2008          -0.2396     0.2233     -1.07  0.28514
## Year2009          -0.1391     0.1720     -0.81  0.42019
## Year2010          -0.0411     0.1096     -0.37  0.70841
## Year2011          -0.3170     0.1000     -3.17  0.00191 **
## Year2012          -0.3888     0.1264     -3.08  0.00258 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.452
## Multiple R-squared:  0.15, Adjusted R-squared:  0.0365
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~ = 1. The remaining 128 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.393  0.872  0.961   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           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"

```

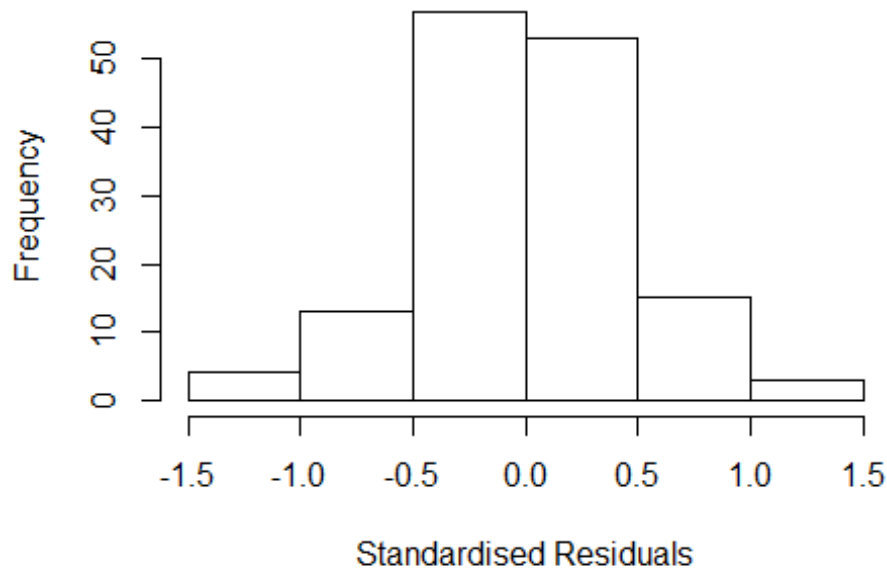
```
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"

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



| | | | | |
|----|------------------|----------|----|---------------------|
| ## | | GVIF | Df | $GVIF^{(1/(2*Df))}$ |
| ## | LastAuthorFemale | 3.26e+13 | 1 | 5.709e+06 |
| ## | Year | 3.26e+13 | 16 | 2.644e+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
## -1.34e+00 -2.44e-01 -1.05e-15 2.77e-01 1.14e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.43300 0.00000 Inf < 2e-16 ***
## LastAuthorFemale1 -0.17920 0.08406 -2.13 0.03494 *
## Year1997 -0.45580 0.08406 -5.42 2.9e-07 ***
## Year1998 -0.28680 0.08406 -3.41 0.00087 ***
## Year1999 -0.63940 0.15894 -4.02 9.8e-05 ***
## Year2000 -0.00250 0.24475 -0.01 0.99187
## Year2001 -0.30526 0.13135 -2.32 0.02171 *
## Year2002 -0.52545 0.18214 -2.88 0.00460 **
## Year2003 -0.98240 0.12313 -7.98 7.5e-13 ***
## Year2004 0.00583 0.14764 0.04 0.96858
## Year2005 -0.40991 0.14357 -2.86 0.00503 **
## Year2006 -0.03176 0.11989 -0.26 0.79149
```

```

## Year2007          -0.12245      0.10112    -1.21   0.22818
## Year2008          -0.24372      0.20705    -1.18   0.24136
## Year2009          -0.08843      0.17693    -0.50   0.61807
## Year2010           0.01145      0.10429     0.11   0.91272
## Year2011          -0.28545      0.10267    -2.78   0.00626 **
## Year2012          -0.35681      0.12692    -2.81   0.00572 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.433
## Multiple R-squared:  0.173, Adjusted R-squared:  0.0625
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 129 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.315  0.862  0.958  0.896  0.987  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      6.90e-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 3613"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 2001 2002 2003 2004 2008 2009 2010 2011 2012
##    2    2    2    1    1    1    2    1    4    2
##
## 1996 2001 2002 2003 2004 2008 2009 2010 2011 2012
##    0    0    0    1    0    1    2    1    2    2
##
## 1996 2001 2002 2003 2004 2008 2009 2010 2011 2012
##    0    0    0    0    0    1    2    1    1    2

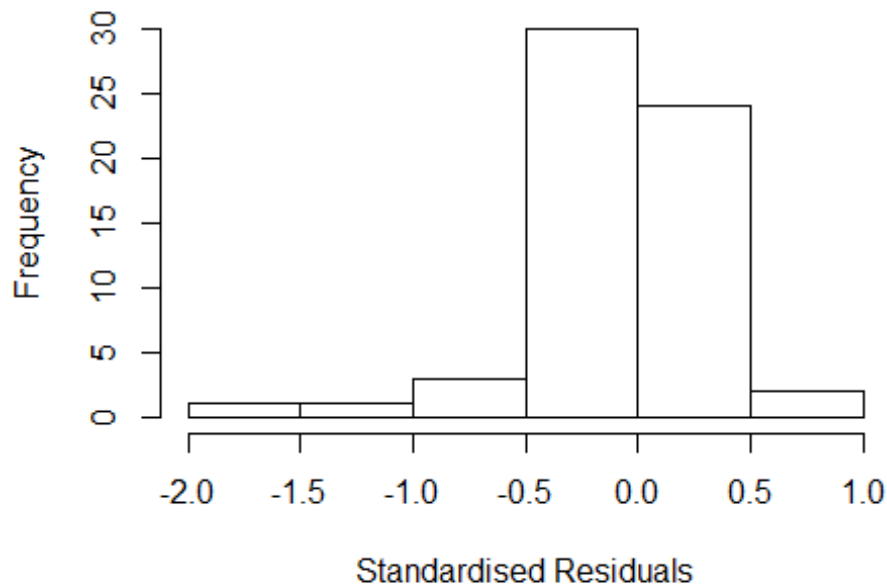
```

```

## [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: 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 3614"
## [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
##      6      2      5      4      9      6      6     11      9     16     10     22     12      7     16
## 2012
##      9
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##      1      0      1      1      1      1      3      6      5      6      5      9     10      4     10
## 2012
##      6
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##      1      0      0      1      1      1      2      5      4      5      4      7     10      4     10
## 2012
##      6
## [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"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 4.392e+13  1      6627404.7
## LastAuthorFemale  3.245e+13  1      5696918.4
## UniqueAuthors    1.385e+28  4        3293.7
## Year              9.138e+41 13         41.1

```


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
## -1.93103 -0.24588 -0.00304 0.22192 0.59825
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.807 0.179 10.08 1.2e-12 ***
## FirstAuthorFemale1 -0.264 0.166 -1.59 0.12042
## LastAuthorFemale1 0.281 0.144 1.96 0.05697 .
## UniqueAuthors2 0.154 0.238 0.65 0.52236
## UniqueAuthors3 0.524 0.178 2.95 0.00525 **
## UniqueAuthors4 0.273 0.152 1.79 0.08013 .
## UniqueAuthors5 0.654 0.211 3.10 0.00350 **
## Year2000 -0.908 0.281 -3.23 0.00244 **
## Year2001 -0.336 0.152 -2.21 0.03290 *
## Year2002 0.183 0.177 1.04 0.30582
```

```

## Year2003          -1.622      0.481   -3.37  0.00165 **
## Year2004          -1.326      0.182   -7.28  6.7e-09 ***
## Year2005          -1.034      0.294   -3.51  0.00110 **
## Year2006          -0.925      0.207   -4.46  6.2e-05 ***
## Year2007          -1.004      0.319   -3.14  0.00311 **
## Year2008          -0.681      0.210   -3.25  0.00233 **
## Year2009          -1.547      0.153  -10.10  1.1e-12 ***
## Year2010          -1.446      0.188   -7.68  1.8e-09 ***
## Year2011          -0.890      0.179   -4.98  1.2e-05 ***
## Year2012          -0.942      0.235   -4.01  0.00025 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.654, Adjusted R-squared:  0.493
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## observation 25 is an outlier with |weight| = 0 ( < 0.0016);
## 9 weights are ~= 1. The remaining 51 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.445  0.903  0.961  0.926  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          1.64e-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 2: First author gender, Last author gender, Year as
factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -1.370e+15  1          NaN
## LastAuthorFemale -1.502e+15  1          NaN
## Year              -4.109e+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
## -1.35e+00 -2.97e-01  8.05e-16  2.61e-01  1.25e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    2.05e+00   2.04e-01   1.01e+01  4.4e-13 ***
## FirstAuthorFemale1 -7.76e-02   1.52e-01  -5.10e-01   0.6116
## LastAuthorFemale1  1.28e-01   1.93e-01   6.60e-01   0.5113
## Year2000        -4.95e-01   2.04e-01  -2.43e+00   0.0192 *
## Year2001        -6.09e-01   6.88e-09  -8.85e+07  < 2e-16 ***
## Year2002         2.16e-01   2.04e-01   1.06e+00   0.2939
## Year2003        -1.32e+00   1.32e+00  -1.00e+00   0.3217
## Year2004        -1.07e+00   1.38e-01  -7.77e+00   7.5e-10 ***
## Year2005        -8.21e-01   2.38e-01  -3.44e+00   0.0013 **
## Year2006        -7.18e-01   2.42e-01  -2.96e+00   0.0048 **
## Year2007        -9.16e-01   4.69e-01  -1.95e+00   0.0573 .
## Year2008        -8.29e-01   3.13e-01  -2.65e+00   0.0111 *
## Year2009        -1.78e+00   1.35e-01  -1.32e+01  < 2e-16 ***
## Year2010        -1.27e+00   2.51e-01  -5.06e+00   7.4e-06 ***
## Year2011        -6.84e-01   1.74e-01  -3.92e+00   0.0003 ***
## Year2012        -8.57e-01   2.47e-01  -3.47e+00   0.0011 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.426
## Multiple R-squared:  0.54,   Adjusted R-squared:  0.387
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~ = 1. The remaining 55 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.298  0.885   0.955   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      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"

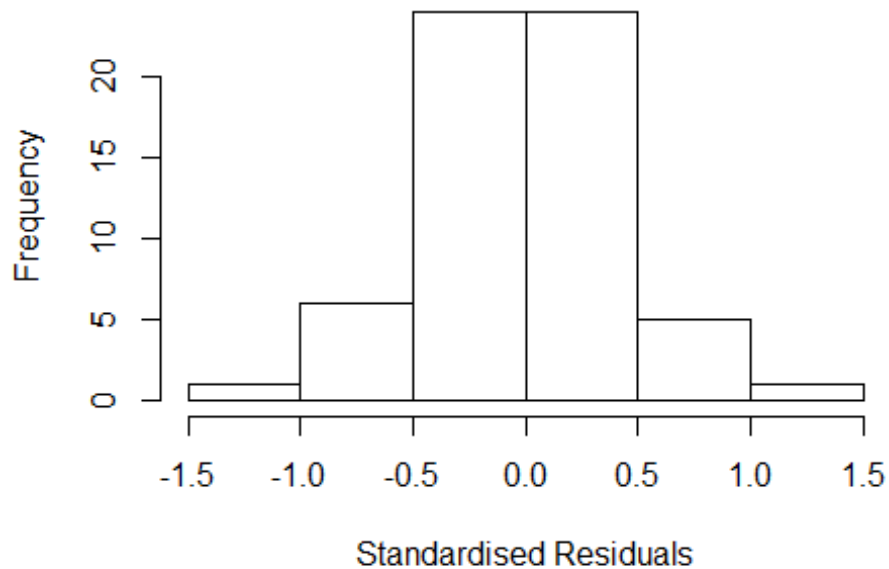
```

```
## compute.outlier.stats
## "SM"
## seed : int(0)
## [1] "Regression 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

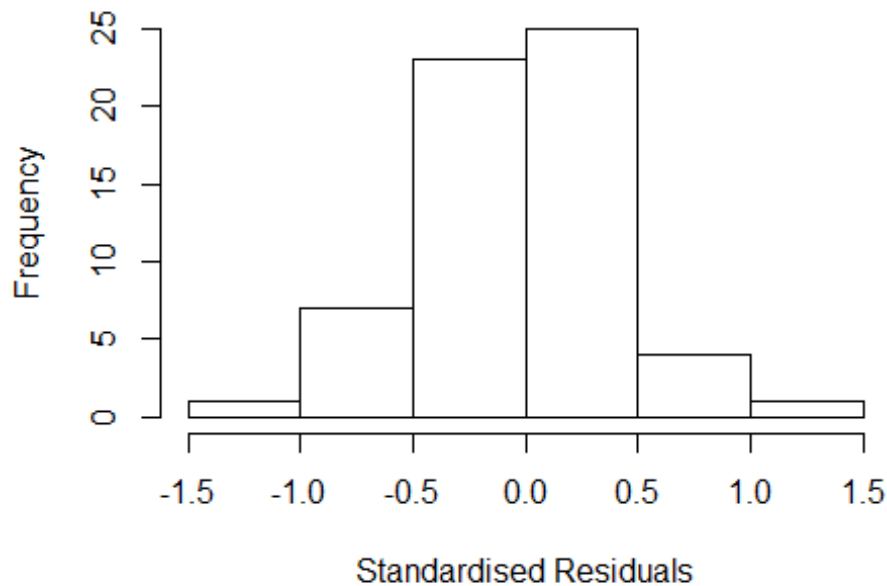
## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

Residuals from first and last author



```
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1          NaN
## Year            NaN 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
## -1.22e+00 -2.75e-01  5.83e-16  2.62e-01  1.20e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.1339     0.1481   14.41 < 2e-16 ***
## FirstAuthorFemale1 -0.0369     0.1481   -0.25  0.80455
## Year2000         -0.5819     0.1481   -3.93  0.00028 ***
## Year2001         -0.6090     0.0000   -Inf < 2e-16 ***
## Year2002          0.1291     0.1481    0.87  0.38770
## Year2003         -1.3449     2.2029   -0.61  0.54454
## Year2004         -1.0982     0.1460   -7.52  1.5e-09 ***
## Year2005         -0.8744     0.2357   -3.71  0.00056 ***
## Year2006         -0.8047     0.1947   -4.13  0.00015 ***
## Year2007         -1.0119     0.4167   -2.43  0.01914 *
## Year2008         -0.9183     0.2414   -3.80  0.00042 ***
## Year2009         -1.8215     0.1309  -13.91 < 2e-16 ***
```

```

## Year2010          -1.3811      0.1848   -7.47  1.8e-09 ***
## Year2011          -0.7288      0.1462   -4.99  9.2e-06 ***
## Year2012          -0.9655      0.1574   -6.13  1.8e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.425
## Multiple R-squared:  0.529, Adjusted R-squared:  0.386
## Convergence in 42 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 55 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.393  0.881  0.958  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      1.64e-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"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.309e+14 1      1.144e+07
## Year              1.309e+14 13      3.491e+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
## -1.30e+00 -2.54e-01  8.88e-16  2.73e-01  1.28e+00
##
## Coefficients:
##      Estimate Std. Error  t value Pr(>|t|)

```

```

## (Intercept)      2.00e+00   1.89e-01   1.06e+01   6.6e-14 ***
## LastAuthorFemale1 9.64e-02   1.89e-01   5.10e-01   0.61272
## Year2000         -4.49e-01   1.89e-01  -2.37e+00   0.02194 *
## Year2001         -6.09e-01   1.19e-08  -5.12e+07   < 2e-16 ***
## Year2002          2.62e-01   1.89e-01   1.39e+00   0.17202
## Year2003         -1.26e+00   1.46e+00  -8.70e-01   0.39121
## Year2004         -1.04e+00   1.40e-01  -7.39e+00   2.4e-09 ***
## Year2005         -7.66e-01   2.21e-01  -3.46e+00   0.00116 **
## Year2006         -6.71e-01   2.31e-01  -2.91e+00   0.00553 **
## Year2007         -8.87e-01   4.42e-01  -2.01e+00   0.05059 .
## Year2008         -7.96e-01   3.16e-01  -2.52e+00   0.01519 *
## Year2009         -1.77e+00   1.38e-01  -1.29e+01   < 2e-16 ***
## Year2010         -1.27e+00   2.59e-01  -4.90e+00   1.2e-05 ***
## Year2011         -6.72e-01   1.81e-01  -3.71e+00   0.00056 ***
## Year2012         -8.52e-01   2.43e-01  -3.50e+00   0.00103 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.426
## Multiple R-squared:  0.537, Adjusted R-squared:  0.397
## Convergence in 29 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 54 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.330  0.884  0.956  0.907  0.974  0.999
## Algorithmic parameters:
##           tuning.chi           bb           tuning.psi           refine.tol
##           1.55e+00           5.00e-01           4.69e+00           1.00e-07
##           rel.tol           solve.tol           eps.outlier           eps.x
##           1.00e-07           1.00e-07           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] "Sample size for the above analysis: 61"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 3616"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"

```

```

##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    3    1    3    1    2    1    1    2    3    3    3    4    8    5
## 2011 2012
##    4    3
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    0    1    2    1    1    1    1    2    3    2    3    4    8    4
## 2011 2012
##    4    3
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    0    1    2    1    1    0    1    2    3    2    3    4    7    4
## 2011 2012
##    4    3
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.94168275343288"
## [1] "Male first author team size 2018 geometric mean: 5.56631536742748"

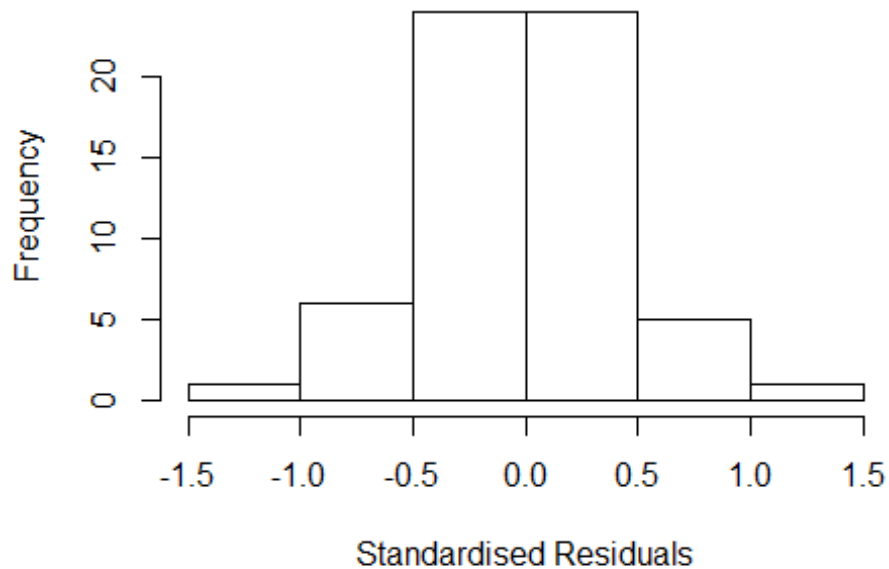
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3, p-value = 0.06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.77454961710008"
## [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

```


Residuals from last author



```
##
## 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"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 4.185e+14 1      20457609
## LastAuthorFemale 3.978e+14 1      19945425
## UniqueAuthors    -1.629e+21 4         NaN
## Year              -8.171e+35 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 +
UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
```

```

##           Min           1Q           Median           3Q           Max
## -9.63e-01 -2.41e-01 -5.27e-16  3.20e-01  7.75e-01
##
## Coefficients:
##               Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   -2.19e-02   2.80e-01  -8.00e-02  0.93868
## FirstAuthorFemale1  9.71e-02   5.02e-01   1.90e-01  0.84895
## LastAuthorFemale1  4.78e-01   3.33e-01   1.44e+00  0.16793
## UniqueAuthors2   -5.42e-01   2.75e-01  -1.97e+00  0.06468 .
## UniqueAuthors3   -5.70e-01   5.68e-01  -1.00e+00  0.32873
## UniqueAuthors4    2.72e-01   4.57e-01   6.00e-01  0.55898
## UniqueAuthors5   -1.02e+00   2.31e-01  -4.41e+00  0.00034 ***
## Year1999         7.11e-01   7.36e-01   9.70e-01  0.34642
## Year2000         4.80e-01   3.82e-08   1.26e+07 < 2e-16 ***
## Year2001         1.84e+00   7.66e-01   2.40e+00  0.02759 *
## Year2003         1.72e+00   2.80e-01   6.12e+00  8.9e-06 ***
## Year2004         1.33e+00   1.17e-01   1.14e+01  1.2e-09 ***
## Year2005         1.14e+00   1.98e-01   5.74e+00  1.9e-05 ***
## Year2006         9.92e-01   3.83e-01   2.59e+00  0.01836 *
## Year2007         9.26e-01   2.93e-01   3.16e+00  0.00544 **
## Year2008         1.16e+00   2.16e-01   5.40e+00  4.0e-05 ***
## Year2009         1.28e+00   3.28e-01   3.89e+00  0.00108 **
## Year2010         5.37e-01   2.89e-01   1.86e+00  0.07997 .
## Year2011         9.52e-01   4.40e-01   2.16e+00  0.04412 *
## Year2012         1.32e+00   5.88e-01   2.25e+00  0.03736 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.644
## Multiple R-squared:  0.438, Adjusted R-squared:  -0.156
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 29 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.807  0.935  0.974  0.951  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           2.63e-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

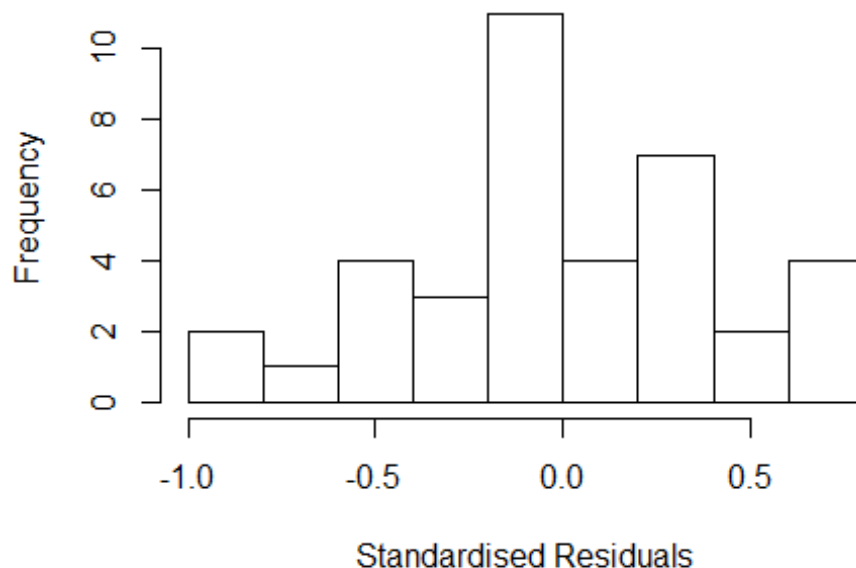
```

```
##                                "SM"
## seed : int(0)
## [1] "Regression 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 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
```

```

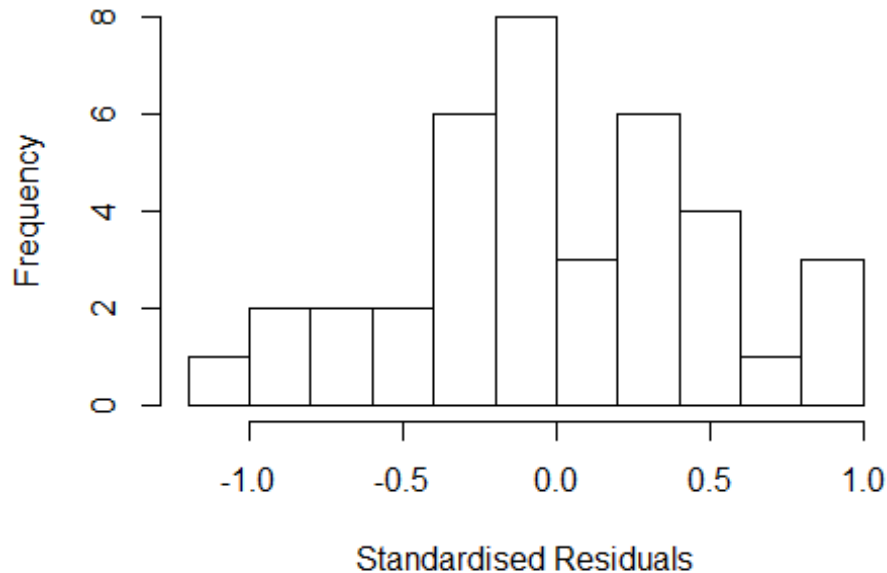
## -1.02e+00 -3.47e-01 -3.05e-16 3.47e-01 8.99e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.307      0.345    0.89 0.38388
## FirstAuthorFemale1 -0.149      0.413   -0.36 0.72142
## LastAuthorFemale1  0.395      0.271    1.46 0.15931
## Year1999          0.519      0.829    0.63 0.53779
## Year2000          0.480      0.000    Inf < 2e-16 ***
## Year2001          1.184      0.271    4.36 0.00025 ***
## Year2003          1.387      0.345    4.02 0.00058 ***
## Year2004          1.018      0.264    3.85 0.00087 ***
## Year2005          0.469      0.515    0.91 0.37204
## Year2006          0.663      0.430    1.54 0.13708
## Year2007          0.357      0.215    1.66 0.11169
## Year2008          0.556      0.183    3.05 0.00590 **
## Year2009          0.823      0.292    2.82 0.01000 *
## Year2010          0.353      0.282    1.25 0.22409
## Year2011          0.232      0.301    0.77 0.44954
## Year2012          0.989      0.270    3.67 0.00136 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.767
## Multiple R-squared: 0.235, Adjusted R-squared: -0.287
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 29 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.845 0.909 0.973 0.952 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      2.63e-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"

```

```
## 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 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
## -1.00e+00 -3.42e-01 -1.11e-16  3.41e-01  1.06e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.3543     0.3306   1.07  0.29494
## FirstAuthorFemale1 0.1987     0.3306   0.60  0.55359
```

```

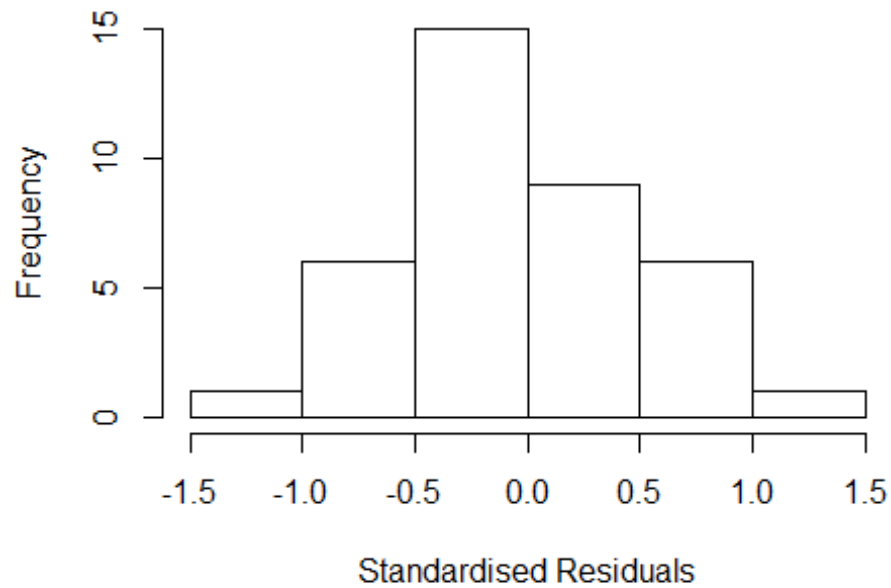
## Year1999      0.4712      0.8340      0.56  0.57755
## Year2000      0.4800      0.0000      Inf < 2e-16 ***
## Year2001      0.7890      0.0000      Inf < 2e-16 ***
## Year2003      1.3397      0.3306      4.05 0.00049 ***
## Year2004      0.8200      0.3944      2.08 0.04893 *
## Year2005      0.2843      0.5180      0.55 0.58846
## Year2006      0.6157      0.4184      1.47 0.15463
## Year2007      0.2410      0.3486      0.69 0.49626
## Year2008      0.5439      0.1811      3.00 0.00634 **
## Year2009      0.8027      0.2872      2.79 0.01030 *
## Year2010      0.3543      0.2820      1.26 0.22156
## Year2011      0.1280      0.2464      0.52 0.60831
## Year2012      0.7241      0.0734      9.87 9.8e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.743
## Multiple R-squared:  0.2, Adjusted R-squared:  -0.287
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 30 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.822  0.927  0.966  0.948  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          2.63e-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"

## 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

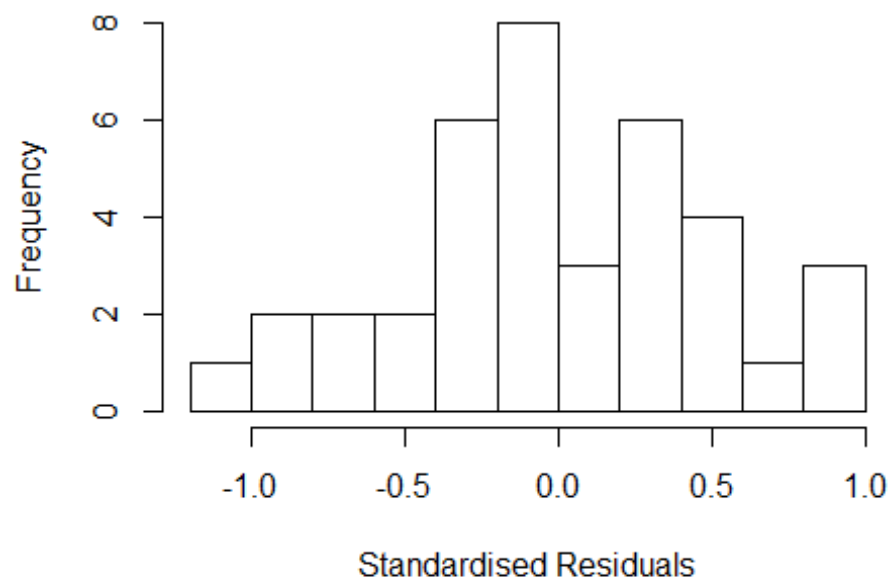
```

Residuals from first author



```
##          GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale  NaN  1          NaN
## Year              NaN 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
## -1.06e+00 -3.48e-01 1.05e-15 3.63e-01 9.44e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.228 0.231 0.99 0.33429
## LastAuthorFemale1 0.325 0.231 1.41 0.17307
## Year1999 0.598 0.793 0.75 0.45887
## Year2000 0.480 0.000 Inf < 2e-16 ***
## Year2001 1.114 0.231 4.82 7.3e-05 ***
## Year2003 1.466 0.231 6.34 1.8e-06 ***
## Year2004 0.983 0.279 3.52 0.00184 **
## Year2005 0.465 0.499 0.93 0.36143
## Year2006 0.742 0.345 2.15 0.04220 *
## Year2007 0.335 0.239 1.40 0.17407
## Year2008 0.577 0.165 3.50 0.00192 **
## Year2009 0.857 0.276 3.11 0.00497 **
## Year2010 0.353 0.282 1.25 0.22249
## Year2011 0.214 0.292 0.73 0.46994
## Year2012 0.942 0.227 4.15 0.00038 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.757
## Multiple R-squared: 0.233, Adjusted R-squared: -0.234
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~ = 1. The remaining 30 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.831 0.930 0.976 0.952 0.981 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 2.63e-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"
```

```
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 | 220 | | 61.82 | | 50.000 | | 47.826 | | -2.1739 | 50.000 | | 43.478 | | -6.522 |
| ## 1100 | 701 | | 61.06 | | 0.000 | | 51.852 | | 51.8519 | 0.000 | | 18.519 | | 18.519 |
| ## 1101 | 151 | | 52.32 | | 50.000 | | 0.000 | | -50.0000 | 50.000 | | 33.333 | | -16.667 |
| ## 1102 | 788 | | 30.71 | | 0.000 | | 33.333 | | 33.3333 | 33.333 | | 25.000 | | -8.333 |
| ## 1103 | 1490 | | 21.28 | | 42.857 | | 44.000 | | 1.1429 | 42.857 | | 36.000 | | -6.857 |
| ## 1104 | 799 | | 51.06 | | 41.667 | | 36.842 | | -4.8246 | 25.000 | | 52.632 | | 27.632 |
| ## 1105 | 795 | | 62.26 | | 25.000 | | 36.842 | | 11.8421 | 33.333 | | 36.842 | | 3.509 |
| ## 1106 | 2898 | | 38.44 | | 11.111 | | 54.545 | | 43.4343 | 11.111 | | 36.364 | | 25.253 |
| ## 1107 | 244 | | 59.84 | | 0.000 | | 37.500 | | 37.5000 | 100.000 | | 25.000 | | -75.000 |
| ## 1108 | 88 | | 29.55 | | 0.000 | | 0.000 | | 0.0000 | 100.000 | | 0.000 | | -100.000 |
| ## 1109 | 114 | | 60.53 | | NaN | | 100.000 | | NaN | NaN | | 0.000 | | NaN |
| ## 1110 | 358 | | 54.47 | | 33.333 | | 12.500 | | -20.8333 | 33.333 | | 18.750 | | -14.583 |
| ## 1111 | 239 | | 39.75 | | 50.000 | | 50.000 | | 0.0000 | 50.000 | | 0.000 | | -50.000 |
| ## 1200 | 130 | | 69.23 | | NaN | | 37.500 | | NaN | NaN | | 37.500 | | NaN |
| ## 1201 | 488 | | 76.64 | | NaN | | 64.286 | | NaN | NaN | | 64.286 | | NaN |
| ## 1202 | 900 | | 81.44 | | 0.000 | | 39.683 | | 39.6825 | 0.000 | | 38.095 | | 38.095 |
| ## 1203 | 592 | | 78.72 | | 25.000 | | 70.370 | | 45.3704 | 25.000 | | 70.370 | | 45.370 |
| ## 1204 | 205 | | 83.41 | | 0.000 | | 42.857 | | 42.8571 | 0.000 | | 35.714 | | 35.714 |
| ## 1205 | 85 | | 81.18 | | 0.000 | | 14.286 | | 14.2857 | 0.000 | | 14.286 | | 14.286 |
| ## 1206 | 16 | | 87.50 | | NaN | | 100.000 | | NaN | NaN | | 100.000 | | NaN |
| ## 1207 | 535 | | 62.06 | | 57.143 | | 30.000 | | -27.1429 | 14.286 | | 25.000 | | 10.714 |
| ## 1208 | 647 | | 83.46 | | 0.000 | | 45.000 | | 45.0000 | 0.000 | | 45.000 | | 45.000 |

| | | | | | | | | |
|---------|------|-------|---------|---------|-----------|---------|---------|----------|
| ## 1209 | 12 | 83.33 | NaN | 100.000 | NaN | NaN | 100.000 | NaN |
| ## 1210 | 94 | 76.60 | 100.000 | 42.857 | -57.1429 | 0.000 | 42.857 | 42.857 |
| ## 1211 | 551 | 87.30 | 11.111 | 23.810 | 12.6984 | 22.222 | 23.810 | 1.587 |
| ## 1212 | 230 | 86.52 | NaN | 25.000 | NaN | NaN | 30.000 | NaN |
| ## 1213 | 275 | 80.36 | NaN | 53.333 | NaN | NaN | 53.333 | NaN |
| ## 1300 | 673 | 63.15 | 100.000 | 46.341 | -53.6585 | 100.000 | 21.951 | -78.049 |
| ## 1301 | 13 | 53.85 | NaN | 100.000 | NaN | NaN | 0.000 | NaN |
| ## 1302 | 138 | 65.94 | 0.000 | 62.500 | 62.5000 | 0.000 | 50.000 | 50.000 |
| ## 1303 | 2295 | 64.62 | 41.667 | 35.135 | -6.5315 | 14.583 | 24.324 | 9.741 |
| ## 1304 | 695 | 52.09 | 42.857 | 57.143 | 14.2857 | 57.143 | 35.714 | -21.429 |
| ## 1305 | 1246 | 47.19 | 20.000 | 62.500 | 42.5000 | 0.000 | 37.500 | 37.500 |
| ## 1306 | 544 | 50.74 | 0.000 | 70.588 | 70.5882 | 0.000 | 41.176 | 41.176 |
| ## 1307 | 865 | 57.46 | 63.636 | 46.667 | -16.9697 | 18.182 | 20.000 | 1.818 |
| ## 1308 | 439 | 59.00 | 42.857 | 33.333 | -9.5238 | 0.000 | 11.111 | 11.111 |
| ## 1309 | 249 | 51.41 | NaN | 12.500 | NaN | NaN | 37.500 | NaN |
| ## 1310 | 402 | 39.30 | NaN | 60.000 | NaN | NaN | 20.000 | NaN |
| ## 1311 | 1129 | 47.21 | 55.556 | 46.154 | -9.4017 | 11.111 | 35.897 | 24.786 |
| ## 1312 | 1557 | 60.82 | 42.857 | 54.545 | 11.6883 | 21.429 | 36.364 | 14.935 |
| ## 1313 | 419 | 64.68 | 0.000 | 41.176 | 41.1765 | 33.333 | 35.294 | 1.961 |
| ## 1314 | 508 | 52.56 | 100.000 | 12.000 | -88.0000 | 0.000 | 20.000 | 20.000 |
| ## 1315 | 118 | 70.34 | 50.000 | 0.000 | -50.0000 | 100.000 | 0.000 | -100.000 |
| ## 1400 | 324 | 82.72 | 0.000 | 22.222 | 22.2222 | 0.000 | 22.222 | 22.222 |
| ## 1401 | 164 | 79.27 | NaN | 33.333 | NaN | NaN | 44.444 | NaN |
| ## 1402 | 135 | 74.81 | 66.667 | 16.667 | -50.0000 | 66.667 | 16.667 | -50.000 |
| ## 1403 | 342 | 77.49 | NaN | 20.000 | NaN | NaN | 16.000 | NaN |
| ## 1404 | 132 | 68.94 | 0.000 | 50.000 | 50.0000 | 0.000 | 0.000 | 0.000 |
| ## 1405 | 272 | 80.15 | 50.000 | 36.842 | -13.1579 | 50.000 | 42.105 | -7.895 |
| ## 1406 | 197 | 77.66 | NaN | 50.000 | NaN | NaN | 37.500 | NaN |
| ## 1407 | 324 | 82.10 | 66.667 | 33.333 | -33.3333 | 66.667 | 55.556 | -11.111 |
| ## 1408 | 575 | 76.35 | 14.286 | 26.471 | 12.1849 | 14.286 | 47.059 | 32.773 |
| ## 1409 | 122 | 72.95 | NaN | 25.000 | NaN | NaN | 25.000 | NaN |
| ## 1410 | 64 | 81.25 | 100.000 | 0.000 | -100.0000 | 100.000 | 0.000 | -100.000 |
| ## 1500 | 676 | 50.44 | 14.286 | 45.455 | 31.1688 | 14.286 | 36.364 | 22.078 |
| ## 1501 | 22 | 59.09 | NaN | 50.000 | NaN | NaN | 50.000 | NaN |
| ## 1502 | 725 | 49.38 | 50.000 | 30.769 | -19.2308 | 0.000 | 7.692 | 7.692 |
| ## 1503 | 440 | 63.64 | 14.286 | 27.273 | 12.9870 | 14.286 | 27.273 | 12.987 |
| ## 1504 | 8 | 62.50 | NaN | NaN | NaN | NaN | NaN | NaN |
| ## 1505 | 139 | 49.64 | NaN | 33.333 | NaN | NaN | 33.333 | NaN |
| ## 1506 | 37 | 67.57 | NaN | NaN | NaN | NaN | NaN | NaN |
| ## 1507 | 120 | 46.67 | NaN | 0.000 | NaN | NaN | 0.000 | NaN |
| ## 1508 | 175 | 54.86 | 25.000 | 33.333 | 8.3333 | 0.000 | 0.000 | 0.000 |
| ## 1600 | 1849 | 56.25 | 16.667 | 29.630 | 12.9630 | 33.333 | 22.222 | -11.111 |
| ## 1601 | 69 | 53.62 | NaN | 0.000 | NaN | NaN | 50.000 | NaN |
| ## 1602 | 844 | 63.74 | 18.750 | 41.667 | 22.9167 | 6.250 | 33.333 | 27.083 |
| ## 1603 | 476 | 60.08 | 16.667 | 72.727 | 56.0606 | 0.000 | 18.182 | 18.182 |
| ## 1604 | 376 | 68.09 | 16.667 | 50.000 | 33.3333 | 0.000 | 30.000 | 30.000 |
| ## 1605 | 1201 | 69.03 | 13.333 | 31.579 | 18.2456 | 13.333 | 15.789 | 2.456 |
| ## 1606 | 1163 | 60.45 | 15.385 | 22.727 | 7.3427 | 0.000 | 27.273 | 27.273 |
| ## 1607 | 389 | 63.75 | 50.000 | 44.444 | -5.5556 | 0.000 | 33.333 | 33.333 |
| ## 1700 | 273 | 59.71 | NaN | 44.444 | NaN | NaN | 22.222 | NaN |

| | | | | | | | | |
|---------|------|-------|---------|---------|----------|---------|---------|---------|
| ## 1701 | 43 | 76.74 | NaN | 0.000 | NaN | NaN | 0.000 | NaN |
| ## 1702 | 274 | 67.15 | 0.000 | 0.000 | 0.0000 | 0.000 | 0.000 | 0.000 |
| ## 1703 | 216 | 61.11 | NaN | 0.000 | NaN | NaN | 0.000 | NaN |
| ## 1704 | 94 | 60.64 | NaN | 25.000 | NaN | NaN | 25.000 | NaN |
| ## 1705 | 462 | 61.04 | 50.000 | 33.333 | -16.6667 | 50.000 | 33.333 | -16.667 |
| ## 1706 | 1090 | 56.61 | 0.000 | 23.529 | 23.5294 | 0.000 | 23.529 | 23.529 |
| ## 1707 | 116 | 64.66 | NaN | 40.000 | NaN | NaN | 60.000 | NaN |
| ## 1708 | 269 | 58.36 | 0.000 | 16.667 | 16.6667 | 0.000 | 16.667 | 16.667 |
| ## 1709 | 167 | 66.47 | 100.000 | 25.000 | -75.0000 | 100.000 | 12.500 | -87.500 |
| ## 1710 | 470 | 66.81 | 33.333 | 26.316 | -7.0175 | 33.333 | 26.316 | -7.018 |
| ## 1711 | 163 | 54.60 | 0.000 | 12.500 | 12.5000 | 0.000 | 25.000 | 25.000 |
| ## 1712 | 672 | 62.20 | 0.000 | 18.750 | 18.7500 | 0.000 | 12.500 | 12.500 |
| ## 1800 | 52 | 78.85 | NaN | 100.000 | NaN | NaN | 100.000 | NaN |
| ## 1801 | 0 | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| ## 1802 | 124 | 65.32 | 0.000 | 50.000 | 50.0000 | 0.000 | 100.000 | 100.000 |
| ## 1803 | 165 | 66.67 | 0.000 | 14.286 | 14.2857 | 0.000 | 14.286 | 14.286 |
| ## 1804 | 142 | 72.54 | 0.000 | 0.000 | 0.0000 | 0.000 | 0.000 | 0.000 |
| ## 1900 | 309 | 71.52 | 11.111 | 36.364 | 25.2525 | 22.222 | 45.455 | 23.232 |
| ## 1901 | 143 | 45.45 | NaN | 50.000 | NaN | NaN | 50.000 | NaN |
| ## 1902 | 172 | 62.79 | 0.000 | 33.333 | 33.3333 | 100.000 | 50.000 | -50.000 |
| ## 1903 | 51 | 62.75 | NaN | NaN | NaN | NaN | NaN | NaN |
| ## 1904 | 160 | 68.75 | NaN | 66.667 | NaN | NaN | 44.444 | NaN |
| ## 1905 | 19 | 42.11 | NaN | 0.000 | NaN | NaN | 0.000 | NaN |
| ## 1906 | 142 | 45.07 | NaN | 0.000 | NaN | NaN | 0.000 | NaN |
| ## 1907 | 202 | 50.99 | 0.000 | 0.000 | 0.0000 | 0.000 | 0.000 | 0.000 |
| ## 1908 | 135 | 51.85 | 0.000 | 33.333 | 33.3333 | 0.000 | 0.000 | 0.000 |
| ## 1909 | 142 | 59.86 | 0.000 | 0.000 | 0.0000 | 0.000 | 0.000 | 0.000 |
| ## 1910 | 247 | 54.66 | 0.000 | 50.000 | 50.0000 | 0.000 | 50.000 | 50.000 |
| ## 1911 | 64 | 71.88 | NaN | 100.000 | NaN | NaN | 0.000 | NaN |
| ## 1912 | 176 | 38.07 | 0.000 | 25.000 | 25.0000 | 0.000 | 50.000 | 50.000 |
| ## 1913 | 15 | 66.67 | NaN | 0.000 | NaN | NaN | 0.000 | NaN |
| ## 2000 | 145 | 79.31 | NaN | 20.000 | NaN | NaN | 10.000 | NaN |
| ## 2001 | 119 | 75.63 | NaN | 28.571 | NaN | NaN | 42.857 | NaN |
| ## 2002 | 1225 | 78.69 | 14.286 | 27.778 | 13.4921 | 14.286 | 24.074 | 9.788 |
| ## 2003 | 305 | 83.93 | 0.000 | 11.765 | 11.7647 | 0.000 | 11.765 | 11.765 |
| ## 2100 | 290 | 53.45 | NaN | 29.412 | NaN | NaN | 41.176 | NaN |
| ## 2101 | 28 | 64.29 | NaN | 33.333 | NaN | NaN | 0.000 | NaN |
| ## 2102 | 232 | 53.45 | NaN | 8.333 | NaN | NaN | 8.333 | NaN |
| ## 2103 | 101 | 31.68 | 0.000 | 50.000 | 50.0000 | 0.000 | 50.000 | 50.000 |
| ## 2104 | 49 | 24.49 | 0.000 | 0.000 | 0.0000 | 0.000 | 0.000 | 0.000 |
| ## 2105 | 500 | 48.80 | 50.000 | 17.647 | -32.3529 | 50.000 | 17.647 | -32.353 |
| ## 2200 | 443 | 52.37 | 0.000 | 20.000 | 20.0000 | 0.000 | 10.000 | 10.000 |
| ## 2201 | 93 | 60.22 | NaN | 100.000 | NaN | NaN | 25.000 | NaN |
| ## 2202 | 105 | 55.24 | 0.000 | 16.667 | 16.6667 | 0.000 | 0.000 | 0.000 |
| ## 2203 | 50 | 64.00 | NaN | 60.000 | NaN | NaN | 40.000 | NaN |
| ## 2204 | 847 | 49.47 | 0.000 | 43.750 | 43.7500 | 0.000 | 21.875 | 21.875 |
| ## 2205 | 477 | 44.23 | NaN | 27.778 | NaN | NaN | 22.222 | NaN |
| ## 2206 | 55 | 40.00 | NaN | NaN | NaN | NaN | NaN | NaN |
| ## 2207 | 295 | 49.83 | 50.000 | 0.000 | -50.0000 | 0.000 | 0.000 | 0.000 |
| ## 2208 | 1886 | 46.45 | 0.000 | 5.882 | 5.8824 | 0.000 | 5.882 | 5.882 |

| | | | | | | | | | |
|----|------|------|--------|---------|---------|----------|--------|---------|---------|
| ## | 2209 | 797 | 38.52 | NaN | 37.500 | NaN | NaN | 16.667 | NaN |
| ## | 2210 | 1091 | 36.85 | 0.000 | 12.500 | 12.5000 | 0.000 | 8.333 | 8.333 |
| ## | 2211 | 840 | 33.57 | 0.000 | 11.111 | 11.1111 | 0.000 | 5.556 | 5.556 |
| ## | 2212 | 102 | 50.98 | NaN | 11.111 | NaN | NaN | 22.222 | NaN |
| ## | 2213 | 266 | 54.14 | 0.000 | 66.667 | 66.6667 | 0.000 | 11.111 | 11.111 |
| ## | 2214 | 91 | 65.93 | NaN | 0.000 | NaN | NaN | 0.000 | NaN |
| ## | 2215 | 262 | 45.80 | NaN | 11.111 | NaN | NaN | 11.111 | NaN |
| ## | 2216 | 37 | 75.68 | NaN | 0.000 | NaN | NaN | 0.000 | NaN |
| ## | 2300 | 623 | 52.97 | 20.000 | 40.000 | 20.0000 | 20.000 | 20.000 | 0.000 |
| ## | 2301 | 106 | 76.42 | NaN | 11.111 | NaN | NaN | 11.111 | NaN |
| ## | 2302 | 111 | 40.54 | NaN | 0.000 | NaN | NaN | 0.000 | NaN |
| ## | 2303 | 881 | 40.41 | 0.000 | 61.538 | 61.5385 | 0.000 | 23.077 | 23.077 |
| ## | 2304 | 502 | 52.19 | 50.000 | 28.571 | -21.4286 | 50.000 | 28.571 | -21.429 |
| ## | 2305 | 408 | 37.99 | NaN | 25.000 | NaN | NaN | 0.000 | NaN |
| ## | 2306 | 80 | 67.50 | 0.000 | 57.143 | 57.1429 | 0.000 | 28.571 | 28.571 |
| ## | 2307 | 211 | 45.97 | 50.000 | 73.333 | 23.3333 | 0.000 | 46.667 | 46.667 |
| ## | 2308 | 726 | 57.99 | 50.000 | 36.364 | -13.6364 | 50.000 | 39.394 | -10.606 |
| ## | 2309 | 241 | 65.98 | 0.000 | 25.000 | 25.0000 | 0.000 | 25.000 | 25.000 |
| ## | 2310 | 448 | 35.49 | 0.000 | 66.667 | 66.6667 | 0.000 | 33.333 | 33.333 |
| ## | 2311 | 392 | 40.82 | 0.000 | 40.000 | 40.0000 | 0.000 | 0.000 | 0.000 |
| ## | 2312 | 410 | 43.17 | 0.000 | 10.000 | 10.0000 | 0.000 | 20.000 | 20.000 |
| ## | 2400 | 101 | 58.42 | NaN | 50.000 | NaN | NaN | 25.000 | NaN |
| ## | 2401 | 1 | 100.00 | NaN | NaN | NaN | NaN | NaN | NaN |
| ## | 2402 | 872 | 43.35 | 27.273 | 42.857 | 15.5844 | 9.091 | 21.429 | 12.338 |
| ## | 2403 | 847 | 52.54 | 58.333 | 40.000 | -18.3333 | 16.667 | 32.000 | 15.333 |
| ## | 2404 | 978 | 52.76 | 66.667 | 34.483 | -32.1839 | 0.000 | 41.379 | 41.379 |
| ## | 2405 | 206 | 51.46 | 100.000 | 33.333 | -66.6667 | 0.000 | 41.667 | 41.667 |
| ## | 2406 | 183 | 51.91 | NaN | 66.667 | NaN | NaN | 0.000 | NaN |
| ## | 2500 | 1280 | 43.91 | 0.000 | 34.043 | 34.0426 | 0.000 | 19.149 | 19.149 |
| ## | 2501 | 17 | 70.59 | NaN | NaN | NaN | NaN | NaN | NaN |
| ## | 2502 | 494 | 46.76 | NaN | 47.619 | NaN | NaN | 9.524 | NaN |
| ## | 2503 | 502 | 37.65 | 0.000 | 28.571 | 28.5714 | 50.000 | 0.000 | -50.000 |
| ## | 2504 | 1472 | 41.51 | 33.333 | 16.667 | -16.6667 | 0.000 | 12.500 | 12.500 |
| ## | 2505 | 1088 | 52.21 | 20.000 | 33.333 | 13.3333 | 10.000 | 0.000 | -10.000 |
| ## | 2506 | 548 | 37.96 | 0.000 | 28.571 | 28.5714 | 50.000 | 0.000 | -50.000 |
| ## | 2507 | 276 | 50.00 | 33.333 | 33.333 | 0.0000 | 0.000 | 11.111 | 11.111 |
| ## | 2508 | 932 | 41.52 | 0.000 | 18.750 | 18.7500 | 20.000 | 6.250 | -13.750 |
| ## | 2600 | 400 | 73.25 | 8.333 | 9.091 | 0.7576 | 8.333 | 18.182 | 9.848 |
| ## | 2601 | 64 | 57.81 | NaN | 60.000 | NaN | NaN | 20.000 | NaN |
| ## | 2602 | 243 | 80.66 | 0.000 | 14.286 | 14.2857 | 0.000 | 14.286 | 14.286 |
| ## | 2603 | 199 | 74.37 | 0.000 | 0.000 | 0.0000 | 0.000 | 11.111 | 11.111 |
| ## | 2604 | 737 | 60.38 | 0.000 | 20.000 | 20.0000 | 0.000 | 12.000 | 12.000 |
| ## | 2605 | 191 | 47.64 | NaN | 0.000 | NaN | NaN | 0.000 | NaN |
| ## | 2606 | 39 | 61.54 | NaN | 0.000 | NaN | NaN | 0.000 | NaN |
| ## | 2607 | 154 | 82.47 | NaN | 25.000 | NaN | NaN | 12.500 | NaN |
| ## | 2608 | 111 | 84.68 | NaN | 25.000 | NaN | NaN | 25.000 | NaN |
| ## | 2609 | 10 | 50.00 | NaN | NaN | NaN | NaN | NaN | NaN |
| ## | 2610 | 183 | 44.81 | 0.000 | 0.000 | 0.0000 | 0.000 | 0.000 | 0.000 |
| ## | 2611 | 433 | 46.42 | 0.000 | 12.500 | 12.5000 | 0.000 | 12.500 | 12.500 |
| ## | 2612 | 90 | 71.11 | NaN | 100.000 | NaN | NaN | 100.000 | NaN |

| | | | | | | | | | |
|----|------|------|--------|---------|---------|----------|---------|---------|---------|
| ## | 2613 | 318 | 60.38 | 20.000 | 22.222 | 2.2222 | 20.000 | 11.111 | -8.889 |
| ## | 2614 | 209 | 71.29 | 0.000 | 25.000 | 25.0000 | 0.000 | 0.000 | 0.000 |
| ## | 2700 | 5068 | 29.36 | 28.571 | 48.235 | 19.6639 | 7.143 | 32.941 | 25.798 |
| ## | 2701 | 670 | 60.30 | 20.000 | 72.973 | 52.9730 | 20.000 | 37.838 | 17.838 |
| ## | 2702 | 126 | 47.62 | 0.000 | 75.000 | 75.0000 | 0.000 | 25.000 | 25.000 |
| ## | 2703 | 461 | 45.55 | 100.000 | 50.000 | -50.0000 | 20.000 | 40.000 | 20.000 |
| ## | 2704 | 119 | 42.86 | 50.000 | NaN | NaN | 0.000 | NaN | NaN |
| ## | 2705 | 830 | 60.36 | 15.385 | 34.483 | 19.0981 | 7.692 | 17.241 | 9.549 |
| ## | 2706 | 272 | 44.49 | 0.000 | 33.333 | 33.3333 | 0.000 | 33.333 | 33.333 |
| ## | 2707 | 73 | 78.08 | NaN | 66.667 | NaN | NaN | 0.000 | NaN |
| ## | 2708 | 309 | 32.69 | 0.000 | 90.000 | 90.0000 | 50.000 | 10.000 | -40.000 |
| ## | 2709 | 10 | 60.00 | NaN | 100.000 | NaN | NaN | 100.000 | NaN |
| ## | 2710 | 60 | 40.00 | NaN | 50.000 | NaN | NaN | 50.000 | NaN |
| ## | 2711 | 328 | 43.29 | NaN | 25.000 | NaN | NaN | 25.000 | NaN |
| ## | 2712 | 420 | 47.14 | NaN | 64.286 | NaN | NaN | 42.857 | NaN |
| ## | 2713 | 235 | 45.11 | 40.000 | 71.429 | 31.4286 | 60.000 | 57.143 | -2.857 |
| ## | 2714 | 102 | 68.63 | 0.000 | 55.556 | 55.5556 | 0.000 | 44.444 | 44.444 |
| ## | 2715 | 680 | 44.85 | 33.333 | 50.000 | 16.6667 | 0.000 | 21.429 | 21.429 |
| ## | 2716 | 252 | 51.98 | 66.667 | 50.000 | -16.6667 | 0.000 | 37.500 | 37.500 |
| ## | 2717 | 394 | 64.72 | 0.000 | 60.000 | 60.0000 | 0.000 | 40.000 | 40.000 |
| ## | 2718 | 141 | 58.16 | NaN | 50.000 | NaN | NaN | 25.000 | NaN |
| ## | 2719 | 404 | 71.78 | NaN | 62.857 | NaN | NaN | 51.429 | NaN |
| ## | 2720 | 322 | 49.07 | 100.000 | 75.000 | -25.0000 | 0.000 | 25.000 | 25.000 |
| ## | 2721 | 186 | 57.53 | NaN | 33.333 | NaN | NaN | 16.667 | NaN |
| ## | 2722 | 146 | 46.58 | 0.000 | 75.000 | 75.0000 | 0.000 | 75.000 | 75.000 |
| ## | 2723 | 479 | 52.40 | 40.000 | 80.000 | 40.0000 | 0.000 | 26.667 | 26.667 |
| ## | 2724 | 275 | 52.73 | 0.000 | 42.857 | 42.8571 | 0.000 | 42.857 | 42.857 |
| ## | 2725 | 698 | 42.98 | 50.000 | 51.852 | 1.8519 | 50.000 | 48.148 | -1.852 |
| ## | 2726 | 513 | 45.42 | NaN | 50.000 | NaN | NaN | 22.727 | NaN |
| ## | 2727 | 180 | 64.44 | NaN | 33.333 | NaN | NaN | 33.333 | NaN |
| ## | 2728 | 854 | 52.69 | NaN | 51.351 | NaN | NaN | 32.432 | NaN |
| ## | 2729 | 798 | 39.10 | 33.333 | 85.000 | 51.6667 | 16.667 | 60.000 | 43.333 |
| ## | 2730 | 818 | 50.24 | 33.333 | 68.571 | 35.2381 | 33.333 | 54.286 | 20.952 |
| ## | 2731 | 310 | 55.81 | 60.000 | 88.889 | 28.8889 | 20.000 | 44.444 | 24.444 |
| ## | 2732 | 1220 | 50.33 | 0.000 | 35.000 | 35.0000 | 0.000 | 18.333 | 18.333 |
| ## | 2733 | 393 | 34.35 | NaN | 55.556 | NaN | NaN | 55.556 | NaN |
| ## | 2734 | 386 | 58.81 | 50.000 | 55.556 | 5.5556 | 0.000 | 38.889 | 38.889 |
| ## | 2735 | 1095 | 52.05 | 50.000 | 64.865 | 14.8649 | 25.000 | 35.135 | 10.135 |
| ## | 2736 | 494 | 54.05 | 0.000 | 65.217 | 65.2174 | 0.000 | 39.130 | 39.130 |
| ## | 2737 | 263 | 60.46 | 100.000 | 18.750 | -81.2500 | 0.000 | 31.250 | 31.250 |
| ## | 2738 | 1580 | 61.58 | 41.176 | 55.385 | 14.2081 | 17.647 | 46.154 | 28.507 |
| ## | 2739 | 1054 | 51.80 | 0.000 | 72.727 | 72.7273 | 100.000 | 56.364 | -43.636 |
| ## | 2740 | 398 | 48.24 | 100.000 | 50.000 | -50.0000 | 0.000 | 16.667 | 16.667 |
| ## | 2741 | 963 | 45.38 | 100.000 | 34.615 | -65.3846 | 50.000 | 38.462 | -11.538 |
| ## | 2742 | 412 | 54.85 | 0.000 | 66.667 | 66.6667 | 0.000 | 50.000 | 50.000 |
| ## | 2743 | 223 | 41.70 | 33.333 | 66.667 | 33.3333 | 33.333 | 33.333 | 0.000 |
| ## | 2744 | 3 | 100.00 | 50.000 | NaN | NaN | 0.000 | NaN | NaN |
| ## | 2745 | 202 | 59.90 | NaN | 66.667 | NaN | NaN | 44.444 | NaN |
| ## | 2746 | 2050 | 40.78 | 0.000 | 36.735 | 36.7347 | 0.000 | 18.367 | 18.367 |
| ## | 2747 | 126 | 39.68 | NaN | 40.000 | NaN | NaN | 0.000 | NaN |

| | | | | | | | | |
|---------|------|-------|---------|---------|----------|---------|---------|---------|
| ## 2748 | 327 | 48.01 | 0.000 | 28.571 | 28.5714 | 0.000 | 28.571 | 28.571 |
| ## 2800 | 512 | 54.69 | 50.000 | 50.000 | 0.0000 | 0.000 | 41.667 | 41.667 |
| ## 2801 | 33 | 75.76 | NaN | 50.000 | NaN | NaN | 50.000 | NaN |
| ## 2802 | 227 | 73.13 | 50.000 | 57.143 | 7.1429 | 0.000 | 50.000 | 50.000 |
| ## 2803 | 192 | 60.42 | 25.000 | 72.727 | 47.7273 | 0.000 | 54.545 | 54.545 |
| ## 2804 | 355 | 59.72 | 50.000 | 54.545 | 4.5455 | 16.667 | 45.455 | 28.788 |
| ## 2805 | 158 | 75.95 | 50.000 | 42.857 | -7.1429 | 25.000 | 57.143 | 32.143 |
| ## 2806 | 52 | 67.31 | NaN | NaN | NaN | NaN | NaN | NaN |
| ## 2807 | 84 | 60.71 | NaN | 25.000 | NaN | NaN | 50.000 | NaN |
| ## 2808 | 441 | 54.88 | NaN | 65.217 | NaN | NaN | 43.478 | NaN |
| ## 2809 | 192 | 54.69 | 60.000 | 83.333 | 23.3333 | 20.000 | 50.000 | 30.000 |
| ## 2900 | 615 | 81.95 | 50.000 | 75.862 | 25.8621 | 75.000 | 62.069 | -12.931 |
| ## 2901 | 46 | 58.70 | NaN | 66.667 | NaN | NaN | 83.333 | NaN |
| ## 2902 | 171 | 73.68 | 0.000 | 66.667 | 66.6667 | 0.000 | 50.000 | 50.000 |
| ## 2903 | 16 | 75.00 | 0.000 | 100.000 | 100.0000 | 0.000 | 0.000 | 0.000 |
| ## 2904 | 9 | 88.89 | NaN | 100.000 | NaN | NaN | 100.000 | NaN |
| ## 2905 | 76 | 77.63 | NaN | 57.143 | NaN | NaN | 71.429 | NaN |
| ## 2906 | 24 | 75.00 | NaN | NaN | NaN | NaN | NaN | NaN |
| ## 2907 | 64 | 81.25 | NaN | 50.000 | NaN | NaN | 16.667 | NaN |
| ## 2908 | 24 | 37.50 | NaN | NaN | NaN | NaN | NaN | NaN |
| ## 2909 | 124 | 59.68 | NaN | 50.000 | NaN | NaN | 0.000 | NaN |
| ## 2910 | 58 | 75.86 | NaN | 80.000 | NaN | NaN | 60.000 | NaN |
| ## 2911 | 64 | 79.69 | NaN | 100.000 | NaN | NaN | 0.000 | NaN |
| ## 2912 | 29 | 89.66 | NaN | 75.000 | NaN | NaN | 75.000 | NaN |
| ## 2913 | 112 | 85.71 | NaN | 92.857 | NaN | NaN | 85.714 | NaN |
| ## 2914 | 62 | 83.87 | NaN | 100.000 | NaN | NaN | 100.000 | NaN |
| ## 2915 | 0 | NA | NA | NA | NA | NA | NA | NA |
| ## 2916 | 448 | 51.12 | 100.000 | 84.000 | -16.0000 | 100.000 | 48.000 | -52.000 |
| ## 2917 | 60 | 83.33 | NaN | 100.000 | NaN | NaN | 66.667 | NaN |
| ## 2918 | 0 | NA | NA | NA | NA | NA | NA | NA |
| ## 2919 | 102 | 76.47 | NaN | 54.545 | NaN | NaN | 63.636 | NaN |
| ## 2920 | 3 | 66.67 | NaN | NaN | NaN | NaN | NaN | NaN |
| ## 2921 | 203 | 55.67 | NaN | 62.500 | NaN | NaN | 75.000 | NaN |
| ## 2922 | 35 | 74.29 | NaN | 33.333 | NaN | NaN | 66.667 | NaN |
| ## 2923 | 18 | 72.22 | NaN | NaN | NaN | NaN | NaN | NaN |
| ## 3000 | 74 | 45.95 | NaN | NaN | NaN | NaN | NaN | NaN |
| ## 3001 | 12 | 75.00 | NaN | 100.000 | NaN | NaN | 100.000 | NaN |
| ## 3002 | 522 | 68.77 | 20.000 | 31.250 | 11.2500 | 20.000 | 43.750 | 23.750 |
| ## 3003 | 562 | 60.68 | 33.333 | 55.172 | 21.8391 | 16.667 | 55.172 | 38.506 |
| ## 3004 | 796 | 58.17 | 50.000 | 58.824 | 8.8235 | 25.000 | 41.176 | 16.176 |
| ## 3005 | 239 | 52.72 | 33.333 | 62.500 | 29.1667 | 0.000 | 37.500 | 37.500 |
| ## 3100 | 877 | 39.57 | NaN | 0.000 | NaN | NaN | 20.000 | NaN |
| ## 3101 | 372 | 33.87 | 0.000 | 0.000 | 0.0000 | 0.000 | 0.000 | 0.000 |
| ## 3102 | 128 | 38.28 | NaN | 40.000 | NaN | NaN | 20.000 | NaN |
| ## 3103 | 160 | 40.00 | 0.000 | 33.333 | 33.3333 | 0.000 | 33.333 | 33.333 |
| ## 3104 | 2026 | 38.55 | 33.333 | 18.868 | -14.4654 | 0.000 | 13.208 | 13.208 |
| ## 3105 | 356 | 46.07 | 0.000 | 50.000 | 50.0000 | 0.000 | 10.000 | 10.000 |
| ## 3106 | 296 | 49.32 | 0.000 | 0.000 | 0.0000 | 0.000 | 0.000 | 0.000 |
| ## 3107 | 1014 | 45.46 | 0.000 | 15.000 | 15.0000 | 0.000 | 15.000 | 15.000 |
| ## 3108 | 146 | 36.99 | NaN | 0.000 | NaN | NaN | 0.000 | NaN |

| | | | | | | | | | |
|----|------|------|-------|---------|---------|----------|---------|---------|---------|
| ## | 3109 | 222 | 38.29 | 0.000 | 0.000 | 0.0000 | 0.000 | 0.000 | 0.000 |
| ## | 3110 | 300 | 36.00 | NaN | 40.000 | NaN | NaN | 20.000 | NaN |
| ## | 3200 | 409 | 76.77 | 35.714 | 68.750 | 33.0357 | 28.571 | 37.500 | 8.929 |
| ## | 3201 | 35 | 80.00 | NaN | 33.333 | NaN | NaN | 66.667 | NaN |
| ## | 3202 | 656 | 60.52 | 41.667 | 52.000 | 10.3333 | 16.667 | 48.000 | 31.333 |
| ## | 3203 | 417 | 79.38 | 0.000 | 51.724 | 51.7241 | 20.000 | 51.724 | 31.724 |
| ## | 3204 | 456 | 72.15 | 33.333 | 53.191 | 19.8582 | 0.000 | 57.447 | 57.447 |
| ## | 3205 | 216 | 79.63 | 66.667 | 50.000 | -16.6667 | 0.000 | 50.000 | 50.000 |
| ## | 3206 | 110 | 63.64 | 50.000 | 25.000 | -25.0000 | 50.000 | 75.000 | 25.000 |
| ## | 3207 | 279 | 74.91 | NaN | 52.941 | NaN | NaN | 41.176 | NaN |
| ## | 3300 | 399 | 79.95 | 50.000 | 52.632 | 2.6316 | 50.000 | 42.105 | -7.895 |
| ## | 3301 | 362 | 80.66 | 0.000 | 70.370 | 70.3704 | 0.000 | 51.852 | 51.852 |
| ## | 3302 | 200 | 79.50 | 0.000 | 50.000 | 50.0000 | 0.000 | 50.000 | 50.000 |
| ## | 3303 | 464 | 79.09 | 14.286 | 36.842 | 22.5564 | 28.571 | 26.316 | -2.256 |
| ## | 3304 | 2117 | 74.07 | 25.000 | 59.873 | 34.8726 | 0.000 | 50.318 | 50.318 |
| ## | 3305 | 1206 | 75.95 | 22.222 | 38.095 | 15.8730 | 22.222 | 34.921 | 12.698 |
| ## | 3306 | 496 | 77.62 | 0.000 | 58.333 | 58.3333 | 0.000 | 55.556 | 55.556 |
| ## | 3307 | 73 | 67.12 | NaN | 66.667 | NaN | NaN | 16.667 | NaN |
| ## | 3308 | 741 | 82.73 | 50.000 | 51.429 | 1.4286 | 0.000 | 45.714 | 45.714 |
| ## | 3309 | 315 | 78.41 | 50.000 | 46.154 | -3.8462 | 50.000 | 61.538 | 11.538 |
| ## | 3310 | 668 | 77.99 | 0.000 | 68.750 | 68.7500 | 0.000 | 65.625 | 65.625 |
| ## | 3311 | 80 | 57.50 | NaN | 50.000 | NaN | NaN | 50.000 | NaN |
| ## | 3312 | 1898 | 81.03 | 7.143 | 42.735 | 35.5922 | 14.286 | 41.880 | 27.595 |
| ## | 3313 | 181 | 58.01 | NaN | 28.571 | NaN | NaN | 42.857 | NaN |
| ## | 3314 | 177 | 76.84 | 50.000 | 52.941 | 2.9412 | 50.000 | 47.059 | -2.941 |
| ## | 3315 | 281 | 75.44 | 100.000 | 56.522 | -43.4783 | 100.000 | 60.870 | -39.130 |
| ## | 3316 | 759 | 82.74 | 0.000 | 51.899 | 51.8987 | 0.000 | 51.899 | 51.899 |
| ## | 3317 | 106 | 77.36 | NaN | 66.667 | NaN | NaN | 33.333 | NaN |
| ## | 3318 | 156 | 82.05 | NaN | 100.000 | NaN | NaN | 100.000 | NaN |
| ## | 3319 | 57 | 68.42 | NaN | 66.667 | NaN | NaN | 50.000 | NaN |
| ## | 3320 | 774 | 82.30 | 25.000 | 24.444 | -0.5556 | 25.000 | 22.222 | -2.778 |
| ## | 3321 | 223 | 74.89 | 50.000 | 40.000 | -10.0000 | 50.000 | 46.667 | -3.333 |
| ## | 3322 | 146 | 76.03 | NaN | 23.077 | NaN | NaN | 23.077 | NaN |
| ## | 3400 | 660 | 33.94 | 30.769 | 37.500 | 6.7308 | 30.769 | 75.000 | 44.231 |
| ## | 3401 | 16 | 6.25 | NaN | 0.000 | NaN | NaN | 100.000 | NaN |
| ## | 3402 | 88 | 10.23 | NaN | 50.000 | NaN | NaN | 50.000 | NaN |
| ## | 3403 | 167 | 23.35 | NaN | 66.667 | NaN | NaN | 16.667 | NaN |
| ## | 3404 | 93 | 15.05 | NaN | 50.000 | NaN | NaN | 50.000 | NaN |
| ## | 3500 | 216 | 31.94 | 0.000 | NaN | NaN | 100.000 | NaN | NaN |
| ## | 3501 | 14 | 21.43 | NaN | NaN | NaN | NaN | NaN | NaN |
| ## | 3502 | 0 | NA | NA | NA | NA | NA | NA | NA |
| ## | 3503 | 0 | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| ## | 3504 | 96 | 39.58 | NaN | 75.000 | NaN | NaN | 0.000 | NaN |
| ## | 3505 | 18 | 50.00 | NaN | NaN | NaN | NaN | NaN | NaN |
| ## | 3506 | 7 | 42.86 | NaN | 100.000 | NaN | NaN | 0.000 | NaN |
| ## | 3600 | 35 | 71.43 | NaN | NaN | NaN | NaN | NaN | NaN |
| ## | 3601 | 73 | 72.60 | NaN | 100.000 | NaN | NaN | 100.000 | NaN |
| ## | 3602 | 6 | 50.00 | NaN | 100.000 | NaN | NaN | 66.667 | NaN |
| ## | 3603 | 10 | 70.00 | NaN | 50.000 | NaN | NaN | 50.000 | NaN |
| ## | 3604 | 2 | 0.00 | NaN | NaN | NaN | NaN | NaN | NaN |

```
## 3605      25      48.00      NaN      0.000      NaN      NaN      0.000      NaN
## 3606       0       NA       NA       NA       NA       NA       NA       NA
## 3607      48      56.25      NaN      0.000      NaN      NaN      0.000      NaN
## 3608       0       NaN      NaN      NaN      NaN      NaN      NaN      NaN
## 3609      97      46.39      NaN 100.000      NaN      NaN 100.000      NaN
## 3610      14      78.57      NaN 100.000      NaN      NaN      0.000      NaN
## 3611      25      64.00      NaN      NaN      NaN      NaN      NaN      NaN
## 3612     546      61.36      0.000  46.341  46.3415      0.000  39.024  39.024
## 3613      24      41.67      NaN      NaN      NaN      NaN      NaN      NaN
## 3614     216      43.98      NaN 100.000      NaN      NaN      0.000      NaN
## 3615       0       NA       NA       NA       NA       NA       NA       NA
## 3616      99      79.80      NaN  60.000      NaN      NaN  90.000      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         47 23.404  50.000   0.000 -50.000  14.894  50.000   0.000
## 1100        207 43.961   0.000  44.444  44.444  21.256   0.000  33.333
## 1101         62 37.097  50.000 100.000  50.000  17.742  50.000   0.000
## 1102        130 28.462   0.000  30.000  30.000  22.308  33.333  30.000
## 1103        174 34.483  42.857  29.412 -13.445  18.966  42.857  41.176
## 1104        283 40.989  41.667  38.462  -3.205  22.615  25.000  19.231
## 1105        333 35.435  25.000  34.483   9.483  21.021  33.333  20.690
## 1106        667 45.727  11.111  45.205  34.094  25.037  11.111  34.247
## 1107         87 28.736   0.000  20.000  20.000  11.494 100.000   0.000
## 1108         15 33.333   0.000      NaN      NaN  33.333 100.000      NaN
## 1109         48 31.250      NaN  20.000      NaN  35.417      NaN  80.000
## 1110        126 30.159  33.333  38.462   5.128  18.254  33.333  23.077
## 1111         54 18.519  50.000  33.333 -16.667   5.556  50.000   0.000
## 1200         46 36.957      NaN  62.500      NaN  36.957      NaN  50.000
## 1201        207 48.309      NaN  34.783      NaN  43.961      NaN  39.130
## 1202        387 31.266   0.000  41.667  41.667  32.817   0.000  41.667
## 1203        278 42.446  25.000  38.710  13.710  39.928  25.000  35.484
## 1204         91 32.967   0.000  37.500  37.500  31.868   0.000  37.500
## 1205         39 23.077   0.000  40.000  40.000  23.077   0.000  40.000
## 1206          5 80.000      NaN      NaN      NaN  80.000      NaN      NaN
## 1207        249 45.783  57.143  50.000  -7.143  30.120  14.286  39.286
## 1208        256 39.062   0.000  48.780  48.780  39.453   0.000  48.780
## 1209          4 75.000      NaN      NaN      NaN  75.000      NaN      NaN
## 1210         38 34.211 100.000  87.500 -12.500  31.579   0.000  87.500
## 1211        217 22.120  11.111  35.714  24.603  20.276  22.222  35.714
## 1212        102 18.627      NaN   0.000      NaN  17.647      NaN   0.000
## 1213        105 44.762      NaN  50.000      NaN  44.762      NaN  50.000
## 1300        181 36.464 100.000  39.130 -60.870  19.337 100.000  28.261
## 1301          3  0.000      NaN   0.000      NaN   0.000      NaN   0.000
## 1302         37 51.351   0.000  50.000  50.000  37.838   0.000  66.667
```


| | | | | | | | | |
|---------|------|--------|---------|---------|---------|--------|---------|---------|
| ## 1303 | 1042 | 43.474 | 41.667 | 40.000 | -1.667 | 23.512 | 14.583 | 25.263 |
| ## 1304 | 225 | 41.778 | 42.857 | 17.647 | -25.210 | 26.667 | 57.143 | 20.588 |
| ## 1305 | 371 | 43.396 | 20.000 | 41.304 | 21.304 | 18.598 | 0.000 | 17.391 |
| ## 1306 | 164 | 56.707 | 0.000 | 52.941 | 52.941 | 28.049 | 0.000 | 29.412 |
| ## 1307 | 322 | 55.901 | 63.636 | 51.613 | -12.023 | 31.056 | 18.182 | 32.258 |
| ## 1308 | 179 | 46.927 | 42.857 | 66.667 | 23.810 | 24.022 | 0.000 | 22.222 |
| ## 1309 | 74 | 55.405 | NaN | 37.500 | NaN | 37.838 | NaN | 62.500 |
| ## 1310 | 95 | 51.579 | NaN | 70.000 | NaN | 21.053 | NaN | 10.000 |
| ## 1311 | 353 | 45.042 | 55.556 | 62.963 | 7.407 | 18.130 | 11.111 | 25.926 |
| ## 1312 | 650 | 47.385 | 42.857 | 41.270 | -1.587 | 23.538 | 21.429 | 20.635 |
| ## 1313 | 164 | 43.902 | 0.000 | 42.857 | 42.857 | 23.780 | 33.333 | 4.762 |
| ## 1314 | 153 | 45.752 | 100.000 | 52.941 | -47.059 | 15.033 | 0.000 | 17.647 |
| ## 1315 | 54 | 35.185 | 50.000 | 75.000 | 25.000 | 22.222 | 100.000 | 25.000 |
| ## 1400 | 179 | 32.961 | 0.000 | 27.778 | 27.778 | 23.464 | 0.000 | 16.667 |
| ## 1401 | 75 | 50.667 | NaN | 33.333 | NaN | 48.000 | NaN | 33.333 |
| ## 1402 | 61 | 26.230 | 66.667 | 0.000 | -66.667 | 21.311 | 66.667 | 14.286 |
| ## 1403 | 144 | 38.889 | NaN | 35.294 | NaN | 30.556 | NaN | 35.294 |
| ## 1404 | 38 | 21.053 | 0.000 | 0.000 | 0.000 | 15.789 | 0.000 | 0.000 |
| ## 1405 | 130 | 20.769 | 50.000 | 25.000 | -25.000 | 23.077 | 50.000 | 37.500 |
| ## 1406 | 97 | 40.206 | NaN | 62.500 | NaN | 37.113 | NaN | 50.000 |
| ## 1407 | 152 | 31.579 | 66.667 | 76.923 | 10.256 | 28.289 | 66.667 | 61.538 |
| ## 1408 | 273 | 26.740 | 14.286 | 43.478 | 29.193 | 22.344 | 14.286 | 30.435 |
| ## 1409 | 52 | 46.154 | NaN | 33.333 | NaN | 48.077 | NaN | 33.333 |
| ## 1410 | 29 | 41.379 | 100.000 | 50.000 | -50.000 | 41.379 | 100.000 | 50.000 |
| ## 1500 | 180 | 31.111 | 14.286 | 33.333 | 19.048 | 13.889 | 14.286 | 20.833 |
| ## 1501 | 4 | 50.000 | NaN | NaN | NaN | 50.000 | NaN | NaN |
| ## 1502 | 208 | 37.500 | 50.000 | 32.432 | -17.568 | 9.615 | 0.000 | 8.108 |
| ## 1503 | 178 | 32.022 | 14.286 | 20.833 | 6.548 | 15.169 | 14.286 | 12.500 |
| ## 1504 | 3 | 0.000 | NaN | 0.000 | NaN | 0.000 | NaN | 0.000 |
| ## 1505 | 51 | 35.294 | NaN | 0.000 | NaN | 9.804 | NaN | 0.000 |
| ## 1506 | 19 | 31.579 | NaN | 100.000 | NaN | 15.789 | NaN | 0.000 |
| ## 1507 | 36 | 16.667 | NaN | 25.000 | NaN | 38.889 | NaN | 75.000 |
| ## 1508 | 55 | 29.091 | 25.000 | 25.000 | 0.000 | 12.727 | 0.000 | 25.000 |
| ## 1600 | 635 | 34.173 | 16.667 | 29.885 | 13.218 | 17.953 | 33.333 | 18.391 |
| ## 1601 | 27 | 33.333 | NaN | 0.000 | NaN | 18.519 | NaN | 0.000 |
| ## 1602 | 378 | 36.243 | 18.750 | 44.118 | 25.368 | 15.079 | 6.250 | 17.647 |
| ## 1603 | 206 | 30.583 | 16.667 | 31.250 | 14.583 | 11.650 | 0.000 | 12.500 |
| ## 1604 | 187 | 30.481 | 16.667 | 13.333 | -3.333 | 20.856 | 0.000 | 13.333 |
| ## 1605 | 556 | 35.612 | 13.333 | 27.083 | 13.750 | 20.504 | 13.333 | 20.833 |
| ## 1606 | 489 | 25.971 | 15.385 | 20.000 | 4.615 | 14.519 | 0.000 | 17.778 |
| ## 1607 | 183 | 37.705 | 50.000 | 33.333 | -16.667 | 8.197 | 0.000 | 5.556 |
| ## 1700 | 112 | 21.429 | NaN | 0.000 | NaN | 16.071 | NaN | 0.000 |
| ## 1701 | 15 | 20.000 | NaN | 50.000 | NaN | 20.000 | NaN | 0.000 |
| ## 1702 | 131 | 14.504 | 0.000 | 16.667 | 16.667 | 10.687 | 0.000 | 16.667 |
| ## 1703 | 94 | 12.766 | NaN | 0.000 | NaN | 13.830 | NaN | 25.000 |
| ## 1704 | 40 | 12.500 | NaN | 0.000 | NaN | 22.500 | NaN | 100.000 |
| ## 1705 | 179 | 13.966 | 50.000 | 14.286 | -35.714 | 8.380 | 50.000 | 0.000 |
| ## 1706 | 375 | 14.667 | 0.000 | 6.061 | 6.061 | 9.600 | 0.000 | 21.212 |
| ## 1707 | 53 | 9.434 | NaN | 0.000 | NaN | 7.547 | NaN | 20.000 |
| ## 1708 | 99 | 8.081 | 0.000 | 14.286 | 14.286 | 6.061 | 0.000 | 14.286 |

| | | | | | | | | |
|---------|-----|--------|---------|---------|---------|--------|---------|---------|
| ## 1709 | 56 | 25.000 | 100.000 | 20.000 | -80.000 | 12.500 | 100.000 | 0.000 |
| ## 1710 | 199 | 19.598 | 33.333 | 5.882 | -27.451 | 17.085 | 33.333 | 17.647 |
| ## 1711 | 54 | 11.111 | 0.000 | 33.333 | 33.333 | 3.704 | 0.000 | 33.333 |
| ## 1712 | 263 | 9.886 | 0.000 | 12.500 | 12.500 | 7.985 | 0.000 | 8.333 |
| ## 1800 | 27 | 7.407 | NaN | 0.000 | NaN | 11.111 | NaN | 33.333 |
| ## 1801 | 0 | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| ## 1802 | 48 | 12.500 | 0.000 | 0.000 | 0.000 | 12.500 | 0.000 | 0.000 |
| ## 1803 | 70 | 15.714 | 0.000 | 0.000 | 0.000 | 8.571 | 0.000 | 0.000 |
| ## 1804 | 69 | 21.739 | 0.000 | 25.000 | 25.000 | 13.043 | 0.000 | 0.000 |
| ## 1900 | 162 | 29.630 | 11.111 | 40.000 | 28.889 | 22.840 | 22.222 | 20.000 |
| ## 1901 | 42 | 19.048 | NaN | 0.000 | NaN | 19.048 | NaN | 100.000 |
| ## 1902 | 70 | 11.429 | 0.000 | 0.000 | 0.000 | 4.286 | 100.000 | 0.000 |
| ## 1903 | 14 | 7.143 | NaN | 0.000 | NaN | 28.571 | NaN | 75.000 |
| ## 1904 | 57 | 36.842 | NaN | 66.667 | NaN | 24.561 | NaN | 66.667 |
| ## 1905 | 7 | 14.286 | NaN | NaN | NaN | 28.571 | NaN | NaN |
| ## 1906 | 39 | 25.641 | NaN | 50.000 | NaN | 5.128 | NaN | 0.000 |
| ## 1907 | 81 | 14.815 | 0.000 | 100.000 | 100.000 | 3.704 | 0.000 | 0.000 |
| ## 1908 | 47 | 14.894 | 0.000 | 25.000 | 25.000 | 8.511 | 0.000 | 0.000 |
| ## 1909 | 47 | 4.255 | 0.000 | 0.000 | 0.000 | 2.128 | 0.000 | 0.000 |
| ## 1910 | 92 | 27.174 | 0.000 | 14.286 | 14.286 | 15.217 | 0.000 | 0.000 |
| ## 1911 | 35 | 28.571 | NaN | 50.000 | NaN | 14.286 | NaN | 0.000 |
| ## 1912 | 52 | 19.231 | 0.000 | 50.000 | 50.000 | 11.538 | 0.000 | 0.000 |
| ## 1913 | 8 | 0.000 | NaN | NaN | NaN | 25.000 | NaN | NaN |
| ## 2000 | 61 | 24.590 | NaN | 33.333 | NaN | 18.033 | NaN | 0.000 |
| ## 2001 | 29 | 27.586 | NaN | 33.333 | NaN | 20.690 | NaN | 16.667 |
| ## 2002 | 596 | 25.839 | 14.286 | 27.869 | 13.583 | 25.336 | 14.286 | 19.672 |
| ## 2003 | 147 | 19.728 | 0.000 | 15.385 | 15.385 | 19.048 | 0.000 | 15.385 |
| ## 2100 | 72 | 27.778 | NaN | 41.667 | NaN | 16.667 | NaN | 8.333 |
| ## 2101 | 2 | 50.000 | NaN | NaN | NaN | 0.000 | NaN | NaN |
| ## 2102 | 44 | 31.818 | NaN | 25.000 | NaN | 6.818 | NaN | 0.000 |
| ## 2103 | 19 | 26.316 | 0.000 | NaN | NaN | 21.053 | 0.000 | NaN |
| ## 2104 | 5 | 20.000 | 0.000 | NaN | NaN | 20.000 | 0.000 | NaN |
| ## 2105 | 81 | 39.506 | 50.000 | 42.857 | -7.143 | 17.284 | 50.000 | 0.000 |
| ## 2200 | 140 | 12.857 | 0.000 | 18.750 | 18.750 | 9.286 | 0.000 | 12.500 |
| ## 2201 | 27 | 25.926 | NaN | 25.000 | NaN | 14.815 | NaN | 25.000 |
| ## 2202 | 34 | 14.706 | 0.000 | 25.000 | 25.000 | 17.647 | 0.000 | 0.000 |
| ## 2203 | 12 | 0.000 | NaN | 0.000 | NaN | 16.667 | NaN | 100.000 |
| ## 2204 | 209 | 23.923 | 0.000 | 16.667 | 16.667 | 8.134 | 0.000 | 9.524 |
| ## 2205 | 82 | 14.634 | NaN | 18.182 | NaN | 12.195 | NaN | 0.000 |
| ## 2206 | 21 | 19.048 | NaN | NaN | NaN | 4.762 | NaN | NaN |
| ## 2207 | 90 | 8.889 | 50.000 | 12.500 | -37.500 | 3.333 | 0.000 | 6.250 |
| ## 2208 | 545 | 15.596 | 0.000 | 10.145 | 10.145 | 7.890 | 0.000 | 4.348 |
| ## 2209 | 175 | 31.429 | NaN | 28.571 | NaN | 21.143 | NaN | 28.571 |
| ## 2210 | 205 | 18.537 | 0.000 | 28.571 | 28.571 | 13.659 | 0.000 | 14.286 |
| ## 2211 | 155 | 22.581 | 0.000 | 29.412 | 29.412 | 8.387 | 0.000 | 0.000 |
| ## 2212 | 19 | 10.526 | NaN | 0.000 | NaN | 10.526 | NaN | 0.000 |
| ## 2213 | 79 | 24.051 | 0.000 | 35.714 | 35.714 | 15.190 | 0.000 | 28.571 |
| ## 2214 | 46 | 4.348 | NaN | 0.000 | NaN | 6.522 | NaN | 0.000 |
| ## 2215 | 39 | 12.821 | NaN | 12.500 | NaN | 10.256 | NaN | 0.000 |
| ## 2216 | 9 | 0.000 | NaN | NaN | NaN | 0.000 | NaN | NaN |

| | | | | | | | | |
|---------|-----|--------|---------|--------|---------|--------|--------|--------|
| ## 2300 | 218 | 30.734 | 20.000 | 34.615 | 14.615 | 16.972 | 20.000 | 19.231 |
| ## 2301 | 42 | 42.857 | NaN | 20.000 | NaN | 33.333 | NaN | 20.000 |
| ## 2302 | 25 | 48.000 | NaN | 28.571 | NaN | 4.000 | NaN | 0.000 |
| ## 2303 | 241 | 42.324 | 0.000 | 36.364 | 36.364 | 14.938 | 0.000 | 22.727 |
| ## 2304 | 178 | 32.022 | 50.000 | 35.000 | -15.000 | 13.483 | 50.000 | 15.000 |
| ## 2305 | 71 | 35.211 | NaN | 27.273 | NaN | 21.127 | NaN | 18.182 |
| ## 2306 | 24 | 33.333 | 0.000 | 50.000 | 50.000 | 4.167 | 0.000 | 0.000 |
| ## 2307 | 55 | 40.000 | 50.000 | 0.000 | -50.000 | 30.909 | 0.000 | 33.333 |
| ## 2308 | 227 | 37.445 | 50.000 | 29.630 | -20.370 | 26.872 | 50.000 | 18.519 |
| ## 2309 | 85 | 44.706 | 0.000 | 11.111 | 11.111 | 17.647 | 0.000 | 22.222 |
| ## 2310 | 84 | 28.571 | 0.000 | 25.000 | 25.000 | 11.905 | 0.000 | 16.667 |
| ## 2311 | 78 | 44.872 | 0.000 | 57.143 | 57.143 | 17.949 | 0.000 | 14.286 |
| ## 2312 | 92 | 28.261 | 0.000 | 23.077 | 23.077 | 19.565 | 0.000 | 30.769 |
| ## 2400 | 28 | 35.714 | NaN | 40.000 | NaN | 14.286 | NaN | 20.000 |
| ## 2401 | 0 | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| ## 2402 | 277 | 50.542 | 27.273 | 45.833 | 18.561 | 14.440 | 9.091 | 12.500 |
| ## 2403 | 311 | 53.376 | 58.333 | 67.742 | 9.409 | 18.971 | 16.667 | 29.032 |
| ## 2404 | 367 | 53.678 | 66.667 | 61.111 | -5.556 | 16.894 | 0.000 | 25.000 |
| ## 2405 | 66 | 48.485 | 100.000 | 83.333 | -16.667 | 22.727 | 0.000 | 33.333 |
| ## 2406 | 74 | 52.703 | NaN | 63.636 | NaN | 10.811 | NaN | 18.182 |
| ## 2500 | 268 | 25.746 | 0.000 | 20.930 | 20.930 | 13.806 | 0.000 | 13.953 |
| ## 2501 | 9 | 33.333 | NaN | NaN | NaN | 33.333 | NaN | NaN |
| ## 2502 | 114 | 26.316 | NaN | 28.000 | NaN | 5.263 | NaN | 4.000 |
| ## 2503 | 133 | 25.564 | 0.000 | 22.727 | 22.727 | 14.286 | 50.000 | 9.091 |
| ## 2504 | 384 | 23.177 | 33.333 | 15.000 | -18.333 | 10.156 | 0.000 | 6.667 |
| ## 2505 | 398 | 26.633 | 20.000 | 22.917 | 2.917 | 11.558 | 10.000 | 8.333 |
| ## 2506 | 146 | 34.247 | 0.000 | 33.333 | 33.333 | 18.493 | 50.000 | 8.333 |
| ## 2507 | 66 | 15.152 | 33.333 | 11.111 | -22.222 | 3.030 | 0.000 | 0.000 |
| ## 2508 | 248 | 29.032 | 0.000 | 30.303 | 30.303 | 15.323 | 20.000 | 15.152 |
| ## 2600 | 215 | 14.419 | 8.333 | 25.000 | 16.667 | 12.558 | 8.333 | 15.000 |
| ## 2601 | 12 | 25.000 | NaN | 0.000 | NaN | 33.333 | NaN | 0.000 |
| ## 2602 | 129 | 8.527 | 0.000 | 11.765 | 11.765 | 18.605 | 0.000 | 5.882 |
| ## 2603 | 100 | 8.000 | 0.000 | 0.000 | 0.000 | 6.000 | 0.000 | 0.000 |
| ## 2604 | 290 | 12.414 | 0.000 | 15.385 | 15.385 | 9.310 | 0.000 | 7.692 |
| ## 2605 | 57 | 14.035 | NaN | 0.000 | NaN | 15.789 | NaN | 12.500 |
| ## 2606 | 5 | 20.000 | NaN | 0.000 | NaN | 0.000 | NaN | 0.000 |
| ## 2607 | 84 | 8.333 | NaN | 0.000 | NaN | 22.619 | NaN | 22.222 |
| ## 2608 | 56 | 8.929 | NaN | 0.000 | NaN | 26.786 | NaN | 14.286 |
| ## 2609 | 3 | 0.000 | NaN | 0.000 | NaN | 0.000 | NaN | 0.000 |
| ## 2610 | 65 | 1.538 | 0.000 | 0.000 | 0.000 | 1.538 | 0.000 | 0.000 |
| ## 2611 | 133 | 17.293 | 0.000 | 13.043 | 13.043 | 9.774 | 0.000 | 17.391 |
| ## 2612 | 46 | 6.522 | NaN | 0.000 | NaN | 28.261 | NaN | 0.000 |
| ## 2613 | 138 | 15.942 | 20.000 | 20.000 | 0.000 | 10.145 | 20.000 | 20.000 |
| ## 2614 | 111 | 11.712 | 0.000 | 0.000 | 0.000 | 11.712 | 0.000 | 0.000 |
| ## 2700 | 712 | 40.730 | 28.571 | 42.593 | 14.021 | 27.669 | 7.143 | 30.556 |
| ## 2701 | 190 | 56.842 | 20.000 | 39.286 | 19.286 | 34.737 | 20.000 | 28.571 |
| ## 2702 | 36 | 47.222 | 0.000 | 50.000 | 50.000 | 41.667 | 0.000 | 50.000 |
| ## 2703 | 150 | 43.333 | 100.000 | 38.462 | -61.538 | 23.333 | 20.000 | 7.692 |
| ## 2704 | 32 | 43.750 | 50.000 | 75.000 | 25.000 | 21.875 | 0.000 | 50.000 |
| ## 2705 | 333 | 27.628 | 15.385 | 27.778 | 12.393 | 16.216 | 7.692 | 22.222 |

| | | | | | | | | |
|---------|-----|--------|---------|---------|---------|--------|---------|---------|
| ## 2706 | 85 | 40.000 | 0.000 | 37.500 | 37.500 | 10.588 | 0.000 | 12.500 |
| ## 2707 | 46 | 13.043 | NaN | 20.000 | NaN | 15.217 | NaN | 20.000 |
| ## 2708 | 51 | 68.627 | 0.000 | 40.000 | 40.000 | 27.451 | 50.000 | 20.000 |
| ## 2709 | 2 | 0.000 | NaN | 0.000 | NaN | 0.000 | NaN | 0.000 |
| ## 2710 | 13 | 46.154 | NaN | 50.000 | NaN | 38.462 | NaN | 100.000 |
| ## 2711 | 87 | 34.483 | NaN | 41.667 | NaN | 17.241 | NaN | 16.667 |
| ## 2712 | 95 | 56.842 | NaN | 62.500 | NaN | 30.526 | NaN | 25.000 |
| ## 2713 | 52 | 48.077 | 40.000 | 75.000 | 35.000 | 42.308 | 60.000 | 50.000 |
| ## 2714 | 38 | 52.632 | 0.000 | 50.000 | 50.000 | 23.684 | 0.000 | 50.000 |
| ## 2715 | 197 | 36.041 | 33.333 | 48.000 | 14.667 | 11.675 | 0.000 | 16.000 |
| ## 2716 | 87 | 49.425 | 66.667 | 58.333 | -8.333 | 21.839 | 0.000 | 25.000 |
| ## 2717 | 121 | 48.760 | 0.000 | 63.636 | 63.636 | 28.099 | 0.000 | 40.909 |
| ## 2718 | 39 | 35.897 | NaN | 50.000 | NaN | 20.513 | NaN | 16.667 |
| ## 2719 | 119 | 52.101 | NaN | 56.522 | NaN | 43.697 | NaN | 52.174 |
| ## 2720 | 111 | 53.153 | 100.000 | 87.500 | -12.500 | 26.126 | 0.000 | 37.500 |
| ## 2721 | 74 | 39.189 | NaN | 37.500 | NaN | 10.811 | NaN | 0.000 |
| ## 2722 | 39 | 43.590 | 0.000 | 50.000 | 50.000 | 43.590 | 0.000 | 0.000 |
| ## 2723 | 163 | 57.055 | 40.000 | 63.158 | 23.158 | 16.564 | 0.000 | 26.316 |
| ## 2724 | 85 | 43.529 | 0.000 | 33.333 | 33.333 | 15.294 | 0.000 | 16.667 |
| ## 2725 | 176 | 55.114 | 50.000 | 58.333 | 8.333 | 23.864 | 50.000 | 33.333 |
| ## 2726 | 139 | 53.957 | NaN | 56.522 | NaN | 28.058 | NaN | 21.739 |
| ## 2727 | 74 | 47.297 | NaN | 55.556 | NaN | 14.865 | NaN | 33.333 |
| ## 2728 | 239 | 36.820 | NaN | 36.000 | NaN | 25.105 | NaN | 24.000 |
| ## 2729 | 169 | 57.988 | 33.333 | 77.778 | 44.444 | 22.485 | 16.667 | 38.889 |
| ## 2730 | 222 | 54.955 | 33.333 | 41.379 | 8.046 | 25.676 | 33.333 | 27.586 |
| ## 2731 | 122 | 28.689 | 60.000 | 66.667 | 6.667 | 17.213 | 20.000 | 25.000 |
| ## 2732 | 355 | 24.789 | 0.000 | 14.583 | 14.583 | 17.465 | 0.000 | 16.667 |
| ## 2733 | 84 | 25.000 | NaN | 20.000 | NaN | 8.333 | NaN | 0.000 |
| ## 2734 | 129 | 57.364 | 50.000 | 63.636 | 13.636 | 46.512 | 0.000 | 27.273 |
| ## 2735 | 305 | 50.820 | 50.000 | 53.125 | 3.125 | 28.852 | 25.000 | 46.875 |
| ## 2736 | 115 | 46.087 | 0.000 | 56.250 | 56.250 | 22.609 | 0.000 | 25.000 |
| ## 2737 | 86 | 40.698 | 100.000 | 50.000 | -50.000 | 17.442 | 0.000 | 25.000 |
| ## 2738 | 578 | 45.156 | 41.176 | 53.571 | 12.395 | 30.104 | 17.647 | 34.524 |
| ## 2739 | 265 | 56.226 | 0.000 | 53.659 | 53.659 | 44.528 | 100.000 | 46.341 |
| ## 2740 | 114 | 28.070 | 100.000 | 7.692 | -92.308 | 5.263 | 0.000 | 7.692 |
| ## 2741 | 280 | 36.071 | 100.000 | 31.034 | -68.966 | 22.500 | 50.000 | 3.448 |
| ## 2742 | 119 | 64.706 | 0.000 | 57.143 | 57.143 | 50.420 | 0.000 | 50.000 |
| ## 2743 | 53 | 60.377 | 33.333 | 66.667 | 33.333 | 32.075 | 33.333 | 50.000 |
| ## 2744 | 3 | 33.333 | 50.000 | NaN | NaN | 0.000 | 0.000 | NaN |
| ## 2745 | 75 | 26.667 | NaN | 18.182 | NaN | 33.333 | NaN | 18.182 |
| ## 2746 | 500 | 22.800 | 0.000 | 26.786 | 26.786 | 8.200 | 0.000 | 5.357 |
| ## 2747 | 34 | 47.059 | NaN | 50.000 | NaN | 8.824 | NaN | 0.000 |
| ## 2748 | 112 | 32.143 | 0.000 | 28.571 | 28.571 | 8.929 | 0.000 | 35.714 |
| ## 2800 | 184 | 39.130 | 50.000 | 50.000 | 0.000 | 19.022 | 0.000 | 22.222 |
| ## 2801 | 4 | 50.000 | NaN | 33.333 | NaN | 75.000 | NaN | 66.667 |
| ## 2802 | 75 | 52.000 | 50.000 | 80.000 | 30.000 | 14.667 | 0.000 | 0.000 |
| ## 2803 | 68 | 50.000 | 25.000 | 71.429 | 46.429 | 13.235 | 0.000 | 7.143 |
| ## 2804 | 141 | 48.936 | 50.000 | 68.750 | 18.750 | 24.113 | 16.667 | 50.000 |
| ## 2805 | 69 | 36.232 | 50.000 | 57.143 | 7.143 | 18.841 | 25.000 | 28.571 |
| ## 2806 | 18 | 55.556 | NaN | 100.000 | NaN | 50.000 | NaN | 100.000 |

| | | | | | | | | |
|---------|-----|---------|---------|---------|---------|--------|---------|---------|
| ## 2807 | 25 | 68.000 | NaN | 87.500 | NaN | 24.000 | NaN | 12.500 |
| ## 2808 | 127 | 42.520 | NaN | 43.750 | NaN | 23.622 | NaN | 18.750 |
| ## 2809 | 82 | 24.390 | 60.000 | 50.000 | -10.000 | 18.293 | 20.000 | 50.000 |
| ## 2900 | 346 | 72.254 | 50.000 | 65.789 | 15.789 | 72.543 | 75.000 | 78.947 |
| ## 2901 | 13 | 76.923 | NaN | 75.000 | NaN | 61.538 | NaN | 62.500 |
| ## 2902 | 76 | 71.053 | 0.000 | 57.143 | 57.143 | 60.526 | 0.000 | 57.143 |
| ## 2903 | 7 | 14.286 | 0.000 | NaN | NaN | 14.286 | 0.000 | NaN |
| ## 2904 | 5 | 40.000 | NaN | 0.000 | NaN | 60.000 | NaN | 0.000 |
| ## 2905 | 30 | 76.667 | NaN | 70.000 | NaN | 76.667 | NaN | 70.000 |
| ## 2906 | 16 | 93.750 | NaN | NaN | NaN | 87.500 | NaN | NaN |
| ## 2907 | 27 | 62.963 | NaN | 100.000 | NaN | 55.556 | NaN | 100.000 |
| ## 2908 | 4 | 100.000 | NaN | NaN | NaN | 75.000 | NaN | NaN |
| ## 2909 | 36 | 63.889 | NaN | 71.429 | NaN | 41.667 | NaN | 28.571 |
| ## 2910 | 18 | 66.667 | NaN | 60.000 | NaN | 72.222 | NaN | 60.000 |
| ## 2911 | 35 | 80.000 | NaN | 66.667 | NaN | 65.714 | NaN | 66.667 |
| ## 2912 | 10 | 80.000 | NaN | 100.000 | NaN | 60.000 | NaN | 100.000 |
| ## 2913 | 36 | 83.333 | NaN | 80.000 | NaN | 69.444 | NaN | 20.000 |
| ## 2914 | 32 | 81.250 | NaN | 60.000 | NaN | 71.875 | NaN | 40.000 |
| ## 2915 | 0 | NA | NA | NA | NA | NA | NA | NA |
| ## 2916 | 100 | 74.000 | 100.000 | 58.333 | -41.667 | 40.000 | 100.000 | 58.333 |
| ## 2917 | 28 | 78.571 | NaN | 66.667 | NaN | 71.429 | NaN | 50.000 |
| ## 2918 | 0 | NA | NA | NA | NA | NA | NA | NA |
| ## 2919 | 29 | 75.862 | NaN | 71.429 | NaN | 68.966 | NaN | 71.429 |
| ## 2920 | 2 | 0.000 | NaN | NaN | NaN | 0.000 | NaN | NaN |
| ## 2921 | 48 | 58.333 | NaN | 55.556 | NaN | 37.500 | NaN | 11.111 |
| ## 2922 | 13 | 84.615 | NaN | 100.000 | NaN | 76.923 | NaN | 100.000 |
| ## 2923 | 7 | 57.143 | NaN | 66.667 | NaN | 71.429 | NaN | 33.333 |
| ## 3000 | 25 | 40.000 | NaN | 50.000 | NaN | 20.000 | NaN | 50.000 |
| ## 3001 | 5 | 0.000 | NaN | 0.000 | NaN | 20.000 | NaN | 100.000 |
| ## 3002 | 256 | 35.938 | 20.000 | 43.478 | 23.478 | 28.516 | 20.000 | 30.435 |
| ## 3003 | 186 | 42.473 | 33.333 | 43.478 | 10.145 | 30.108 | 16.667 | 39.130 |
| ## 3004 | 315 | 45.079 | 50.000 | 64.706 | 14.706 | 17.460 | 25.000 | 32.353 |
| ## 3005 | 81 | 50.617 | 33.333 | 60.000 | 26.667 | 30.864 | 0.000 | 60.000 |
| ## 3100 | 223 | 13.901 | NaN | 11.111 | NaN | 7.623 | NaN | 7.407 |
| ## 3101 | 91 | 10.989 | 0.000 | 33.333 | 33.333 | 7.692 | 0.000 | 0.000 |
| ## 3102 | 29 | 6.897 | NaN | 0.000 | NaN | 10.345 | NaN | 0.000 |
| ## 3103 | 46 | 19.565 | 0.000 | 50.000 | 50.000 | 13.043 | 0.000 | 0.000 |
| ## 3104 | 438 | 18.950 | 33.333 | 15.254 | -18.079 | 12.329 | 0.000 | 20.339 |
| ## 3105 | 105 | 24.762 | 0.000 | 15.385 | 15.385 | 16.190 | 0.000 | 15.385 |
| ## 3106 | 124 | 5.645 | 0.000 | 14.286 | 14.286 | 4.839 | 0.000 | 0.000 |
| ## 3107 | 287 | 11.847 | 0.000 | 2.500 | 2.500 | 6.620 | 0.000 | 5.000 |
| ## 3108 | 35 | 42.857 | NaN | 0.000 | NaN | 34.286 | NaN | 0.000 |
| ## 3109 | 64 | 1.562 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| ## 3110 | 67 | 28.358 | NaN | 0.000 | NaN | 8.955 | NaN | 0.000 |
| ## 3200 | 203 | 47.291 | 35.714 | 60.000 | 24.286 | 39.409 | 28.571 | 40.000 |
| ## 3201 | 11 | 18.182 | NaN | NaN | NaN | 18.182 | NaN | NaN |
| ## 3202 | 275 | 46.182 | 41.667 | 53.571 | 11.905 | 27.273 | 16.667 | 42.857 |
| ## 3203 | 166 | 46.386 | 0.000 | 45.455 | 45.455 | 36.145 | 20.000 | 36.364 |
| ## 3204 | 142 | 64.085 | 33.333 | 85.714 | 52.381 | 50.704 | 0.000 | 61.905 |
| ## 3205 | 98 | 46.939 | 66.667 | 60.000 | -6.667 | 37.755 | 0.000 | 40.000 |

| | | | | | | | | |
|---------|-----|---------|---------|---------|---------|---------|---------|---------|
| ## 3206 | 50 | 42.000 | 50.000 | 75.000 | 25.000 | 28.000 | 50.000 | 0.000 |
| ## 3207 | 91 | 43.956 | NaN | 47.059 | NaN | 40.659 | NaN | 52.941 |
| ## 3300 | 183 | 38.251 | 50.000 | 33.333 | -16.667 | 38.251 | 50.000 | 33.333 |
| ## 3301 | 154 | 43.506 | 0.000 | 55.000 | 55.000 | 37.662 | 0.000 | 55.000 |
| ## 3302 | 90 | 32.222 | 0.000 | 57.143 | 57.143 | 34.444 | 0.000 | 57.143 |
| ## 3303 | 216 | 38.426 | 14.286 | 31.818 | 17.532 | 32.407 | 28.571 | 31.818 |
| ## 3304 | 762 | 55.381 | 25.000 | 64.602 | 39.602 | 49.606 | 0.000 | 53.097 |
| ## 3305 | 530 | 31.887 | 22.222 | 36.364 | 14.141 | 30.377 | 22.222 | 34.545 |
| ## 3306 | 186 | 54.839 | 0.000 | 58.824 | 58.824 | 52.688 | 0.000 | 61.765 |
| ## 3307 | 25 | 24.000 | NaN | 33.333 | NaN | 8.000 | NaN | 33.333 |
| ## 3308 | 332 | 39.759 | 50.000 | 41.667 | -8.333 | 33.735 | 0.000 | 33.333 |
| ## 3309 | 156 | 42.949 | 50.000 | 44.444 | -5.556 | 41.026 | 50.000 | 50.000 |
| ## 3310 | 312 | 42.308 | 0.000 | 41.935 | 41.935 | 40.705 | 0.000 | 38.710 |
| ## 3311 | 18 | 38.889 | NaN | 33.333 | NaN | 33.333 | NaN | 33.333 |
| ## 3312 | 809 | 34.240 | 7.143 | 39.535 | 32.392 | 33.498 | 14.286 | 40.698 |
| ## 3313 | 56 | 33.929 | NaN | 50.000 | NaN | 21.429 | NaN | 16.667 |
| ## 3314 | 62 | 51.613 | 50.000 | 57.143 | 7.143 | 53.226 | 50.000 | 57.143 |
| ## 3315 | 98 | 40.816 | 100.000 | 38.889 | -61.111 | 45.918 | 100.000 | 55.556 |
| ## 3316 | 248 | 38.710 | 0.000 | 50.000 | 50.000 | 37.500 | 0.000 | 50.000 |
| ## 3317 | 47 | 40.426 | NaN | 28.571 | NaN | 48.936 | NaN | 28.571 |
| ## 3318 | 60 | 75.000 | NaN | 77.778 | NaN | 70.000 | NaN | 66.667 |
| ## 3319 | 17 | 70.588 | NaN | 75.000 | NaN | 64.706 | NaN | 50.000 |
| ## 3320 | 331 | 29.305 | 25.000 | 21.277 | -3.723 | 30.211 | 25.000 | 23.404 |
| ## 3321 | 89 | 31.461 | 50.000 | 25.000 | -25.000 | 33.708 | 50.000 | 37.500 |
| ## 3322 | 48 | 41.667 | NaN | 33.333 | NaN | 35.417 | NaN | 33.333 |
| ## 3400 | 161 | 39.752 | 30.769 | 53.333 | 22.564 | 28.571 | 30.769 | 26.667 |
| ## 3401 | 0 | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| ## 3402 | 1 | 0.000 | NaN | NaN | NaN | 0.000 | NaN | NaN |
| ## 3403 | 14 | 28.571 | NaN | 0.000 | NaN | 28.571 | NaN | 0.000 |
| ## 3404 | 7 | 14.286 | NaN | NaN | NaN | 28.571 | NaN | NaN |
| ## 3500 | 50 | 32.000 | 0.000 | 20.000 | 20.000 | 30.000 | 100.000 | 60.000 |
| ## 3501 | 0 | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| ## 3502 | 0 | NA | NA | NA | NA | NA | NA | NA |
| ## 3503 | 0 | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| ## 3504 | 25 | 20.000 | NaN | 0.000 | NaN | 8.000 | NaN | 0.000 |
| ## 3505 | 4 | 50.000 | NaN | 100.000 | NaN | 0.000 | NaN | 0.000 |
| ## 3506 | 1 | 100.000 | NaN | NaN | NaN | 100.000 | NaN | NaN |
| ## 3600 | 16 | 62.500 | NaN | NaN | NaN | 75.000 | NaN | NaN |
| ## 3601 | 25 | 80.000 | NaN | 100.000 | NaN | 64.000 | NaN | 85.714 |
| ## 3602 | 0 | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| ## 3603 | 2 | 50.000 | NaN | 100.000 | NaN | 100.000 | NaN | 100.000 |
| ## 3604 | 0 | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| ## 3605 | 5 | 40.000 | NaN | NaN | NaN | 80.000 | NaN | NaN |
| ## 3606 | 0 | NA | NA | NA | NA | NA | NA | NA |
| ## 3607 | 17 | 41.176 | NaN | NaN | NaN | 70.588 | NaN | NaN |
| ## 3608 | 0 | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| ## 3609 | 23 | 73.913 | NaN | 66.667 | NaN | 60.870 | NaN | 0.000 |
| ## 3610 | 5 | 20.000 | NaN | NaN | NaN | 20.000 | NaN | NaN |
| ## 3611 | 7 | 71.429 | NaN | 100.000 | NaN | 42.857 | NaN | 100.000 |
| ## 3612 | 145 | 50.345 | 0.000 | 33.333 | 33.333 | 42.069 | 0.000 | 38.095 |

| | | | | | | | | |
|---------|-----------|----------|----------|---------|-----|--------|-----|--------|
| ## 3613 | 7 | 14.286 | NaN | 0.000 | NaN | 0.000 | NaN | 0.000 |
| ## 3614 | 61 | 37.705 | NaN | 50.000 | NaN | 31.148 | NaN | 0.000 |
| ## 3615 | 0 | NA | NA | NA | NA | NA | NA | NA |
| ## 3616 | 38 | 73.684 | NaN | 100.000 | NaN | 55.263 | NaN | 33.333 |
| ## | ChgN | 1CtyMF96 | 1CtyMF14 | | | | | |
| ## 1000 | -50.0000 | 2 | 3 | | | | | |
| ## 1100 | 33.3333 | 4 | 45 | | | | | |
| ## 1101 | -50.0000 | 2 | 1 | | | | | |
| ## 1102 | -3.3333 | 3 | 10 | | | | | |
| ## 1103 | -1.6807 | 7 | 17 | | | | | |
| ## 1104 | -5.7692 | 12 | 26 | | | | | |
| ## 1105 | -12.6437 | 12 | 29 | | | | | |
| ## 1106 | 23.1355 | 9 | 73 | | | | | |
| ## 1107 | -100.0000 | 1 | 5 | | | | | |
| ## 1108 | NaN | 1 | 0 | | | | | |
| ## 1109 | NaN | 0 | 5 | | | | | |
| ## 1110 | -10.2564 | 6 | 13 | | | | | |
| ## 1111 | -50.0000 | 2 | 6 | | | | | |
| ## 1200 | NaN | 0 | 8 | | | | | |
| ## 1201 | NaN | 0 | 23 | | | | | |
| ## 1202 | 41.6667 | 4 | 48 | | | | | |
| ## 1203 | 10.4839 | 4 | 31 | | | | | |
| ## 1204 | 37.5000 | 2 | 8 | | | | | |
| ## 1205 | 40.0000 | 1 | 5 | | | | | |
| ## 1206 | NaN | 0 | 0 | | | | | |
| ## 1207 | 25.0000 | 7 | 28 | | | | | |
| ## 1208 | 48.7805 | 2 | 41 | | | | | |
| ## 1209 | NaN | 0 | 0 | | | | | |
| ## 1210 | 87.5000 | 1 | 8 | | | | | |
| ## 1211 | 13.4921 | 9 | 28 | | | | | |
| ## 1212 | NaN | 0 | 15 | | | | | |
| ## 1213 | NaN | 0 | 16 | | | | | |
| ## 1300 | -71.7391 | 1 | 46 | | | | | |
| ## 1301 | NaN | 0 | 1 | | | | | |
| ## 1302 | 66.6667 | 1 | 6 | | | | | |
| ## 1303 | 10.6798 | 48 | 95 | | | | | |
| ## 1304 | -36.5546 | 7 | 34 | | | | | |
| ## 1305 | 17.3913 | 10 | 46 | | | | | |
| ## 1306 | 29.4118 | 1 | 17 | | | | | |
| ## 1307 | 14.0762 | 11 | 31 | | | | | |
| ## 1308 | 22.2222 | 7 | 18 | | | | | |
| ## 1309 | NaN | 0 | 8 | | | | | |
| ## 1310 | NaN | 0 | 10 | | | | | |
| ## 1311 | 14.8148 | 9 | 27 | | | | | |
| ## 1312 | -0.7937 | 14 | 63 | | | | | |
| ## 1313 | -28.5714 | 3 | 21 | | | | | |
| ## 1314 | 17.6471 | 2 | 17 | | | | | |
| ## 1315 | -75.0000 | 2 | 4 | | | | | |
| ## 1400 | 16.6667 | 2 | 18 | | | | | |
| ## 1401 | NaN | 0 | 3 | | | | | |

| | | | |
|---------|-----------|----|----|
| ## 1402 | -52.3810 | 3 | 7 |
| ## 1403 | NaN | 0 | 17 |
| ## 1404 | 0.0000 | 1 | 1 |
| ## 1405 | -12.5000 | 2 | 8 |
| ## 1406 | NaN | 0 | 8 |
| ## 1407 | -5.1282 | 3 | 13 |
| ## 1408 | 16.1491 | 7 | 23 |
| ## 1409 | NaN | 0 | 6 |
| ## 1410 | -50.0000 | 1 | 2 |
| ## 1500 | 6.5476 | 7 | 24 |
| ## 1501 | NaN | 0 | 0 |
| ## 1502 | 8.1081 | 4 | 37 |
| ## 1503 | -1.7857 | 7 | 24 |
| ## 1504 | NaN | 0 | 1 |
| ## 1505 | NaN | 0 | 6 |
| ## 1506 | NaN | 0 | 1 |
| ## 1507 | NaN | 0 | 4 |
| ## 1508 | 25.0000 | 4 | 8 |
| ## 1600 | -14.9425 | 18 | 87 |
| ## 1601 | NaN | 0 | 1 |
| ## 1602 | 11.3971 | 16 | 34 |
| ## 1603 | 12.5000 | 6 | 16 |
| ## 1604 | 13.3333 | 6 | 15 |
| ## 1605 | 7.5000 | 15 | 48 |
| ## 1606 | 17.7778 | 13 | 45 |
| ## 1607 | 5.5556 | 6 | 18 |
| ## 1700 | NaN | 0 | 10 |
| ## 1701 | NaN | 0 | 2 |
| ## 1702 | 16.6667 | 3 | 6 |
| ## 1703 | NaN | 0 | 4 |
| ## 1704 | NaN | 0 | 2 |
| ## 1705 | -50.0000 | 2 | 14 |
| ## 1706 | 21.2121 | 5 | 33 |
| ## 1707 | NaN | 0 | 5 |
| ## 1708 | 14.2857 | 3 | 7 |
| ## 1709 | -100.0000 | 1 | 5 |
| ## 1710 | -15.6863 | 3 | 17 |
| ## 1711 | 33.3333 | 1 | 3 |
| ## 1712 | 8.3333 | 6 | 24 |
| ## 1800 | NaN | 0 | 3 |
| ## 1801 | NaN | 0 | 0 |
| ## 1802 | 0.0000 | 2 | 2 |
| ## 1803 | 0.0000 | 5 | 4 |
| ## 1804 | 0.0000 | 2 | 4 |
| ## 1900 | -2.2222 | 9 | 10 |
| ## 1901 | NaN | 0 | 1 |
| ## 1902 | -100.0000 | 1 | 6 |
| ## 1903 | NaN | 0 | 4 |
| ## 1904 | NaN | 0 | 6 |
| ## 1905 | NaN | 0 | 0 |

| | | | |
|---------|----------|----|----|
| ## 1906 | NaN | 0 | 2 |
| ## 1907 | 0.0000 | 5 | 2 |
| ## 1908 | 0.0000 | 2 | 4 |
| ## 1909 | 0.0000 | 1 | 2 |
| ## 1910 | 0.0000 | 3 | 7 |
| ## 1911 | NaN | 0 | 2 |
| ## 1912 | 0.0000 | 1 | 2 |
| ## 1913 | NaN | 0 | 0 |
| ## 2000 | NaN | 0 | 3 |
| ## 2001 | NaN | 0 | 6 |
| ## 2002 | 5.3864 | 14 | 61 |
| ## 2003 | 15.3846 | 5 | 13 |
| ## 2100 | NaN | 0 | 12 |
| ## 2101 | NaN | 0 | 0 |
| ## 2102 | NaN | 0 | 4 |
| ## 2103 | NaN | 1 | 0 |
| ## 2104 | NaN | 1 | 0 |
| ## 2105 | -50.0000 | 2 | 7 |
| ## 2200 | 12.5000 | 7 | 16 |
| ## 2201 | NaN | 0 | 4 |
| ## 2202 | 0.0000 | 2 | 4 |
| ## 2203 | NaN | 0 | 2 |
| ## 2204 | 9.5238 | 2 | 42 |
| ## 2205 | NaN | 0 | 11 |
| ## 2206 | NaN | 0 | 0 |
| ## 2207 | 6.2500 | 2 | 16 |
| ## 2208 | 4.3478 | 6 | 69 |
| ## 2209 | NaN | 0 | 21 |
| ## 2210 | 14.2857 | 3 | 35 |
| ## 2211 | 0.0000 | 2 | 17 |
| ## 2212 | NaN | 0 | 4 |
| ## 2213 | 28.5714 | 3 | 14 |
| ## 2214 | NaN | 0 | 3 |
| ## 2215 | NaN | 0 | 8 |
| ## 2216 | NaN | 0 | 0 |
| ## 2300 | -0.7692 | 5 | 26 |
| ## 2301 | NaN | 0 | 5 |
| ## 2302 | NaN | 0 | 7 |
| ## 2303 | 22.7273 | 9 | 22 |
| ## 2304 | -35.0000 | 2 | 20 |
| ## 2305 | NaN | 0 | 11 |
| ## 2306 | 0.0000 | 1 | 2 |
| ## 2307 | 33.3333 | 2 | 3 |
| ## 2308 | -31.4815 | 2 | 27 |
| ## 2309 | 22.2222 | 2 | 9 |
| ## 2310 | 16.6667 | 1 | 12 |
| ## 2311 | 14.2857 | 1 | 7 |
| ## 2312 | 30.7692 | 1 | 13 |
| ## 2400 | NaN | 0 | 5 |
| ## 2401 | NaN | 0 | 0 |

| | | | |
|---------|----------|----|-----|
| ## 2402 | 3.4091 | 11 | 24 |
| ## 2403 | 12.3656 | 12 | 31 |
| ## 2404 | 25.0000 | 3 | 36 |
| ## 2405 | 33.3333 | 1 | 6 |
| ## 2406 | NaN | 0 | 11 |
| ## 2500 | 13.9535 | 2 | 43 |
| ## 2501 | NaN | 0 | 0 |
| ## 2502 | NaN | 0 | 25 |
| ## 2503 | -40.9091 | 2 | 22 |
| ## 2504 | 6.6667 | 3 | 60 |
| ## 2505 | -1.6667 | 10 | 48 |
| ## 2506 | -41.6667 | 2 | 24 |
| ## 2507 | 0.0000 | 3 | 9 |
| ## 2508 | -4.8485 | 5 | 33 |
| ## 2600 | 6.6667 | 12 | 20 |
| ## 2601 | NaN | 0 | 2 |
| ## 2602 | 5.8824 | 2 | 17 |
| ## 2603 | 0.0000 | 7 | 7 |
| ## 2604 | 7.6923 | 13 | 26 |
| ## 2605 | NaN | 0 | 8 |
| ## 2606 | NaN | 0 | 1 |
| ## 2607 | NaN | 0 | 9 |
| ## 2608 | NaN | 0 | 7 |
| ## 2609 | NaN | 0 | 2 |
| ## 2610 | 0.0000 | 3 | 3 |
| ## 2611 | 17.3913 | 1 | 23 |
| ## 2612 | NaN | 0 | 5 |
| ## 2613 | 0.0000 | 5 | 15 |
| ## 2614 | 0.0000 | 1 | 8 |
| ## 2700 | 23.4127 | 14 | 108 |
| ## 2701 | 8.5714 | 5 | 28 |
| ## 2702 | 50.0000 | 1 | 2 |
| ## 2703 | -12.3077 | 5 | 13 |
| ## 2704 | 50.0000 | 2 | 4 |
| ## 2705 | 14.5299 | 13 | 18 |
| ## 2706 | 12.5000 | 1 | 8 |
| ## 2707 | NaN | 0 | 5 |
| ## 2708 | -30.0000 | 2 | 5 |
| ## 2709 | NaN | 0 | 1 |
| ## 2710 | NaN | 0 | 2 |
| ## 2711 | NaN | 0 | 12 |
| ## 2712 | NaN | 0 | 16 |
| ## 2713 | -10.0000 | 5 | 4 |
| ## 2714 | 50.0000 | 2 | 2 |
| ## 2715 | 16.0000 | 3 | 25 |
| ## 2716 | 25.0000 | 3 | 12 |
| ## 2717 | 40.9091 | 3 | 22 |
| ## 2718 | NaN | 0 | 6 |
| ## 2719 | NaN | 0 | 23 |
| ## 2720 | 37.5000 | 1 | 8 |

| | | | |
|---------|----------|----|----|
| ## 2721 | NaN | 0 | 8 |
| ## 2722 | 0.0000 | 2 | 2 |
| ## 2723 | 26.3158 | 5 | 19 |
| ## 2724 | 16.6667 | 6 | 6 |
| ## 2725 | -16.6667 | 2 | 24 |
| ## 2726 | NaN | 0 | 23 |
| ## 2727 | NaN | 0 | 9 |
| ## 2728 | NaN | 0 | 25 |
| ## 2729 | 22.2222 | 6 | 18 |
| ## 2730 | -5.7471 | 3 | 29 |
| ## 2731 | 5.0000 | 5 | 12 |
| ## 2732 | 16.6667 | 5 | 48 |
| ## 2733 | NaN | 0 | 5 |
| ## 2734 | 27.2727 | 2 | 11 |
| ## 2735 | 21.8750 | 4 | 32 |
| ## 2736 | 25.0000 | 1 | 16 |
| ## 2737 | 25.0000 | 1 | 12 |
| ## 2738 | 16.8768 | 17 | 84 |
| ## 2739 | -53.6585 | 3 | 41 |
| ## 2740 | 7.6923 | 1 | 13 |
| ## 2741 | -46.5517 | 2 | 29 |
| ## 2742 | 50.0000 | 1 | 14 |
| ## 2743 | 16.6667 | 3 | 12 |
| ## 2744 | NaN | 2 | 0 |
| ## 2745 | NaN | 0 | 11 |
| ## 2746 | 5.3571 | 7 | 56 |
| ## 2747 | NaN | 0 | 4 |
| ## 2748 | 35.7143 | 1 | 14 |
| ## 2800 | 22.2222 | 2 | 18 |
| ## 2801 | NaN | 0 | 3 |
| ## 2802 | 0.0000 | 2 | 5 |
| ## 2803 | 7.1429 | 4 | 14 |
| ## 2804 | 33.3333 | 6 | 16 |
| ## 2805 | 3.5714 | 4 | 7 |
| ## 2806 | NaN | 0 | 1 |
| ## 2807 | NaN | 0 | 8 |
| ## 2808 | NaN | 0 | 16 |
| ## 2809 | 30.0000 | 5 | 6 |
| ## 2900 | 3.9474 | 4 | 38 |
| ## 2901 | NaN | 0 | 8 |
| ## 2902 | 57.1429 | 2 | 7 |
| ## 2903 | NaN | 2 | 0 |
| ## 2904 | NaN | 0 | 2 |
| ## 2905 | NaN | 0 | 10 |
| ## 2906 | NaN | 0 | 0 |
| ## 2907 | NaN | 0 | 3 |
| ## 2908 | NaN | 0 | 0 |
| ## 2909 | NaN | 0 | 14 |
| ## 2910 | NaN | 0 | 5 |
| ## 2911 | NaN | 0 | 3 |

| | | | |
|---------|----------|----|-----|
| ## 2912 | NaN | 0 | 1 |
| ## 2913 | NaN | 0 | 5 |
| ## 2914 | NaN | 0 | 5 |
| ## 2915 | NA | NA | NA |
| ## 2916 | -41.6667 | 1 | 12 |
| ## 2917 | NaN | 0 | 6 |
| ## 2918 | NA | NA | NA |
| ## 2919 | NaN | 0 | 7 |
| ## 2920 | NaN | 0 | 0 |
| ## 2921 | NaN | 0 | 9 |
| ## 2922 | NaN | 0 | 1 |
| ## 2923 | NaN | 0 | 3 |
| ## 3000 | NaN | 0 | 2 |
| ## 3001 | NaN | 0 | 1 |
| ## 3002 | 10.4348 | 10 | 23 |
| ## 3003 | 22.4638 | 6 | 23 |
| ## 3004 | 7.3529 | 12 | 34 |
| ## 3005 | 60.0000 | 3 | 5 |
| ## 3100 | NaN | 0 | 27 |
| ## 3101 | 0.0000 | 3 | 3 |
| ## 3102 | NaN | 0 | 2 |
| ## 3103 | 0.0000 | 1 | 2 |
| ## 3104 | 20.3390 | 9 | 59 |
| ## 3105 | 15.3846 | 1 | 13 |
| ## 3106 | 0.0000 | 1 | 7 |
| ## 3107 | 5.0000 | 2 | 40 |
| ## 3108 | NaN | 0 | 5 |
| ## 3109 | 0.0000 | 3 | 4 |
| ## 3110 | NaN | 0 | 4 |
| ## 3200 | 11.4286 | 14 | 30 |
| ## 3201 | NaN | 0 | 0 |
| ## 3202 | 26.1905 | 12 | 28 |
| ## 3203 | 16.3636 | 5 | 22 |
| ## 3204 | 61.9048 | 3 | 21 |
| ## 3205 | 40.0000 | 3 | 10 |
| ## 3206 | -50.0000 | 2 | 4 |
| ## 3207 | NaN | 0 | 17 |
| ## 3300 | -16.6667 | 4 | 21 |
| ## 3301 | 55.0000 | 2 | 20 |
| ## 3302 | 57.1429 | 2 | 7 |
| ## 3303 | 3.2468 | 7 | 22 |
| ## 3304 | 53.0973 | 4 | 113 |
| ## 3305 | 12.3232 | 9 | 55 |
| ## 3306 | 61.7647 | 1 | 34 |
| ## 3307 | NaN | 0 | 3 |
| ## 3308 | 33.3333 | 2 | 36 |
| ## 3309 | 0.0000 | 2 | 18 |
| ## 3310 | 38.7097 | 4 | 31 |
| ## 3311 | NaN | 0 | 3 |
| ## 3312 | 26.4120 | 14 | 86 |

```

## 3313      NaN      0      6
## 3314      7.1429      2      7
## 3315     -44.4444      1     18
## 3316     50.0000      2     36
## 3317      NaN      0      7
## 3318      NaN      0      9
## 3319      NaN      0      4
## 3320     -1.5957      4     47
## 3321    -12.5000      2      8
## 3322      NaN      0      6
## 3400     -4.1026     13     15
## 3401      NaN      0      0
## 3402      NaN      0      0
## 3403      NaN      0      2
## 3404      NaN      0      0
## 3500    -40.0000      1      5
## 3501      NaN      0      0
## 3502       NA      NA      NA
## 3503      NaN      0      0
## 3504      NaN      0      1
## 3505      NaN      0      1
## 3506      NaN      0      0
## 3600      NaN      0      0
## 3601      NaN      0      7
## 3602      NaN      0      0
## 3603      NaN      0      1
## 3604      NaN      0      0
## 3605      NaN      0      0
## 3606       NA      NA      NA
## 3607      NaN      0      0
## 3608      NaN      0      0
## 3609      NaN      0      3
## 3610      NaN      0      0
## 3611      NaN      0      1
## 3612     38.0952      1     21
## 3613      NaN      0      2
## 3614      NaN      0      6
## 3615       NA      NA      NA
## 3616      NaN      0      3

```

```
print("GenderTeamSize in 2018")
```

```
## [1] "GenderTeamSize in 2018"
```

```
print(GenderTeamSize)
```

```

##      Articles FirstF FirstM FirstP LastF LastM LastP
## 1000         23  3.752  3.374 0.75576 3.621 3.496 0.802009
## 1100         27  5.139  3.232 0.16940 5.985 3.774 0.283207
## 1101          NA     NA     NA     NA     NA     NA     NA
## 1102         12  4.527  3.177 0.25232 5.192 3.156 0.068548

```

| | | | | | | | |
|---------|----|-------|-------|---------|-------|-------|----------|
| ## 1103 | 25 | 4.379 | 3.493 | 0.40197 | 4.143 | 3.707 | 0.707230 |
| ## 1104 | 19 | 2.034 | 3.456 | 0.14338 | 2.844 | 2.842 | 0.739410 |
| ## 1105 | 19 | 2.380 | 2.785 | 0.82550 | 2.686 | 2.595 | 0.929726 |
| ## 1106 | 77 | 4.353 | 3.863 | 0.37831 | 4.072 | 4.152 | 0.729833 |
| ## 1107 | 8 | 3.634 | 3.178 | 0.28992 | 4.000 | 3.147 | 0.075927 |
| ## 1108 | NA | NA | NA | NA | NA | NA | NA |
| ## 1109 | NA | NA | NA | NA | NA | NA | NA |
| ## 1110 | 16 | 3.000 | 2.958 | 0.93374 | 4.160 | 2.740 | 0.204782 |
| ## 1111 | 4 | 3.873 | 2.449 | 0.41422 | NA | NA | NA |
| ## 1200 | 8 | 1.000 | 1.000 | NaN | 1.000 | 1.000 | NaN |
| ## 1201 | 28 | 2.216 | 1.431 | 0.09935 | 2.203 | 1.446 | 0.156694 |
| ## 1202 | 63 | 1.028 | 1.037 | 0.83409 | 1.000 | 1.055 | 0.173195 |
| ## 1203 | 27 | 1.611 | 1.530 | 0.82458 | 1.691 | 1.364 | 0.568693 |
| ## 1204 | 14 | 1.260 | 1.000 | 0.31232 | 1.000 | 1.167 | 0.550985 |
| ## 1205 | 7 | 1.000 | 1.000 | NaN | 1.000 | 1.000 | NaN |
| ## 1206 | NA | NA | NA | NA | NA | NA | NA |
| ## 1207 | 20 | 1.122 | 1.000 | 0.15607 | 1.000 | 1.047 | 0.644167 |
| ## 1208 | 40 | 1.000 | 1.032 | 0.39295 | 1.000 | 1.032 | 0.392949 |
| ## 1209 | NA | NA | NA | NA | NA | NA | NA |
| ## 1210 | 7 | 1.260 | 1.732 | 0.55411 | 1.260 | 1.732 | 0.554113 |
| ## 1211 | 42 | 1.072 | 1.022 | 0.40068 | 1.072 | 1.022 | 0.400679 |
| ## 1212 | 20 | 1.000 | 1.203 | 0.32520 | 1.122 | 1.160 | 1.000000 |
| ## 1213 | 15 | 1.189 | 1.170 | 1.00000 | 1.147 | 1.219 | 0.922085 |
| ## 1300 | 41 | 4.642 | 4.085 | 0.43616 | 3.194 | 4.723 | 0.161747 |
| ## 1301 | NA | NA | NA | NA | NA | NA | NA |
| ## 1302 | 8 | 5.875 | 5.241 | 0.53864 | 6.300 | 5.030 | 0.297483 |
| ## 1303 | 37 | 4.614 | 3.220 | 0.01707 | 5.157 | 3.271 | 0.010907 |
| ## 1304 | 14 | 3.778 | 4.107 | 0.55377 | 4.338 | 3.699 | 0.496833 |
| ## 1305 | 32 | 4.245 | 3.022 | 0.19960 | 4.110 | 3.530 | 0.553841 |
| ## 1306 | 17 | 7.710 | 6.093 | 0.33951 | 8.180 | 6.576 | 0.350604 |
| ## 1307 | 30 | 5.533 | 4.515 | 0.12937 | 5.085 | 4.935 | 0.895495 |
| ## 1308 | 9 | 6.868 | 3.689 | 0.14863 | 7.000 | 4.299 | 0.324723 |
| ## 1309 | 8 | 9.000 | 4.409 | 0.26946 | 5.646 | 4.384 | 0.762838 |
| ## 1310 | 5 | 6.694 | 5.477 | 0.76091 | 6.000 | 6.223 | 1.000000 |
| ## 1311 | 39 | 4.088 | 4.951 | 0.55064 | 3.921 | 4.914 | 0.574702 |
| ## 1312 | 44 | 4.950 | 4.292 | 0.23850 | 5.566 | 4.180 | 0.075944 |
| ## 1313 | 17 | 5.507 | 5.459 | 0.72615 | 5.334 | 5.559 | 0.570824 |
| ## 1314 | 25 | 2.289 | 3.182 | 0.49859 | 2.491 | 3.220 | 0.491952 |
| ## 1315 | NA | NA | NA | NA | NA | NA | NA |
| ## 1400 | 9 | 1.732 | 2.034 | 0.87903 | 1.732 | 2.034 | 0.879026 |
| ## 1401 | 9 | 2.621 | 2.040 | 0.78549 | 2.913 | 1.783 | 0.196706 |
| ## 1402 | 6 | 3.000 | 1.888 | 0.34278 | 3.000 | 1.888 | 0.342782 |
| ## 1403 | 25 | 2.000 | 1.726 | 0.60720 | 1.682 | 1.797 | 0.779133 |
| ## 1404 | 4 | 2.000 | 4.000 | 0.19393 | NA | NA | NA |
| ## 1405 | 19 | 1.739 | 1.972 | 0.52553 | 1.861 | 1.898 | 0.964662 |
| ## 1406 | 8 | 2.913 | 1.861 | 0.17203 | 2.289 | 2.352 | 0.753945 |
| ## 1407 | 18 | 1.906 | 1.698 | 0.68846 | 2.107 | 1.414 | 0.070801 |
| ## 1408 | 34 | 1.876 | 1.783 | 1.00000 | 2.190 | 1.523 | 0.045915 |
| ## 1409 | 4 | 3.000 | 2.080 | 1.00000 | 3.000 | 2.080 | 1.000000 |
| ## 1410 | NA | NA | NA | NA | NA | NA | NA |

| | | | | | | | |
|---------|----|-------|-------|---------|--------|-------|----------|
| ## 1500 | 22 | 4.494 | 3.535 | 0.54879 | 4.967 | 3.455 | 0.255209 |
| ## 1501 | 2 | 4.000 | 1.000 | 1.00000 | 4.000 | 1.000 | 1.000000 |
| ## 1502 | 13 | 3.751 | 4.064 | 0.58341 | 11.000 | 3.642 | 0.135525 |
| ## 1503 | 11 | 4.481 | 4.166 | 0.67459 | 5.013 | 3.994 | 1.000000 |
| ## 1504 | NA | NA | NA | NA | NA | NA | NA |
| ## 1505 | 3 | 5.000 | 4.243 | 1.00000 | 3.000 | 5.477 | 0.666667 |
| ## 1506 | NA | NA | NA | NA | NA | NA | NA |
| ## 1507 | NA | NA | NA | NA | NA | NA | NA |
| ## 1508 | 6 | 3.873 | 3.936 | 1.00000 | NA | NA | NA |
| ## 1600 | 54 | 4.293 | 3.463 | 0.39452 | 5.230 | 3.341 | 0.027343 |
| ## 1601 | NA | NA | NA | NA | NA | NA | NA |
| ## 1602 | 12 | 5.261 | 3.040 | 0.16298 | 6.160 | 3.009 | 0.059003 |
| ## 1603 | 11 | 4.441 | 2.714 | 0.16616 | 3.464 | 3.982 | 1.000000 |
| ## 1604 | 10 | 3.641 | 6.044 | 0.33978 | 4.820 | 4.637 | 0.728454 |
| ## 1605 | 19 | 4.679 | 3.315 | 0.08469 | 4.380 | 3.580 | 0.524644 |
| ## 1606 | 22 | 4.618 | 3.723 | 0.26222 | 5.492 | 3.442 | 0.070523 |
| ## 1607 | 9 | 3.936 | 2.569 | 0.30614 | 5.518 | 2.330 | 0.177530 |
| ## 1700 | 9 | 2.449 | 1.516 | 0.07810 | 2.449 | 1.739 | 0.331096 |
| ## 1701 | NA | NA | NA | NA | NA | NA | NA |
| ## 1702 | NA | NA | NA | NA | NA | NA | NA |
| ## 1703 | NA | NA | NA | NA | NA | NA | NA |
| ## 1704 | 4 | 3.000 | 4.160 | 0.63735 | 3.000 | 4.160 | 0.637352 |
| ## 1705 | 12 | 2.060 | 3.342 | 0.11093 | 2.449 | 3.064 | 0.657924 |
| ## 1706 | 17 | 2.060 | 2.390 | 0.59583 | 2.280 | 2.316 | 1.000000 |
| ## 1707 | 5 | 4.243 | 2.289 | 0.22354 | 3.780 | 2.000 | 0.128147 |
| ## 1708 | 6 | 3.000 | 2.352 | 1.00000 | 3.000 | 2.352 | 1.000000 |
| ## 1709 | 8 | 2.000 | 1.906 | 1.00000 | 1.000 | 2.119 | 0.241835 |
| ## 1710 | 19 | 2.048 | 3.071 | 0.11213 | 2.402 | 2.901 | 0.630214 |
| ## 1711 | 8 | 3.000 | 3.504 | 1.00000 | 3.873 | 3.302 | 0.862684 |
| ## 1712 | 16 | 2.884 | 2.245 | 0.29591 | 3.464 | 2.227 | 0.094100 |
| ## 1800 | NA | NA | NA | NA | NA | NA | NA |
| ## 1801 | NA | NA | NA | NA | NA | NA | NA |
| ## 1802 | 2 | 3.000 | 3.000 | NaN | NA | NA | NA |
| ## 1803 | 7 | 1.000 | 1.414 | 0.75519 | 1.000 | 1.414 | 0.755189 |
| ## 1804 | NA | NA | NA | NA | NA | NA | NA |
| ## 1900 | 11 | 1.189 | 2.318 | 0.05907 | 1.431 | 2.221 | 0.249405 |
| ## 1901 | 2 | 2.000 | 2.000 | NaN | 2.000 | 2.000 | NaN |
| ## 1902 | 6 | 3.000 | 3.568 | 1.00000 | 2.621 | 4.327 | 0.642835 |
| ## 1903 | NA | NA | NA | NA | NA | NA | NA |
| ## 1904 | 9 | 1.906 | 1.000 | 0.11794 | 1.565 | 1.516 | 1.000000 |
| ## 1905 | NA | NA | NA | NA | NA | NA | NA |
| ## 1906 | NA | NA | NA | NA | NA | NA | NA |
| ## 1907 | NA | NA | NA | NA | NA | NA | NA |
| ## 1908 | 3 | 4.000 | 3.162 | 1.00000 | NA | NA | NA |
| ## 1909 | NA | NA | NA | NA | NA | NA | NA |
| ## 1910 | 2 | 1.000 | 3.000 | 1.00000 | 1.000 | 3.000 | 1.000000 |
| ## 1911 | NA | NA | NA | NA | NA | NA | NA |
| ## 1912 | 4 | 3.000 | 2.621 | 1.00000 | 3.000 | 2.449 | 0.617075 |
| ## 1913 | NA | NA | NA | NA | NA | NA | NA |
| ## 2000 | 10 | 2.000 | 1.622 | 0.67608 | 2.000 | 1.661 | 0.852684 |

| | | | | | | | |
|---------|----|-------|-------|---------|--------|-------|----------|
| ## 2001 | 7 | 1.000 | 2.169 | 0.05187 | 1.260 | 2.213 | 0.114653 |
| ## 2002 | 54 | 1.769 | 1.705 | 0.85283 | 1.631 | 1.753 | 0.634741 |
| ## 2003 | 17 | 2.000 | 1.783 | 0.69176 | 1.732 | 1.817 | 1.000000 |
| ## 2100 | 17 | 2.491 | 3.033 | 0.43215 | 3.061 | 2.731 | 0.716139 |
| ## 2101 | 3 | 3.000 | 3.873 | 1.00000 | NA | NA | NA |
| ## 2102 | 12 | 6.000 | 2.451 | 0.12134 | 6.000 | 2.451 | 0.121335 |
| ## 2103 | 2 | 6.000 | 4.000 | 1.00000 | 6.000 | 4.000 | 1.000000 |
| ## 2104 | NA | NA | NA | NA | NA | NA | NA |
| ## 2105 | 34 | 3.026 | 2.936 | 0.92488 | 3.780 | 2.800 | 0.333826 |
| ## 2200 | 10 | 3.162 | 1.707 | 0.27332 | 5.000 | 1.737 | 0.144127 |
| ## 2201 | NA | NA | NA | NA | NA | NA | NA |
| ## 2202 | 6 | 3.000 | 2.352 | 0.76302 | NA | NA | NA |
| ## 2203 | 5 | 1.817 | 4.899 | 0.20000 | 1.414 | 4.160 | 0.200000 |
| ## 2204 | 32 | 3.365 | 3.039 | 0.83121 | 3.963 | 2.987 | 0.113836 |
| ## 2205 | 18 | 2.048 | 2.378 | 0.44098 | 2.060 | 2.349 | 0.618417 |
| ## 2206 | NA | NA | NA | NA | NA | NA | NA |
| ## 2207 | NA | NA | NA | NA | NA | NA | NA |
| ## 2208 | 17 | 6.000 | 3.477 | 0.29899 | 4.000 | 3.567 | 0.917281 |
| ## 2209 | 24 | 3.465 | 2.937 | 0.46379 | 2.913 | 3.169 | 0.500232 |
| ## 2210 | 24 | 3.208 | 3.454 | 0.82403 | 3.317 | 3.432 | 1.000000 |
| ## 2211 | 18 | 5.745 | 3.128 | 0.34497 | 11.000 | 3.120 | 0.110781 |
| ## 2212 | 9 | 5.000 | 2.632 | 0.15304 | 4.472 | 2.479 | 0.064078 |
| ## 2213 | 9 | 3.360 | 1.260 | 0.03565 | 3.000 | 2.359 | 1.000000 |
| ## 2214 | NA | NA | NA | NA | NA | NA | NA |
| ## 2215 | 9 | 3.000 | 3.035 | 1.00000 | 3.000 | 3.035 | 1.000000 |
| ## 2216 | NA | NA | NA | NA | NA | NA | NA |
| ## 2300 | 15 | 2.696 | 2.395 | 0.84919 | 2.621 | 2.484 | 0.697911 |
| ## 2301 | 9 | 1.000 | 1.609 | 0.64343 | 1.000 | 1.609 | 0.643429 |
| ## 2302 | NA | NA | NA | NA | NA | NA | NA |
| ## 2303 | 13 | 2.769 | 2.221 | 0.44752 | 2.000 | 2.734 | 0.539255 |
| ## 2304 | 7 | 3.464 | 4.043 | 0.84367 | 6.000 | 3.245 | 0.236724 |
| ## 2305 | 8 | 3.000 | 2.493 | 0.86355 | NA | NA | NA |
| ## 2306 | 7 | 1.682 | 3.000 | 0.35434 | 1.000 | 2.930 | 0.067808 |
| ## 2307 | 15 | 3.811 | 2.213 | 0.20276 | 3.420 | 3.193 | 0.439820 |
| ## 2308 | 33 | 2.284 | 2.371 | 0.80334 | 2.559 | 2.206 | 0.473602 |
| ## 2309 | 8 | 2.000 | 2.570 | 0.38715 | 2.828 | 2.289 | 0.729406 |
| ## 2310 | 9 | 3.915 | 3.915 | 1.00000 | 3.915 | 3.915 | 1.000000 |
| ## 2311 | 5 | 3.000 | 3.915 | 0.33292 | NA | NA | NA |
| ## 2312 | 10 | 5.000 | 2.636 | 0.14546 | 4.472 | 2.502 | 0.101481 |
| ## 2400 | 8 | 5.384 | 3.936 | 0.55657 | 5.292 | 4.394 | 0.734271 |
| ## 2401 | NA | NA | NA | NA | NA | NA | NA |
| ## 2402 | 14 | 3.086 | 4.070 | 0.18874 | 3.634 | 3.609 | 1.000000 |
| ## 2403 | 25 | 6.803 | 5.331 | 0.29954 | 6.740 | 5.510 | 0.331184 |
| ## 2404 | 29 | 5.725 | 4.704 | 0.35359 | 5.323 | 4.839 | 0.487792 |
| ## 2405 | 12 | 6.687 | 5.038 | 0.54941 | 6.407 | 4.988 | 0.461753 |
| ## 2406 | 3 | 7.071 | 8.000 | 1.00000 | NA | NA | NA |
| ## 2500 | 47 | 3.552 | 3.443 | 0.92743 | 4.139 | 3.340 | 0.260876 |
| ## 2501 | NA | NA | NA | NA | NA | NA | NA |
| ## 2502 | 21 | 4.558 | 3.460 | 0.26715 | 6.325 | 3.754 | 0.161297 |
| ## 2503 | 7 | 2.449 | 3.728 | 0.15548 | NA | NA | NA |

| | | | | | | | |
|---------|----|-------|--------|---------|--------|-------|----------|
| ## 2504 | 24 | 5.384 | 3.678 | 0.34701 | 5.000 | 3.785 | 0.377179 |
| ## 2505 | 6 | 3.873 | 2.213 | 0.13361 | NA | NA | NA |
| ## 2506 | 7 | 3.464 | 4.338 | 0.51694 | NA | NA | NA |
| ## 2507 | 9 | 3.175 | 4.234 | 0.49822 | 4.000 | 3.828 | 1.000000 |
| ## 2508 | 16 | 3.915 | 4.064 | 0.88804 | 7.000 | 3.890 | 0.173195 |
| ## 2600 | 11 | 2.000 | 1.196 | 0.22587 | 2.000 | 1.130 | 0.050490 |
| ## 2601 | 5 | 2.289 | 1.732 | 1.00000 | 2.000 | 2.060 | 1.000000 |
| ## 2602 | 14 | 1.414 | 1.414 | 1.00000 | 1.414 | 1.414 | 1.000000 |
| ## 2603 | NA | NA | NA | NA | NA | NA | NA |
| ## 2604 | 25 | 2.491 | 1.588 | 0.09170 | 2.884 | 1.621 | 0.070156 |
| ## 2605 | NA | NA | NA | NA | NA | NA | NA |
| ## 2606 | NA | NA | NA | NA | NA | NA | NA |
| ## 2607 | 8 | 2.000 | 1.414 | 0.36131 | 2.000 | 1.486 | 0.632585 |
| ## 2608 | 4 | 2.000 | 1.260 | 0.61708 | 2.000 | 1.260 | 0.617075 |
| ## 2609 | NA | NA | NA | NA | NA | NA | NA |
| ## 2610 | NA | NA | NA | NA | NA | NA | NA |
| ## 2611 | 8 | 4.000 | 2.284 | 0.36808 | 4.000 | 2.284 | 0.368083 |
| ## 2612 | NA | NA | NA | NA | NA | NA | NA |
| ## 2613 | 9 | 2.236 | 2.560 | 1.00000 | 1.000 | 2.783 | 0.237302 |
| ## 2614 | 4 | 2.000 | 2.000 | NaN | NA | NA | NA |
| ## 2700 | 85 | 3.925 | 3.493 | 0.33480 | 3.653 | 3.716 | 0.660324 |
| ## 2701 | 37 | 4.989 | 6.095 | 0.31571 | 5.506 | 5.126 | 0.333911 |
| ## 2702 | 4 | 2.520 | 5.000 | 0.34578 | 5.000 | 2.520 | 0.345779 |
| ## 2703 | 10 | 3.438 | 4.919 | 0.45781 | 2.913 | 5.175 | 0.159341 |
| ## 2704 | NA | NA | NA | NA | NA | NA | NA |
| ## 2705 | 29 | 3.942 | 5.099 | 0.29757 | 4.416 | 4.720 | 0.930369 |
| ## 2706 | 3 | 4.000 | 3.873 | 1.00000 | 3.000 | 4.472 | 0.666667 |
| ## 2707 | 3 | 7.000 | 2.000 | 0.47950 | NA | NA | NA |
| ## 2708 | 10 | 4.589 | 4.000 | 0.85654 | 2.000 | 4.957 | 0.205718 |
| ## 2709 | NA | NA | NA | NA | NA | NA | NA |
| ## 2710 | 2 | 2.000 | 5.000 | 1.00000 | 5.000 | 2.000 | 1.000000 |
| ## 2711 | 4 | 1.000 | 4.718 | 0.50000 | 1.000 | 4.718 | 0.500000 |
| ## 2712 | 14 | 5.243 | 5.502 | 0.94595 | 3.990 | 6.632 | 0.212333 |
| ## 2713 | 7 | 3.728 | 3.162 | 0.68190 | 3.464 | 3.684 | 0.851596 |
| ## 2714 | 9 | 4.043 | 6.160 | 0.17245 | 4.162 | 5.533 | 0.385261 |
| ## 2715 | 14 | 7.694 | 7.571 | 1.00000 | 6.000 | 8.150 | 0.115805 |
| ## 2716 | 8 | 4.229 | 7.454 | 0.88242 | 3.037 | 8.119 | 0.066798 |
| ## 2717 | 20 | 4.843 | 4.796 | 0.84520 | 4.804 | 4.837 | 0.784593 |
| ## 2718 | 8 | 4.738 | 3.464 | 0.45960 | 4.000 | 4.069 | 1.000000 |
| ## 2719 | 35 | 3.596 | 2.857 | 0.42796 | 3.002 | 3.652 | 0.395528 |
| ## 2720 | 4 | 7.094 | 5.000 | 1.00000 | 17.000 | 4.718 | 0.500000 |
| ## 2721 | 6 | 6.928 | 10.440 | 0.63859 | 6.000 | 9.899 | 0.552453 |
| ## 2722 | 4 | 2.884 | 5.000 | 1.00000 | 3.915 | 2.000 | 0.637352 |
| ## 2723 | 15 | 6.012 | 4.762 | 0.77022 | 7.969 | 5.092 | 0.112919 |
| ## 2724 | 7 | 8.143 | 4.681 | 0.07183 | 6.000 | 5.886 | 1.000000 |
| ## 2725 | 27 | 4.742 | 4.283 | 0.58972 | 5.174 | 3.978 | 0.162371 |
| ## 2726 | 22 | 4.790 | 5.046 | 0.84184 | 7.164 | 4.402 | 0.019225 |
| ## 2727 | 12 | 4.229 | 5.403 | 0.44067 | 2.991 | 6.425 | 0.020713 |
| ## 2728 | 37 | 5.617 | 4.649 | 0.40021 | 6.987 | 4.414 | 0.039619 |
| ## 2729 | 20 | 3.790 | 4.160 | 0.82986 | 3.721 | 4.033 | 0.906494 |

| | | | | | | | |
|---------|----|-------|-------|---------|-------|-------|----------|
| ## 2730 | 35 | 5.978 | 5.177 | 0.40914 | 5.274 | 6.285 | 0.340529 |
| ## 2731 | 9 | 3.562 | 3.000 | 0.68447 | 3.568 | 3.438 | 1.000000 |
| ## 2732 | 60 | 3.844 | 3.643 | 0.58131 | 2.982 | 3.899 | 0.129787 |
| ## 2733 | 9 | 4.282 | 3.936 | 0.70982 | 4.919 | 3.310 | 0.619796 |
| ## 2734 | 18 | 3.137 | 4.456 | 0.17668 | 3.868 | 3.543 | 0.613603 |
| ## 2735 | 37 | 4.058 | 3.105 | 0.29199 | 3.324 | 3.911 | 0.230269 |
| ## 2736 | 23 | 4.376 | 4.582 | 0.60279 | 4.707 | 4.287 | 0.546455 |
| ## 2737 | 16 | 5.241 | 3.929 | 0.41259 | 4.095 | 4.170 | 1.000000 |
| ## 2738 | 65 | 4.306 | 3.238 | 0.10975 | 3.821 | 3.766 | 0.889024 |
| ## 2739 | 55 | 2.813 | 2.251 | 0.31637 | 2.719 | 2.557 | 0.944820 |
| ## 2740 | 6 | 3.780 | 5.518 | 1.00000 | 1.000 | 6.188 | 0.234764 |
| ## 2741 | 26 | 4.423 | 4.656 | 0.80654 | 3.954 | 5.010 | 0.192300 |
| ## 2742 | 12 | 4.059 | 3.130 | 0.41084 | 3.813 | 3.634 | 1.000000 |
| ## 2743 | 3 | 3.000 | 6.000 | 0.47950 | 3.000 | 4.243 | 1.000000 |
| ## 2744 | NA | NA | NA | NA | NA | NA | NA |
| ## 2745 | 9 | 5.536 | 4.579 | 0.89553 | 5.826 | 4.743 | 0.900865 |
| ## 2746 | 49 | 4.930 | 4.052 | 0.23159 | 5.496 | 4.133 | 0.055992 |
| ## 2747 | 5 | 7.483 | 6.868 | 0.76710 | NA | NA | NA |
| ## 2748 | 7 | 3.873 | 3.949 | 1.00000 | 2.828 | 4.478 | 0.324127 |
| ## 2800 | 24 | 3.822 | 3.507 | 0.66203 | 3.207 | 4.024 | 0.574432 |
| ## 2801 | 4 | 4.899 | 5.916 | 1.00000 | 7.483 | 3.873 | 0.333333 |
| ## 2802 | 14 | 5.856 | 4.679 | 0.47426 | 6.586 | 4.296 | 0.107481 |
| ## 2803 | 11 | 4.236 | 2.621 | 0.14169 | 4.160 | 3.245 | 0.573230 |
| ## 2804 | 11 | 5.210 | 5.833 | 1.00000 | 6.544 | 4.734 | 0.387672 |
| ## 2805 | 7 | 3.915 | 3.663 | 0.85445 | 6.062 | 2.000 | 0.043597 |
| ## 2806 | NA | NA | NA | NA | NA | NA | NA |
| ## 2807 | 4 | 8.000 | 5.593 | 0.34578 | 7.483 | 5.000 | 0.220671 |
| ## 2808 | 23 | 5.776 | 4.187 | 0.26735 | 7.094 | 4.046 | 0.038577 |
| ## 2809 | 6 | 3.898 | 3.000 | 0.76302 | 4.481 | 3.107 | 0.368688 |
| ## 2900 | 29 | 3.694 | 4.759 | 0.37895 | 3.860 | 4.039 | 0.945415 |
| ## 2901 | 6 | 2.828 | 2.828 | 1.00000 | 2.639 | 4.000 | 0.729034 |
| ## 2902 | 6 | 2.340 | 4.243 | 0.63859 | 2.154 | 3.780 | 0.506555 |
| ## 2903 | NA | NA | NA | NA | NA | NA | NA |
| ## 2904 | NA | NA | NA | NA | NA | NA | NA |
| ## 2905 | 7 | 2.213 | 1.710 | 0.57840 | 2.605 | 1.000 | 0.067808 |
| ## 2906 | NA | NA | NA | NA | NA | NA | NA |
| ## 2907 | 6 | 1.587 | 3.476 | 0.50018 | 1.000 | 2.787 | 0.365712 |
| ## 2908 | NA | NA | NA | NA | NA | NA | NA |
| ## 2909 | 4 | 5.745 | 3.464 | 1.00000 | NA | NA | NA |
| ## 2910 | 5 | 3.310 | 1.000 | 0.46816 | 2.884 | 2.236 | 1.000000 |
| ## 2911 | NA | NA | NA | NA | NA | NA | NA |
| ## 2912 | 4 | 3.302 | 5.000 | 0.34578 | 3.915 | 3.000 | 0.637352 |
| ## 2913 | 14 | 2.686 | 2.000 | 0.61349 | 2.661 | 2.449 | 0.780393 |
| ## 2914 | NA | NA | NA | NA | NA | NA | NA |
| ## 2915 | NA | NA | NA | NA | NA | NA | NA |
| ## 2916 | 25 | 5.394 | 4.559 | 0.49860 | 5.966 | 4.667 | 0.204449 |
| ## 2917 | NA | NA | NA | NA | NA | NA | NA |
| ## 2918 | NA | NA | NA | NA | NA | NA | NA |
| ## 2919 | 11 | 2.449 | 2.268 | 0.84999 | 2.712 | 1.861 | 0.556987 |
| ## 2920 | NA | NA | NA | NA | NA | NA | NA |

| | | | | | | | |
|---------|-----|-------|-------|---------|-------|-------|----------|
| ## 2921 | 16 | 3.979 | 3.689 | 0.95583 | 3.575 | 4.899 | 0.239387 |
| ## 2922 | 3 | 4.000 | 2.236 | 1.00000 | 4.472 | 1.000 | 0.666667 |
| ## 2923 | NA | NA | NA | NA | NA | NA | NA |
| ## 3000 | NA | NA | NA | NA | NA | NA | NA |
| ## 3001 | NA | NA | NA | NA | NA | NA | NA |
| ## 3002 | 16 | 5.985 | 3.582 | 0.18736 | 6.241 | 3.093 | 0.053838 |
| ## 3003 | 29 | 3.447 | 4.344 | 0.51432 | 3.434 | 4.365 | 0.543746 |
| ## 3004 | 17 | 4.885 | 2.712 | 0.04884 | 4.366 | 3.501 | 0.587993 |
| ## 3005 | 8 | 4.789 | 1.260 | 0.04983 | 2.520 | 3.160 | 0.762838 |
| ## 3100 | NA | NA | NA | NA | NA | NA | NA |
| ## 3101 | NA | NA | NA | NA | NA | NA | NA |
| ## 3102 | 5 | 4.472 | 2.289 | 0.23614 | 4.000 | 2.783 | 1.000000 |
| ## 3103 | 6 | 1.732 | 2.213 | 1.00000 | 1.732 | 2.213 | 1.000000 |
| ## 3104 | 53 | 5.016 | 3.349 | 0.01781 | 4.942 | 3.446 | 0.063348 |
| ## 3105 | 10 | 3.288 | 3.170 | 1.00000 | 2.000 | 3.404 | 0.475332 |
| ## 3106 | NA | NA | NA | NA | NA | NA | NA |
| ## 3107 | 40 | 4.530 | 3.597 | 0.43179 | 4.817 | 3.558 | 0.257940 |
| ## 3108 | NA | NA | NA | NA | NA | NA | NA |
| ## 3109 | NA | NA | NA | NA | NA | NA | NA |
| ## 3110 | 5 | 5.477 | 1.587 | 0.13864 | 6.000 | 2.115 | 0.276500 |
| ## 3200 | 16 | 2.405 | 2.759 | 0.41441 | 2.289 | 2.653 | 0.867021 |
| ## 3201 | 3 | 2.000 | 2.000 | 1.00000 | 2.828 | 1.000 | 0.666667 |
| ## 3202 | 25 | 3.105 | 1.878 | 0.05298 | 2.531 | 2.357 | 0.977626 |
| ## 3203 | 29 | 3.067 | 2.206 | 0.24792 | 2.646 | 2.585 | 0.964555 |
| ## 3204 | 47 | 2.796 | 1.866 | 0.02515 | 2.887 | 1.715 | 0.004733 |
| ## 3205 | 18 | 2.395 | 1.793 | 0.35631 | 2.706 | 1.587 | 0.042426 |
| ## 3206 | 4 | 3.000 | 2.884 | 1.00000 | 3.302 | 2.000 | 0.345779 |
| ## 3207 | 17 | 1.737 | 1.454 | 0.49197 | 1.795 | 1.473 | 0.872227 |
| ## 3300 | 19 | 1.625 | 1.260 | 0.50552 | 1.297 | 1.554 | 0.441567 |
| ## 3301 | 27 | 2.192 | 1.454 | 0.10353 | 2.153 | 1.736 | 0.377468 |
| ## 3302 | 16 | 1.091 | 1.091 | 1.00000 | 1.091 | 1.091 | 1.000000 |
| ## 3303 | 19 | 2.000 | 1.565 | 0.36590 | 1.888 | 1.655 | 0.620428 |
| ## 3304 | 157 | 2.034 | 1.618 | 0.02417 | 1.872 | 1.839 | 0.994107 |
| ## 3305 | 63 | 1.648 | 1.836 | 0.42110 | 1.638 | 1.832 | 0.505183 |
| ## 3306 | 36 | 3.143 | 1.865 | 0.01449 | 2.977 | 2.063 | 0.093570 |
| ## 3307 | 6 | 2.449 | 2.000 | 0.40174 | 3.000 | 2.169 | 0.288844 |
| ## 3308 | 35 | 1.193 | 1.308 | 0.61967 | 1.044 | 1.449 | 0.057832 |
| ## 3309 | 13 | 1.513 | 2.000 | 0.34348 | 1.622 | 2.000 | 0.517540 |
| ## 3310 | 32 | 1.471 | 1.534 | 0.80978 | 1.450 | 1.571 | 0.588993 |
| ## 3311 | 2 | 1.000 | 1.000 | NaN | 1.000 | 1.000 | NaN |
| ## 3312 | 117 | 1.539 | 1.291 | 0.06438 | 1.500 | 1.318 | 0.148778 |
| ## 3313 | 7 | 1.414 | 1.644 | 0.83432 | 1.817 | 1.414 | 0.566765 |
| ## 3314 | 17 | 1.080 | 1.000 | 0.40940 | 1.000 | 1.080 | 0.409395 |
| ## 3315 | 23 | 1.466 | 1.282 | 0.58264 | 1.499 | 1.220 | 0.333279 |
| ## 3316 | 79 | 1.052 | 1.068 | 0.60897 | 1.045 | 1.076 | 0.592795 |
| ## 3317 | 3 | 2.000 | 1.000 | 1.00000 | 1.000 | 2.000 | 1.000000 |
| ## 3318 | NA | NA | NA | NA | NA | NA | NA |
| ## 3319 | 6 | 1.682 | 1.000 | 0.41131 | 1.260 | 1.587 | 1.000000 |
| ## 3320 | 45 | 1.335 | 1.167 | 0.28211 | 1.282 | 1.185 | 0.597093 |
| ## 3321 | 15 | 1.587 | 1.737 | 0.70517 | 1.739 | 1.622 | 0.852628 |

```
## 3322      13  1.260  1.835  0.32218  1.260  1.835  0.322183
## 3400       8  4.309  3.728  0.42078  4.107  3.464  0.368066
## 3401      NA      NA      NA      NA      NA      NA      NA
## 3402       2  5.000  5.000      NaN  5.000  5.000      NaN
## 3403       6  5.958  3.162  0.15266  5.000  4.789  0.763025
## 3404       2  5.000  5.000      NaN  5.000  5.000      NaN
## 3500      NA      NA      NA      NA      NA      NA      NA
## 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       4  4.932  1.000  0.50000      NA      NA      NA
## 3505      NA      NA      NA      NA      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      NA      NA      NA      NA      NA      NA      NA
## 3603       2  1.000  2.000  1.00000  1.000  2.000  1.000000
## 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      NA      NA      NA      NA      NA      NA      NA
## 3611      NA      NA      NA      NA      NA      NA      NA
## 3612      41  3.303  3.170  0.70039  2.934  3.437  0.439427
## 3613      NA      NA      NA      NA      NA      NA      NA
## 3614      NA      NA      NA      NA      NA      NA      NA
## 3615      NA      NA      NA      NA      NA      NA      NA
## 3616      10  2.942  5.566  0.06488  3.775  4.000  1.000000
```

```
print(RegCoef)
```

```
##          FFA1      FLA1      2      3      4      5+
## 1000      NA      NA      NA      NA      NA      NA
## 1100  0.007417 -0.0579331  0.397019  0.6807147  0.546034  7.240e-01
## 1101      NA      NA      NA      NA      NA      NA
## 1102  0.009568  0.0501670  0.316490  0.4439671  0.371661  5.314e-01
## 1103 -0.059368 -0.1515358  0.421272  0.5129509  0.452060  6.231e-01
## 1104 -0.084012  0.0092248  0.078715 -0.0222257  0.091246  2.097e-02
## 1105 -0.084684  0.0486111  0.280096  0.3016911  0.255924  2.840e-01
## 1106  0.008363 -0.0169585  0.700839  0.7856010  0.808974  8.769e-01
## 1107  0.018577  0.1072160 -0.253699  0.0082091 -0.040804 -1.378e-01
## 1108      NA      NA      NA      NA      NA      NA
## 1109  0.142315  0.1135551  0.219840  0.0838397  0.368311  2.748e-01
## 1110 -0.037065  0.0488494  0.517101  0.2860209  0.532002  6.898e-01
## 1111 -0.039959 -0.2602808 -0.044168 -0.1055768  0.074968 -1.784e-01
## 1200  0.684182 -1.4945556  1.384008  0.8276357  1.133600  2.576e-15
## 1201  0.048222  0.0045173  0.200257  0.1578955  0.412697  4.880e-01
## 1202 -0.097013  0.1466534  0.547105  0.1440925  0.686128 -3.236e-01
```

| | | | | | | | |
|----|------|-----------|------------|-----------|------------|-----------|------------|
| ## | 1203 | -0.205443 | 0.3050428 | 0.260557 | 0.3664452 | 0.434503 | -9.064e-03 |
| ## | 1204 | -0.171060 | 0.2917311 | 0.425902 | 0.2003803 | -0.352811 | 5.039e-02 |
| ## | 1205 | NA | NA | NA | NA | NA | NA |
| ## | 1206 | NA | NA | NA | NA | NA | NA |
| ## | 1207 | 0.140032 | 0.0769070 | -0.125631 | -0.2358508 | -0.205161 | -1.048e-01 |
| ## | 1208 | -0.594093 | 0.6126364 | 1.431391 | -0.2036803 | -0.526500 | -1.065e+00 |
| ## | 1209 | NA | NA | NA | NA | NA | NA |
| ## | 1210 | -1.198838 | 1.2235296 | 0.012549 | 1.4193881 | -2.340289 | -2.797e+00 |
| ## | 1211 | 0.124876 | 0.2080899 | 0.341920 | -0.3070437 | 0.002468 | -1.978e-01 |
| ## | 1212 | 1.197633 | -0.8720000 | 0.497438 | -0.7577356 | -1.240000 | -1.469e+00 |
| ## | 1213 | NA | NA | NA | NA | NA | NA |
| ## | 1300 | 0.172493 | 0.0165625 | 0.368076 | 0.1661552 | 0.215042 | 3.070e-01 |
| ## | 1301 | NA | NA | NA | NA | NA | NA |
| ## | 1302 | NA | NA | NA | NA | NA | NA |
| ## | 1303 | -0.047612 | -0.0065528 | 0.209417 | 0.2291596 | 0.258571 | 3.050e-01 |
| ## | 1304 | -0.024536 | -0.0376288 | -0.034106 | -0.0393783 | 0.126913 | 8.182e-02 |
| ## | 1305 | -0.022156 | 0.0079759 | 0.567414 | 0.7168616 | 0.681662 | 7.840e-01 |
| ## | 1306 | 0.034061 | 0.1403398 | 0.374669 | 0.5066736 | 0.479541 | 5.435e-01 |
| ## | 1307 | -0.063482 | -0.0263840 | 0.215800 | 0.1420152 | 0.220436 | 2.210e-01 |
| ## | 1308 | -0.083930 | 0.1132636 | 0.555312 | 0.4778956 | 0.528582 | 5.849e-01 |
| ## | 1309 | -0.039423 | 0.1630741 | 0.244390 | 0.2099234 | 0.232017 | 5.127e-01 |
| ## | 1310 | 0.051979 | 0.0650866 | 0.498425 | 0.6582541 | 0.633797 | 7.428e-01 |
| ## | 1311 | -0.093410 | 0.0032593 | 0.389520 | 0.3928079 | 0.419970 | 4.672e-01 |
| ## | 1312 | -0.012929 | 0.0125523 | 0.355696 | 0.4220143 | 0.449986 | 4.874e-01 |
| ## | 1313 | -0.029938 | 0.1146319 | 0.199276 | 0.3169690 | 0.372233 | 3.899e-01 |
| ## | 1314 | -0.016457 | 0.1747147 | 0.121382 | 0.0120759 | 0.164456 | 2.849e-01 |
| ## | 1315 | -0.325674 | 0.1122613 | 0.444028 | 0.4665438 | 0.684470 | 2.100e-01 |
| ## | 1400 | -0.032044 | -0.0351850 | 0.164507 | -0.0768763 | 0.262507 | 5.175e-01 |
| ## | 1401 | 0.350678 | 0.0334757 | 0.151610 | 0.1076884 | 0.491732 | -1.138e-01 |
| ## | 1402 | -0.004137 | -0.2243941 | 0.375265 | -0.6116278 | -0.109575 | 3.041e-01 |
| ## | 1403 | 0.060571 | -0.0169611 | 0.008206 | -0.0724533 | -0.260645 | 4.308e-01 |
| ## | 1404 | NA | NA | NA | NA | NA | NA |
| ## | 1405 | -0.110625 | 0.0658776 | 0.061256 | -0.1634397 | 0.234331 | 7.160e-02 |
| ## | 1406 | -0.081192 | -0.0522260 | 0.280234 | 0.1608034 | 0.391818 | 5.065e-01 |
| ## | 1407 | 0.027905 | -0.0610796 | -0.016921 | 0.1560709 | -0.078296 | -4.309e-01 |
| ## | 1408 | 0.208573 | -0.0519818 | 0.045846 | 0.1158247 | 0.195282 | 1.730e-02 |
| ## | 1409 | -0.066775 | 0.1044932 | 0.006694 | 0.1289371 | -1.027480 | -6.460e-01 |
| ## | 1410 | NA | NA | NA | NA | NA | NA |
| ## | 1500 | 0.107771 | 0.0327322 | 0.396980 | 0.3738360 | 0.365070 | 3.900e-01 |
| ## | 1501 | NA | NA | NA | NA | NA | NA |
| ## | 1502 | -0.110337 | -0.0271799 | 0.588878 | 0.7100873 | 0.753881 | 9.104e-01 |
| ## | 1503 | -0.034384 | -0.0812698 | -0.332828 | -0.2929370 | -0.364496 | -1.415e-02 |
| ## | 1504 | NA | NA | NA | NA | NA | NA |
| ## | 1505 | NA | NA | NA | NA | NA | NA |
| ## | 1506 | NA | NA | NA | NA | NA | NA |
| ## | 1507 | -1.552208 | 0.0173601 | -0.451765 | -1.2931468 | -1.480780 | -9.046e-01 |
| ## | 1508 | -0.316423 | 0.2484478 | 0.747697 | 1.0898861 | 1.135445 | 1.222e+00 |
| ## | 1600 | 0.078582 | -0.1757870 | 0.445595 | 0.5136531 | 0.454860 | 5.181e-01 |
| ## | 1601 | NA | NA | NA | NA | NA | NA |
| ## | 1602 | -0.032844 | 0.0415690 | 0.221629 | 0.1872804 | 0.217587 | 2.386e-01 |

| | | | | | | |
|---------|-----------|------------|-----------|------------|-----------|------------|
| ## 1603 | 0.014512 | 0.0113999 | 0.293244 | 0.3405121 | 0.244346 | 3.554e-01 |
| ## 1604 | -0.050786 | -0.0298608 | -0.094474 | -0.0219852 | -0.154635 | 1.180e-01 |
| ## 1605 | -0.027943 | -0.0098409 | 0.150598 | 0.1495547 | 0.120508 | 2.393e-01 |
| ## 1606 | 0.030663 | -0.0672919 | -0.104853 | -0.1367683 | -0.203774 | -7.140e-02 |
| ## 1607 | 0.019816 | -0.0800145 | -0.042219 | -0.1848166 | -0.195275 | -1.924e-01 |
| ## 1700 | 0.591670 | -0.0922861 | 0.135951 | 0.3366072 | 0.539915 | 3.711e-01 |
| ## 1701 | NA | NA | NA | NA | NA | NA |
| ## 1702 | -0.247386 | 0.3588975 | 0.147651 | -0.0021949 | 0.305457 | 2.126e-01 |
| ## 1703 | -0.394307 | 0.1058284 | 0.229494 | 0.4558337 | 0.648300 | 5.376e-01 |
| ## 1704 | -0.518311 | 0.2915904 | 0.213065 | 0.7330743 | 0.688673 | 1.728e+00 |
| ## 1705 | -0.054219 | -0.2055750 | 0.132421 | -0.0891378 | 0.354100 | 4.170e-02 |
| ## 1706 | -0.129600 | 0.0935175 | 0.226464 | 0.3253419 | 0.337344 | 2.988e-01 |
| ## 1707 | -0.421026 | 0.6346682 | 0.287842 | 0.1432923 | -0.190620 | 3.319e-01 |
| ## 1708 | 0.574843 | -0.5388671 | -0.092113 | -0.1040212 | 0.054982 | -6.754e-01 |
| ## 1709 | -0.001471 | 0.4919049 | 0.101110 | 0.2341711 | 0.406972 | 8.205e-01 |
| ## 1710 | -0.257693 | -0.0940398 | -0.070922 | -0.0797467 | 0.375032 | -4.725e-01 |
| ## 1711 | -0.394863 | -0.0297356 | 0.541349 | 0.5945066 | 0.760368 | 1.369e+00 |
| ## 1712 | -0.140828 | 0.0685598 | -0.065738 | 0.1130667 | 0.314702 | -9.728e-02 |
| ## 1800 | NA | NA | NA | NA | NA | NA |
| ## 1801 | NA | NA | NA | NA | NA | NA |
| ## 1802 | -0.094559 | -0.1304547 | 0.021746 | -0.1963146 | 0.551696 | 6.485e-01 |
| ## 1803 | 0.264082 | 0.4300264 | 0.070633 | -0.2365966 | 0.267980 | 8.138e-02 |
| ## 1804 | -0.435990 | 0.7815045 | 0.337807 | 0.0271656 | -0.348196 | 7.956e-01 |
| ## 1900 | 0.104271 | 0.0374947 | 0.386245 | 0.3947963 | 0.147897 | 1.762e-01 |
| ## 1901 | 0.315035 | -0.0955877 | 0.168798 | 0.2375178 | -0.234032 | 1.082e+00 |
| ## 1902 | -0.142209 | 0.9790165 | 0.628602 | 0.6599127 | 0.863903 | 7.637e-01 |
| ## 1903 | NA | NA | NA | NA | NA | NA |
| ## 1904 | 0.214424 | -0.4614986 | -0.125466 | 0.0336532 | 0.098656 | 6.449e-01 |
| ## 1905 | NA | NA | NA | NA | NA | NA |
| ## 1906 | NA | NA | NA | NA | NA | NA |
| ## 1907 | -0.168820 | -0.0980248 | -0.041424 | 0.5214873 | 0.392777 | 5.238e-01 |
| ## 1908 | NA | NA | NA | NA | NA | NA |
| ## 1909 | -0.122525 | 0.5627520 | 0.089862 | 0.0354120 | 0.748333 | -8.986e-02 |
| ## 1910 | 0.188432 | 0.0614707 | 0.029244 | 0.2364905 | 0.184123 | 2.610e-01 |
| ## 1911 | NA | NA | NA | NA | NA | NA |
| ## 1912 | 0.069942 | -0.3269670 | 0.125576 | 0.1193545 | 0.425372 | -2.455e-01 |
| ## 1913 | NA | NA | NA | NA | NA | NA |
| ## 2000 | 0.148197 | 0.0156896 | 0.584193 | 1.2582802 | 0.837882 | -1.060e+00 |
| ## 2001 | NA | NA | NA | NA | NA | NA |
| ## 2002 | 0.042082 | -0.0129717 | 0.174058 | 0.2053687 | 0.202567 | 5.037e-01 |
| ## 2003 | 0.104653 | 0.0106000 | 0.297061 | 0.2440696 | 0.049831 | -4.682e-01 |
| ## 2100 | 0.050197 | -0.0433721 | 0.270874 | 0.2648995 | 0.214954 | 2.538e-01 |
| ## 2101 | NA | NA | NA | NA | NA | NA |
| ## 2102 | 0.115664 | -0.0483777 | 1.204557 | 0.9743846 | 1.555079 | 1.953e-01 |
| ## 2103 | NA | NA | NA | NA | NA | NA |
| ## 2104 | NA | NA | NA | NA | NA | NA |
| ## 2105 | 0.001102 | 0.2783895 | 0.347742 | 0.5393361 | 0.544620 | -6.131e-02 |
| ## 2200 | 0.186621 | -0.1420625 | 0.069374 | -0.0206410 | 0.296021 | 9.637e-01 |
| ## 2201 | NA | NA | NA | NA | NA | NA |
| ## 2202 | NA | NA | NA | NA | NA | NA |

| | | | | | | |
|---------|-----------|------------|-----------|------------|-----------|------------|
| ## 2203 | NA | NA | NA | NA | NA | NA |
| ## 2204 | NA | NA | NA | NA | NA | NA |
| ## 2205 | -0.180935 | 0.5708931 | 0.306705 | 0.4711825 | 0.608839 | 6.841e-01 |
| ## 2206 | NA | NA | NA | NA | NA | NA |
| ## 2207 | -0.440568 | -0.0455742 | 0.427552 | 0.2073379 | 0.287409 | 3.538e-01 |
| ## 2208 | 0.032398 | -0.0241141 | 0.361340 | 0.5559287 | 0.394622 | 4.101e-01 |
| ## 2209 | 0.015387 | -0.1247426 | 0.146431 | 0.3978309 | 0.283793 | 2.691e-01 |
| ## 2210 | -0.104887 | 0.1567293 | 0.191370 | 0.4579948 | 0.452281 | 3.960e-01 |
| ## 2211 | -0.067681 | -0.0714563 | 0.153592 | 0.2461122 | 0.495228 | 2.264e-01 |
| ## 2212 | NA | NA | NA | NA | NA | NA |
| ## 2213 | -0.051877 | 0.4414543 | 0.081147 | -0.0141893 | 0.391490 | 3.497e-01 |
| ## 2214 | 0.888239 | -0.2189456 | 0.639178 | 0.9273224 | 0.550166 | 7.874e-01 |
| ## 2215 | NA | NA | NA | NA | NA | NA |
| ## 2216 | NA | NA | NA | NA | NA | NA |
| ## 2300 | 0.035671 | -0.1494127 | 0.032416 | 0.2434573 | 0.073722 | 1.429e-01 |
| ## 2301 | -0.311635 | 0.0419462 | -0.102491 | -0.0751326 | -0.078108 | -1.103e+00 |
| ## 2302 | NA | NA | NA | NA | NA | NA |
| ## 2303 | 0.004640 | 0.0174515 | 0.316354 | 0.4931082 | 0.378011 | 5.710e-01 |
| ## 2304 | 0.059573 | -0.1504039 | 0.056064 | 0.1050740 | 0.262618 | 2.601e-01 |
| ## 2305 | 0.078477 | -0.1835950 | 0.703172 | 0.4071335 | 0.319469 | 5.910e-01 |
| ## 2306 | NA | NA | NA | NA | NA | NA |
| ## 2307 | 0.097001 | 0.2571325 | -0.189769 | -0.1044985 | -0.551271 | -3.460e-01 |
| ## 2308 | 0.066574 | 0.0091025 | 0.053753 | 0.1680665 | 0.058738 | 1.051e-01 |
| ## 2309 | 0.018066 | 0.1143388 | -0.143908 | -0.0738070 | -0.188174 | -4.713e-03 |
| ## 2310 | NA | NA | NA | NA | NA | NA |
| ## 2311 | 0.023971 | 0.0009549 | 0.140051 | 0.1252873 | 0.024920 | 2.553e-01 |
| ## 2312 | 0.094701 | -0.1661911 | 0.017884 | 0.0008323 | 0.112579 | 2.350e-01 |
| ## 2400 | NA | NA | NA | NA | NA | NA |
| ## 2401 | NA | NA | NA | NA | NA | NA |
| ## 2402 | -0.021907 | -0.1307328 | 0.457961 | 0.5948620 | 0.638494 | 7.867e-01 |
| ## 2403 | 0.045771 | -0.0693540 | 0.125052 | 0.1517332 | 0.321351 | 3.840e-01 |
| ## 2404 | -0.055472 | -0.0177093 | 0.323135 | 0.4699808 | 0.490479 | 4.994e-01 |
| ## 2405 | -0.046739 | -0.1176722 | -0.194995 | -0.1440645 | 0.146453 | 7.012e-02 |
| ## 2406 | 0.041482 | 0.1490921 | -0.428427 | -0.3472978 | -0.105409 | -2.692e-01 |
| ## 2500 | NA | NA | NA | NA | NA | NA |
| ## 2501 | NA | NA | NA | NA | NA | NA |
| ## 2502 | -0.318250 | 0.1245487 | 0.990206 | 1.3211663 | 1.313376 | 1.337e+00 |
| ## 2503 | -0.033142 | 0.0330980 | 0.146088 | 0.0963022 | 0.081367 | -3.916e-03 |
| ## 2504 | 0.042349 | 0.0279161 | -0.122412 | -0.0939454 | -0.074123 | -4.889e-02 |
| ## 2505 | 0.120147 | -0.0087706 | -0.283927 | -0.2837689 | -0.299079 | -1.736e-01 |
| ## 2506 | 0.026410 | 0.1424435 | -0.420880 | -0.4676530 | -0.484385 | -4.690e-01 |
| ## 2507 | 0.020378 | 0.0616629 | 1.319330 | 1.4980608 | 1.159477 | 1.341e+00 |
| ## 2508 | 0.111277 | 0.0009046 | -0.418835 | -0.3893586 | -0.258337 | -2.549e-01 |
| ## 2600 | -0.242811 | -0.1414508 | 0.220242 | 0.6077835 | 0.134805 | -4.909e-02 |
| ## 2601 | NA | NA | NA | NA | NA | NA |
| ## 2602 | -0.373839 | 0.2055829 | 0.214580 | 0.1839736 | 0.826602 | -9.099e-02 |
| ## 2603 | -0.320479 | -0.5399217 | 0.577284 | -0.0950209 | -0.266504 | -6.704e-02 |
| ## 2604 | 0.089814 | 0.1054573 | 0.278234 | 0.2144690 | 0.562427 | 6.630e-01 |
| ## 2605 | 0.022599 | 0.1940531 | 0.173274 | 0.3152522 | 0.337032 | 3.532e-01 |
| ## 2606 | NA | NA | NA | NA | NA | NA |

| | | | | | | |
|---------|-----------|------------|-----------|------------|-----------|------------|
| ## 2607 | 0.160908 | 0.2844100 | -0.194820 | -0.3216106 | 0.173524 | 1.217e+00 |
| ## 2608 | NA | NA | NA | NA | NA | NA |
| ## 2609 | NA | NA | NA | NA | NA | NA |
| ## 2610 | NA | NA | NA | NA | NA | NA |
| ## 2611 | -0.030864 | 0.1559579 | 0.213001 | 0.1632984 | 0.338835 | 1.819e-01 |
| ## 2612 | NA | NA | NA | NA | NA | NA |
| ## 2613 | -0.075020 | 0.2156466 | 0.247472 | 0.2902680 | 0.186871 | 4.799e-01 |
| ## 2614 | -0.489439 | 0.0598412 | -0.030906 | 0.0054391 | 0.334920 | 1.616e-01 |
| ## 2700 | NA | NA | NA | NA | NA | NA |
| ## 2701 | -0.028331 | 0.1208439 | 0.537776 | 0.4958611 | 0.454347 | 6.370e-01 |
| ## 2702 | NA | NA | NA | NA | NA | NA |
| ## 2703 | -0.146201 | -0.1496037 | 0.434819 | 0.2799548 | 0.370309 | 5.179e-01 |
| ## 2704 | NA | NA | NA | NA | NA | NA |
| ## 2705 | 0.180652 | -0.0328647 | 0.003112 | -0.0528547 | 0.082767 | 2.573e-01 |
| ## 2706 | -0.006935 | 0.0726366 | -0.348734 | 0.1530661 | 0.345905 | 3.593e-01 |
| ## 2707 | NA | NA | NA | NA | NA | NA |
| ## 2708 | -0.460423 | 0.3552243 | 0.912047 | 0.9064645 | 1.035204 | 1.095e+00 |
| ## 2709 | NA | NA | NA | NA | NA | NA |
| ## 2710 | NA | NA | NA | NA | NA | NA |
| ## 2711 | -0.093183 | 0.3899582 | -0.469900 | -0.4075076 | 0.044486 | 1.788e-01 |
| ## 2712 | -0.035123 | 0.3223010 | 0.068941 | 0.8860886 | 0.967751 | 1.149e+00 |
| ## 2713 | 0.227877 | -0.0344200 | 0.347534 | 0.3314632 | 0.098097 | 2.771e-01 |
| ## 2714 | 0.388919 | -0.2707322 | 0.383071 | -0.0149962 | -0.114398 | 2.274e-01 |
| ## 2715 | 0.087196 | 0.0920525 | -0.170551 | -0.4365583 | -0.310908 | -3.955e-02 |
| ## 2716 | -0.190028 | -0.1652493 | 0.334194 | 0.4485494 | 0.213282 | 3.422e-01 |
| ## 2717 | 0.010728 | 0.1035898 | 0.201879 | 0.4241865 | 0.326634 | 4.891e-01 |
| ## 2718 | 0.119578 | -0.0469320 | 0.357362 | 0.5128959 | -0.251901 | 4.838e-01 |
| ## 2719 | 0.218148 | -0.1644137 | 0.011544 | 0.2086301 | 0.302380 | -9.802e-02 |
| ## 2720 | 0.083878 | -0.0107325 | 0.573429 | 0.6763606 | 0.739170 | 5.591e-01 |
| ## 2721 | -0.083249 | -0.2546614 | 0.069521 | -0.4499471 | -0.245893 | -7.277e-02 |
| ## 2722 | -0.125087 | 0.3049641 | 0.484485 | 0.2440000 | 0.190270 | -2.479e-02 |
| ## 2723 | 0.080825 | -0.0861008 | -0.249534 | -0.2137543 | -0.129947 | 3.996e-03 |
| ## 2724 | 0.271128 | 0.2726839 | 0.155051 | 0.2356840 | 0.049991 | 2.578e-01 |
| ## 2725 | -0.048735 | -0.0216868 | -0.322006 | -0.1431122 | 0.003464 | -5.071e-02 |
| ## 2726 | 0.080072 | -0.0757746 | 0.353700 | 0.4843407 | 0.453059 | 4.501e-01 |
| ## 2727 | 0.102402 | 0.1812481 | 0.251852 | 0.0659801 | 0.126212 | 4.386e-01 |
| ## 2728 | -0.017634 | 0.0545987 | 0.332060 | 0.3242713 | 0.331377 | 4.338e-01 |
| ## 2729 | 0.193554 | 0.1048874 | 0.298646 | 0.3309968 | 0.418674 | 2.674e-01 |
| ## 2730 | 0.069793 | 0.1965497 | 0.457476 | 0.4331017 | 0.602263 | 6.306e-01 |
| ## 2731 | -0.029339 | 0.0248960 | 0.252552 | 0.1205300 | 0.317537 | 2.376e-01 |
| ## 2732 | -0.028205 | -0.0192346 | -0.007692 | 0.0937196 | 0.080447 | 1.350e-01 |
| ## 2733 | -0.040127 | 0.2078548 | -0.056546 | 0.0248465 | -0.066269 | -8.059e-02 |
| ## 2734 | -0.232372 | 0.0031925 | -0.038880 | 0.0953870 | 0.205063 | 3.023e-01 |
| ## 2735 | 0.075110 | -0.1561803 | -0.052022 | 0.0316996 | -0.048539 | 1.197e-01 |
| ## 2736 | -0.084226 | -0.0568162 | -0.106174 | -0.0673509 | -0.043379 | 2.019e-01 |
| ## 2737 | -0.004948 | 0.1080277 | 0.022115 | 0.0422361 | 0.324369 | 3.726e-01 |
| ## 2738 | -0.037031 | -0.0072718 | 0.155111 | 0.1468212 | 0.375024 | 4.613e-01 |
| ## 2739 | 0.047767 | 0.0570825 | 0.334087 | 0.4019684 | 0.467630 | 3.529e-01 |
| ## 2740 | 0.456467 | 0.0020978 | -0.466097 | -0.4484228 | -0.103265 | -9.892e-02 |
| ## 2741 | 0.070998 | 0.0690240 | 0.234117 | 0.2588169 | 0.437873 | 5.866e-01 |

| | | | | | | | |
|----|------|-----------|------------|-----------|------------|-----------|------------|
| ## | 2742 | -0.243113 | -0.0216910 | 0.687188 | 0.8170493 | 0.996440 | 8.932e-01 |
| ## | 2743 | 0.245926 | 0.1157953 | 0.330206 | 0.3411578 | 0.422288 | 3.032e-01 |
| ## | 2744 | NA | NA | NA | NA | NA | NA |
| ## | 2745 | 0.266817 | -0.0175122 | 0.108251 | 0.1617922 | -0.025297 | 4.160e-01 |
| ## | 2746 | 0.124514 | -0.1319264 | 0.092596 | 0.2367168 | 0.341841 | 4.784e-01 |
| ## | 2747 | 0.615340 | 0.3301880 | -0.128098 | 0.4045462 | 0.482061 | 9.713e-01 |
| ## | 2748 | -0.127033 | 0.3190600 | 0.346372 | 0.4725695 | 0.382414 | 4.676e-01 |
| ## | 2800 | -0.118331 | 0.1317968 | 0.108248 | 0.2786558 | 0.210817 | 2.394e-01 |
| ## | 2801 | NA | NA | NA | NA | NA | NA |
| ## | 2802 | NA | NA | NA | NA | NA | NA |
| ## | 2803 | -0.028444 | -0.0737756 | 0.067830 | -0.0581848 | 0.246059 | -1.025e-01 |
| ## | 2804 | NA | NA | NA | NA | NA | NA |
| ## | 2805 | 0.083095 | -0.0493590 | 0.234536 | 0.2244342 | 0.709114 | 4.378e-01 |
| ## | 2806 | NA | NA | NA | NA | NA | NA |
| ## | 2807 | NA | NA | NA | NA | NA | NA |
| ## | 2808 | 0.005333 | -0.0685982 | 0.334731 | 0.2783205 | 0.146701 | 3.453e-01 |
| ## | 2809 | 0.179510 | 0.0542250 | 0.063610 | 0.1169675 | 0.374975 | 1.636e-01 |
| ## | 2900 | 0.258054 | -0.0098842 | 0.079348 | 0.2429040 | 0.094640 | 1.525e-01 |
| ## | 2901 | NA | NA | NA | NA | NA | NA |
| ## | 2902 | 0.055076 | 0.1877405 | 0.064203 | 0.4489775 | 0.625271 | 6.509e-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 | 0.468020 | -0.2880961 | -0.203775 | 0.1722772 | -0.551066 | 1.746e-02 |
| ## | 2910 | NA | NA | NA | NA | NA | NA |
| ## | 2911 | NA | NA | NA | NA | NA | NA |
| ## | 2912 | NA | NA | NA | NA | NA | NA |
| ## | 2913 | 0.377591 | 0.3417093 | 0.855663 | 1.8618196 | 1.293168 | -1.450e-02 |
| ## | 2914 | NA | NA | NA | NA | NA | NA |
| ## | 2915 | NA | NA | NA | NA | NA | NA |
| ## | 2916 | 0.069943 | 0.0300784 | 0.365831 | 0.4861604 | 0.307102 | 4.701e-01 |
| ## | 2917 | 0.010361 | 0.8487810 | 0.685900 | 0.4983009 | -0.308731 | 5.530e-01 |
| ## | 2918 | NA | NA | NA | NA | NA | NA |
| ## | 2919 | -0.483963 | 0.5269644 | -0.168785 | 0.9925662 | 0.399568 | -2.579e+00 |
| ## | 2920 | NA | NA | NA | NA | NA | NA |
| ## | 2921 | NA | NA | NA | NA | NA | NA |
| ## | 2922 | NA | NA | NA | NA | NA | NA |
| ## | 2923 | NA | NA | NA | NA | NA | NA |
| ## | 3000 | NA | NA | NA | NA | NA | NA |
| ## | 3001 | NA | NA | NA | NA | NA | NA |
| ## | 3002 | 0.006980 | 0.0249971 | 0.235612 | 0.2342906 | 0.307835 | 3.217e-01 |
| ## | 3003 | 0.040377 | 0.0500475 | 0.609655 | 0.6152676 | 0.563919 | 5.969e-01 |
| ## | 3004 | -0.069916 | -0.0322924 | 0.445399 | 0.4159210 | 0.499488 | 5.421e-01 |
| ## | 3005 | -0.133737 | -0.0004676 | 0.212937 | -0.0429804 | 0.397475 | 1.662e-01 |
| ## | 3100 | 0.022277 | -0.0205215 | 0.090952 | 0.1763649 | 0.199724 | 4.714e-01 |
| ## | 3101 | -0.074791 | 0.1530828 | -0.003157 | 0.3451977 | 0.127582 | 5.285e-01 |
| ## | 3102 | NA | NA | NA | NA | NA | NA |

| | | | | | | |
|---------|-----------|------------|-----------|------------|-----------|------------|
| ## 3103 | 0.134367 | -0.5075439 | 0.191067 | 0.0738521 | 0.718099 | -1.978e-01 |
| ## 3104 | 0.103672 | -0.0198499 | 0.138833 | 0.1796633 | 0.238388 | 2.333e-01 |
| ## 3105 | -0.019347 | 0.2063135 | 0.042636 | 0.2968997 | 0.107711 | 2.109e-01 |
| ## 3106 | -0.525905 | 0.7632845 | 0.273060 | 0.5789238 | -0.532388 | 2.890e-01 |
| ## 3107 | -0.070569 | -0.1379085 | 0.427279 | 0.6165159 | 0.486557 | 4.554e-01 |
| ## 3108 | NA | NA | NA | NA | NA | NA |
| ## 3109 | NA | NA | NA | NA | NA | NA |
| ## 3110 | 0.122931 | 0.1622656 | -0.272105 | -0.2236234 | -0.229649 | -2.320e-01 |
| ## 3200 | -0.094106 | 0.0351392 | 0.256083 | 0.0525218 | 0.402612 | 2.409e-01 |
| ## 3201 | NA | NA | NA | NA | NA | NA |
| ## 3202 | 0.020909 | 0.0266941 | 0.033499 | -0.1303385 | -0.021413 | -1.254e-01 |
| ## 3203 | 0.112195 | -0.0045923 | 0.188126 | 0.3517795 | 0.446441 | 3.954e-01 |
| ## 3204 | -0.057429 | -0.0871169 | -0.044874 | 0.4093664 | 0.278780 | 7.546e-02 |
| ## 3205 | 0.099837 | -0.0106876 | 0.071005 | 0.1439628 | 0.397929 | 2.158e-01 |
| ## 3206 | 0.136343 | -0.4400036 | 0.289911 | 0.2922721 | 0.600588 | 1.312e-01 |
| ## 3207 | -0.230457 | 0.3315436 | 0.243330 | 0.4059653 | 0.123684 | -1.510e-01 |
| ## 3300 | 0.282299 | -0.2002835 | -0.080995 | 0.1243030 | 0.081827 | -7.590e-01 |
| ## 3301 | -0.050488 | -0.0831955 | 0.187236 | 0.3876447 | -0.149846 | 9.710e-02 |
| ## 3302 | 0.493947 | -0.3910646 | 0.379967 | 0.6764199 | -0.064965 | -3.162e-01 |
| ## 3303 | 0.022748 | -0.0136877 | 0.059640 | -0.1266484 | 0.617279 | 4.551e-01 |
| ## 3304 | 0.082884 | 0.0589856 | 0.165847 | 0.2180893 | 0.231949 | 2.357e-01 |
| ## 3305 | -0.024032 | 0.0480381 | 0.234647 | 0.1707646 | 0.278888 | -6.744e-02 |
| ## 3306 | -0.148326 | 0.1164246 | 0.073220 | 0.1226410 | 0.051483 | 3.105e-01 |
| ## 3307 | NA | NA | NA | NA | NA | NA |
| ## 3308 | -0.083230 | -0.1080269 | 0.227508 | 0.0277379 | 0.169917 | -2.767e-02 |
| ## 3309 | 0.196109 | -0.2126512 | 0.125579 | 0.2278235 | -0.016999 | 3.199e-01 |
| ## 3310 | -0.121936 | 0.2684163 | 0.248694 | 0.3674724 | 0.703282 | 1.697e-02 |
| ## 3311 | NA | NA | NA | NA | NA | NA |
| ## 3312 | 0.018196 | 0.0176084 | 0.107153 | 0.0260011 | 0.064971 | -3.588e-02 |
| ## 3313 | -0.127066 | 0.3024201 | 0.159628 | 0.0641288 | -0.067610 | 6.457e-01 |
| ## 3314 | 1.088628 | -0.9274629 | 0.447128 | 0.7352042 | 0.242000 | 1.529e-01 |
| ## 3315 | -0.475268 | 0.4118391 | 0.152089 | 0.4223629 | 0.192069 | -6.107e-01 |
| ## 3316 | 0.030042 | -0.1034379 | 0.399185 | 0.2738871 | 0.526045 | 6.262e-01 |
| ## 3317 | 0.188181 | -0.2746588 | 0.182357 | -0.0055637 | 0.414469 | 8.359e-01 |
| ## 3318 | 0.351675 | -0.2162407 | 0.536418 | 0.6607333 | -0.145235 | 6.752e-01 |
| ## 3319 | NA | NA | NA | NA | NA | NA |
| ## 3320 | -0.218538 | 0.2663516 | 0.119462 | 0.0960713 | 0.264509 | -4.053e-01 |
| ## 3321 | 0.173028 | -0.0290488 | 0.070181 | -0.0599793 | 0.293779 | 3.850e-01 |
| ## 3322 | 0.192300 | -0.3016938 | -0.135868 | 0.2228297 | 0.415828 | 1.301e-01 |
| ## 3400 | 0.025173 | -0.0349482 | 0.535975 | 0.7238778 | 0.652406 | 7.718e-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 | 0.047149 | -0.3401187 | 0.613249 | 0.6479809 | 0.308835 | -1.114e-01 |
| ## 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 | NA | NA | NA | NA | NA | NA |
| ## 3612 | -0.023719 | -0.1231929 | 0.434885 | 0.4043370 | 0.550738 | 5.175e-01 |
| ## 3613 | NA | NA | NA | NA | NA | NA |
| ## 3614 | -0.264003 | 0.2812793 | 0.153648 | 0.5237142 | 0.273041 | 6.537e-01 |
| ## 3615 | NA | NA | NA | NA | NA | NA |
| ## 3616 | 0.097067 | 0.4778116 | -0.541820 | -0.5698758 | 0.272256 | -1.018e+00 |
| ## | FFA2 | FLA2 | FFA3 | FLA4 | | |
| ## 1000 | NA | NA | NA | NA | | |
| ## 1100 | 0.0661375 | -0.0231004 | 0.064212 | -0.0173275 | | |
| ## 1101 | 0.2659232 | 0.1803586 | 0.281687 | NA | | |
| ## 1102 | 0.0347689 | 0.0882048 | 0.050128 | 0.0955576 | | |
| ## 1103 | -0.0085138 | -0.1426835 | -0.019813 | -0.1440025 | | |
| ## 1104 | -0.0815973 | 0.0127915 | -0.079668 | -0.0040231 | | |
| ## 1105 | -0.0524388 | 0.0526868 | -0.042045 | 0.0397828 | | |
| ## 1106 | 0.0430723 | -0.0299586 | 0.038485 | -0.0215672 | | |
| ## 1107 | 0.0072247 | 0.1946051 | 0.045232 | 0.1967250 | | |
| ## 1108 | NA | NA | NA | NA | | |
| ## 1109 | 0.2794725 | 0.1574253 | 0.302810 | 0.1759745 | | |
| ## 1110 | -0.0875918 | 0.1093095 | -0.076869 | 0.0972872 | | |
| ## 1111 | -0.0417516 | -0.2143259 | -0.039558 | -0.2135249 | | |
| ## 1200 | 0.4350696 | -0.9847397 | -0.470613 | -0.5990245 | | |
| ## 1201 | 0.0792582 | -0.0721583 | 0.035975 | -0.0260713 | | |
| ## 1202 | -0.2493922 | 0.3071515 | 0.034890 | 0.0800893 | | |
| ## 1203 | -0.1645428 | 0.2706520 | 0.060305 | 0.1255105 | | |
| ## 1204 | -0.0295422 | 0.1160210 | 0.067464 | 0.0904602 | | |
| ## 1205 | NA | NA | NA | NA | | |
| ## 1206 | NA | NA | NA | NA | | |
| ## 1207 | 0.1015994 | 0.0979871 | 0.127469 | 0.1307272 | | |
| ## 1208 | 0.0809548 | -0.0653151 | 0.017491 | 0.0137928 | | |
| ## 1209 | NA | NA | NA | NA | | |
| ## 1210 | -1.1970147 | 1.1200274 | -0.426694 | -0.0783411 | | |
| ## 1211 | 0.2343809 | 0.1109654 | 0.323682 | 0.3192833 | | |
| ## 1212 | 1.6288130 | -1.2657484 | 0.435540 | 0.3239897 | | |
| ## 1213 | NA | NA | 0.069217 | 0.0692174 | | |
| ## 1300 | 0.2333945 | 0.0679936 | 0.229341 | 0.0381507 | | |
| ## 1301 | NA | NA | NA | NA | | |
| ## 1302 | -0.2134681 | -0.0843109 | -0.190358 | -0.0073604 | | |
| ## 1303 | -0.0394297 | -0.0040991 | -0.039788 | -0.0106369 | | |

| | | | | |
|---------|------------|------------|-----------|------------|
| ## 1304 | -0.0179692 | -0.0219304 | -0.022100 | -0.0264456 |
| ## 1305 | 0.0267436 | 0.0371594 | 0.027935 | 0.0383470 |
| ## 1306 | 0.0975004 | 0.1419992 | 0.125739 | 0.1620448 |
| ## 1307 | -0.0468657 | -0.0230105 | -0.049783 | -0.0292809 |
| ## 1308 | -0.0832443 | 0.1079970 | -0.070626 | 0.0955380 |
| ## 1309 | 0.0457605 | 0.1794607 | 0.058654 | 0.1824694 |
| ## 1310 | 0.0325461 | 0.0271185 | 0.037278 | 0.0386833 |
| ## 1311 | -0.0650284 | -0.0095688 | -0.065125 | -0.0197793 |
| ## 1312 | 0.0089243 | 0.0210291 | 0.011422 | 0.0224291 |
| ## 1313 | -0.0083044 | 0.1047611 | -0.002683 | 0.1047563 |
| ## 1314 | 0.0235781 | 0.1324896 | 0.048999 | 0.1403019 |
| ## 1315 | -0.2405436 | 0.1104822 | -0.199565 | 0.0263550 |
| ## 1400 | 0.0322043 | -0.0796400 | -0.004238 | -0.0615306 |
| ## 1401 | 0.3459175 | 0.0269545 | 0.364806 | 0.2958244 |
| ## 1402 | 0.1831750 | -0.2579123 | 0.008041 | -0.1166342 |
| ## 1403 | 0.0653081 | -0.0170878 | 0.055820 | 0.0259895 |
| ## 1404 | -1.2430346 | 0.1327289 | -1.134249 | -0.5950971 |
| ## 1405 | -0.0698672 | 0.0409777 | -0.043704 | 0.0084044 |
| ## 1406 | -0.1200079 | -0.1000028 | -0.156814 | -0.1491270 |
| ## 1407 | -0.0042067 | -0.0226579 | -0.014320 | -0.0238812 |
| ## 1408 | 0.2173237 | -0.0578367 | 0.184140 | 0.0832622 |
| ## 1409 | -0.1404020 | 0.2038675 | -0.036046 | 0.1252015 |
| ## 1410 | NA | NA | NA | NA |
| ## 1500 | 0.1418913 | 0.0363448 | 0.146619 | 0.0665140 |
| ## 1501 | NA | NA | NA | NA |
| ## 1502 | -0.0798911 | 0.0005476 | -0.080139 | 0.0057728 |
| ## 1503 | -0.0363778 | -0.1077852 | -0.024628 | -0.1015917 |
| ## 1504 | NA | NA | NA | NA |
| ## 1505 | 0.1923350 | 0.0404751 | 0.202560 | -0.0009776 |
| ## 1506 | NA | NA | NA | NA |
| ## 1507 | -0.4003540 | -0.0808116 | -0.425842 | -0.1634660 |
| ## 1508 | -0.1935442 | 0.1931144 | -0.253836 | 0.3143754 |
| ## 1600 | 0.0978231 | -0.1797913 | 0.085734 | -0.1703840 |
| ## 1601 | NA | NA | NA | NA |
| ## 1602 | -0.0316631 | 0.0281949 | -0.029955 | 0.0245260 |
| ## 1603 | 0.0153140 | -0.0264649 | 0.015813 | -0.0285285 |
| ## 1604 | -0.0550777 | -0.0581420 | -0.058920 | -0.0631860 |
| ## 1605 | -0.0197397 | -0.0106693 | -0.020520 | -0.0126952 |
| ## 1606 | 0.0253178 | -0.0753082 | 0.021844 | -0.0736794 |
| ## 1607 | 0.0136595 | -0.0374471 | 0.010736 | -0.0336177 |
| ## 1700 | 0.6360085 | -0.0803857 | 0.599451 | 0.2736575 |
| ## 1701 | NA | NA | NA | NA |
| ## 1702 | -0.2677417 | 0.3311906 | -0.170240 | NA |
| ## 1703 | -0.3197767 | 0.1475965 | -0.294806 | 0.0960065 |
| ## 1704 | -0.2232089 | 0.2551849 | -0.100132 | 0.1867146 |
| ## 1705 | -0.0873132 | -0.2175063 | -0.177984 | -0.2720413 |
| ## 1706 | -0.1303732 | 0.0496650 | -0.120291 | 0.0127135 |
| ## 1707 | -0.4299552 | 0.5830467 | -0.327232 | 0.5049398 |
| ## 1708 | 0.6188347 | -0.5193928 | 0.420178 | -0.2115840 |
| ## 1709 | -0.1388485 | 0.5606851 | 0.039970 | 0.4195027 |

| | | | | | |
|----|------|------------|------------|-----------|------------|
| ## | 1710 | -0.2280565 | -0.0952734 | -0.268447 | -0.1984281 |
| ## | 1711 | -0.3002175 | -0.0399110 | -0.296689 | 0.0522712 |
| ## | 1712 | -0.1453355 | 0.0090670 | -0.143883 | -0.0190900 |
| ## | 1800 | NA | NA | NA | NA |
| ## | 1801 | NA | NA | NA | NA |
| ## | 1802 | -0.0800763 | -0.1918208 | -0.220255 | -0.2310824 |
| ## | 1803 | 0.1793907 | 0.4929292 | 0.291531 | 0.5796737 |
| ## | 1804 | -0.1604226 | 0.5531496 | 0.049894 | 0.4636079 |
| ## | 1900 | 0.1677550 | -0.0733031 | 0.124121 | 0.0416545 |
| ## | 1901 | 0.4502785 | -0.2778372 | 0.281065 | 0.0249127 |
| ## | 1902 | 0.2413802 | 0.7073801 | 0.411233 | 0.9207687 |
| ## | 1903 | NA | NA | NA | NA |
| ## | 1904 | 0.2090048 | -0.4837561 | -0.067616 | -0.3188643 |
| ## | 1905 | NA | NA | NA | NA |
| ## | 1906 | 0.1782939 | 0.1475000 | 0.147610 | 0.1475000 |
| ## | 1907 | 0.1500840 | -0.1540861 | 0.145070 | -0.1396984 |
| ## | 1908 | -0.3207026 | -0.2220848 | -0.298490 | -0.1864424 |
| ## | 1909 | -0.1319974 | 0.5916924 | -0.132004 | NA |
| ## | 1910 | 0.1752730 | 0.0068191 | 0.176067 | 0.0818640 |
| ## | 1911 | -0.1535481 | 0.4416583 | -0.075306 | 0.4049245 |
| ## | 1912 | -0.0186058 | -0.2229432 | -0.065193 | -0.2286796 |
| ## | 1913 | NA | NA | NA | NA |
| ## | 2000 | 0.1924976 | -0.0887641 | NA | NA |
| ## | 2001 | 0.0512803 | 0.5351643 | 0.119190 | 0.5464973 |
| ## | 2002 | 0.0489578 | -0.0215975 | 0.035555 | 0.0087347 |
| ## | 2003 | 0.1419510 | -0.0296345 | 0.124377 | 0.0587611 |
| ## | 2100 | 0.0683722 | -0.0864698 | 0.045669 | -0.0563719 |
| ## | 2101 | NA | NA | NA | NA |
| ## | 2102 | 0.2429702 | 0.2509467 | 0.229962 | 0.2106211 |
| ## | 2103 | NA | NA | NA | NA |
| ## | 2104 | NA | NA | NA | NA |
| ## | 2105 | -0.0574208 | 0.1735748 | -0.003570 | 0.1498393 |
| ## | 2200 | 0.2946094 | -0.0238687 | 0.291598 | 0.0285836 |
| ## | 2201 | -0.0963408 | -0.2646221 | -0.154771 | -0.3310333 |
| ## | 2202 | -0.2950768 | 0.4273805 | -0.189122 | 0.3529041 |
| ## | 2203 | NA | NA | NA | NA |
| ## | 2204 | -0.0084045 | 0.0382681 | -0.004632 | 0.0361950 |
| ## | 2205 | -0.2869904 | 0.5208207 | -0.057959 | 0.3618709 |
| ## | 2206 | NA | NA | NA | NA |
| ## | 2207 | -0.2913908 | -0.1796419 | -0.348861 | -0.3872113 |
| ## | 2208 | 0.0911765 | 0.0275397 | 0.091722 | 0.0317821 |
| ## | 2209 | 0.0283705 | -0.1057058 | 0.012157 | -0.1001325 |
| ## | 2210 | -0.0384104 | 0.1923046 | -0.045859 | 0.1946679 |
| ## | 2211 | 0.0473132 | -0.0079063 | 0.046223 | -0.0037296 |
| ## | 2212 | NA | NA | NA | NA |
| ## | 2213 | -0.0399639 | 0.3817265 | 0.105288 | 0.3602571 |
| ## | 2214 | 0.8250221 | -0.4692886 | 0.602097 | -0.2063807 |
| ## | 2215 | -0.5621027 | 0.5154900 | -0.494966 | 0.4730467 |
| ## | 2216 | NA | NA | NA | NA |
| ## | 2300 | 0.0404399 | -0.1524316 | 0.021739 | -0.1464878 |

| | | | | | |
|----|------|------------|------------|-----------|------------|
| ## | 2301 | -0.3318973 | 0.0890064 | -0.273774 | -0.0800253 |
| ## | 2302 | NA | NA | NA | NA |
| ## | 2303 | 0.0513361 | -0.0256690 | 0.050279 | -0.0217136 |
| ## | 2304 | 0.0471261 | -0.1867985 | 0.017930 | -0.1721771 |
| ## | 2305 | 0.1333809 | -0.1601873 | 0.102801 | -0.1301426 |
| ## | 2306 | NA | NA | NA | NA |
| ## | 2307 | -0.1026845 | 0.1592854 | -0.124092 | 0.1733155 |
| ## | 2308 | 0.0583961 | -0.0098736 | 0.053615 | 0.0274407 |
| ## | 2309 | 0.0286674 | 0.1493068 | 0.065613 | 0.1551682 |
| ## | 2310 | 0.0899949 | -0.2387313 | 0.083705 | -0.2332010 |
| ## | 2311 | 0.0002358 | 0.0129522 | 0.001203 | 0.0102819 |
| ## | 2312 | 0.1000105 | -0.1631962 | 0.096216 | -0.1617186 |
| ## | 2400 | 0.1902775 | -0.1846829 | 0.082458 | -0.0868234 |
| ## | 2401 | NA | NA | NA | NA |
| ## | 2402 | 0.0364397 | -0.0647693 | 0.033710 | -0.0616009 |
| ## | 2403 | 0.0954628 | -0.0546361 | 0.093399 | -0.0489414 |
| ## | 2404 | -0.0282034 | 0.0037254 | -0.027925 | -0.0001277 |
| ## | 2405 | -0.0023362 | -0.1068967 | 0.018315 | -0.1063644 |
| ## | 2406 | 0.0242047 | 0.1554669 | 0.035204 | 0.1611722 |
| ## | 2500 | 0.2504207 | -0.2146370 | 0.245258 | -0.2097433 |
| ## | 2501 | NA | NA | NA | NA |
| ## | 2502 | -0.2482557 | 0.1323902 | -0.242170 | 0.0926402 |
| ## | 2503 | -0.0420421 | 0.0529145 | -0.044709 | 0.0557322 |
| ## | 2504 | 0.0374625 | 0.0273689 | 0.038910 | 0.0311266 |
| ## | 2505 | 0.1164276 | -0.0242632 | 0.116358 | -0.0228715 |
| ## | 2506 | -0.0015546 | 0.1157394 | -0.007156 | 0.1154112 |
| ## | 2507 | 0.0674996 | 0.1657320 | 0.048705 | 0.1397770 |
| ## | 2508 | 0.0952999 | -0.0215269 | 0.094594 | -0.0160205 |
| ## | 2600 | -0.2658176 | -0.0157757 | -0.274698 | -0.1735646 |
| ## | 2601 | NA | NA | NA | NA |
| ## | 2602 | NA | NA | -0.255760 | 0.1364719 |
| ## | 2603 | 0.0116763 | -0.2193914 | 0.010695 | -0.2236408 |
| ## | 2604 | 0.2286181 | 0.0722835 | 0.254861 | 0.1760188 |
| ## | 2605 | 0.0733031 | 0.2438056 | 0.134180 | 0.2605388 |
| ## | 2606 | NA | NA | NA | NA |
| ## | 2607 | 0.1533303 | 0.1733483 | 0.195328 | 0.1871143 |
| ## | 2608 | -0.7503068 | 0.3969267 | -0.540144 | 0.1963890 |
| ## | 2609 | NA | NA | NA | NA |
| ## | 2610 | NA | NA | -0.371126 | -0.3711264 |
| ## | 2611 | 0.0299837 | 0.1593702 | 0.044860 | 0.1649565 |
| ## | 2612 | 0.0983519 | 0.5483388 | 0.319310 | 0.5671743 |
| ## | 2613 | 0.0259862 | 0.1691390 | 0.083611 | 0.1810879 |
| ## | 2614 | -0.5132829 | 0.0469537 | -0.493881 | -0.1382864 |
| ## | 2700 | 0.1067139 | -0.0506332 | 0.086003 | -0.0006497 |
| ## | 2701 | 0.0891624 | 0.0389471 | 0.099971 | 0.0706209 |
| ## | 2702 | -0.2133220 | 0.1088722 | -0.147062 | 0.0146494 |
| ## | 2703 | -0.1438218 | -0.2346889 | -0.185935 | -0.2731839 |
| ## | 2704 | -0.1435481 | 0.0350382 | -0.129441 | -0.0393732 |
| ## | 2705 | 0.1796123 | -0.0639236 | 0.174029 | -0.0386623 |
| ## | 2706 | 0.0736964 | -0.2550170 | -0.014061 | -0.2281176 |

| | | | | | |
|----|------|------------|------------|-----------|------------|
| ## | 2707 | -0.0810000 | 0.6380000 | 0.949000 | NA |
| ## | 2708 | -0.2360926 | 0.1664397 | -0.208019 | 0.1280234 |
| ## | 2709 | NA | NA | NA | NA |
| ## | 2710 | NA | NA | NA | NA |
| ## | 2711 | -0.1733974 | 0.3721216 | -0.071322 | 0.3095390 |
| ## | 2712 | -0.0220731 | 0.2349559 | 0.047867 | 0.2281770 |
| ## | 2713 | 0.2470733 | -0.0796896 | 0.233018 | -0.0196380 |
| ## | 2714 | 0.3853107 | -0.2787123 | 0.374022 | -0.3143485 |
| ## | 2715 | 0.0480910 | 0.0889671 | 0.058448 | 0.1002183 |
| ## | 2716 | -0.1614560 | -0.1651884 | -0.160743 | -0.1673888 |
| ## | 2717 | 0.0252466 | 0.0581972 | 0.020201 | 0.0538240 |
| ## | 2718 | 0.2999056 | -0.0039704 | 0.298034 | 0.1597163 |
| ## | 2719 | 0.1561870 | -0.1106199 | 0.106390 | -0.0397606 |
| ## | 2720 | 0.1261206 | -0.0148733 | 0.127287 | -0.0218725 |
| ## | 2721 | -0.1260973 | -0.1719476 | -0.132974 | -0.1811467 |
| ## | 2722 | -0.1965918 | 0.1663766 | -0.182827 | 0.1522228 |
| ## | 2723 | 0.0973629 | -0.1192329 | 0.085989 | -0.1042254 |
| ## | 2724 | 0.3115851 | 0.2719376 | 0.384754 | 0.3965325 |
| ## | 2725 | -0.0157266 | 0.0192180 | -0.013192 | 0.0160127 |
| ## | 2726 | 0.0828725 | -0.0992844 | 0.077638 | -0.0973703 |
| ## | 2727 | 0.0898994 | 0.1167157 | 0.104001 | 0.1390027 |
| ## | 2728 | -0.0040305 | 0.0642996 | 0.009040 | 0.0632251 |
| ## | 2729 | 0.1405866 | 0.0240188 | 0.142287 | 0.0317295 |
| ## | 2730 | 0.0974408 | 0.1627809 | 0.144190 | 0.1945190 |
| ## | 2731 | NA | NA | NA | 0.0248236 |
| ## | 2732 | -0.0524266 | -0.0415959 | -0.070234 | -0.0707117 |
| ## | 2733 | -0.0641001 | 0.1970188 | -0.022418 | 0.1634320 |
| ## | 2734 | -0.2131392 | 0.0131073 | -0.211242 | -0.0216137 |
| ## | 2735 | 0.0597763 | -0.1607736 | 0.041289 | -0.1517172 |
| ## | 2736 | -0.0134995 | -0.0819577 | -0.023921 | -0.0838456 |
| ## | 2737 | 0.0149123 | 0.0747951 | 0.038846 | 0.0831223 |
| ## | 2738 | -0.0045156 | -0.0362606 | -0.009775 | -0.0370546 |
| ## | 2739 | 0.0668058 | 0.0379154 | 0.077620 | 0.0554338 |
| ## | 2740 | 0.4611617 | 0.0391645 | 0.459886 | -0.0459628 |
| ## | 2741 | 0.1079011 | -0.0338965 | 0.102672 | -0.0086268 |
| ## | 2742 | -0.2003412 | -0.0559778 | -0.217616 | -0.1185635 |
| ## | 2743 | 0.2448919 | 0.0488338 | 0.247194 | 0.0595250 |
| ## | 2744 | NA | NA | NA | NA |
| ## | 2745 | 0.3694703 | -0.0442474 | 0.367868 | -0.0266064 |
| ## | 2746 | 0.1003563 | -0.1617477 | 0.093001 | -0.1502519 |
| ## | 2747 | 0.5807886 | 0.2012167 | 0.611533 | 0.3675002 |
| ## | 2748 | -0.1210522 | 0.3130146 | -0.080903 | 0.2652109 |
| ## | 2800 | -0.1098124 | 0.1471646 | -0.090773 | 0.1264228 |
| ## | 2801 | NA | NA | NA | NA |
| ## | 2802 | NA | NA | -0.060775 | NA |
| ## | 2803 | 0.0002510 | -0.0165087 | 0.001116 | -0.0164742 |
| ## | 2804 | 0.0391497 | 0.1558176 | 0.059693 | 0.1634137 |
| ## | 2805 | 0.1564402 | -0.1508366 | 0.106634 | -0.0673804 |
| ## | 2806 | NA | NA | NA | NA |
| ## | 2807 | 0.1569802 | -0.2330121 | 0.201803 | -0.2529734 |

| | | | | | |
|----|------|------------|------------|-----------|------------|
| ## | 2808 | 0.0019943 | -0.0627864 | -0.007922 | -0.0625965 |
| ## | 2809 | 0.1908530 | -0.0400127 | 0.180060 | 0.0097527 |
| ## | 2900 | 0.2669484 | 0.0093284 | 0.271376 | 0.1395202 |
| ## | 2901 | NA | NA | NA | NA |
| ## | 2902 | -0.0792017 | 0.1404295 | -0.017268 | 0.1140592 |
| ## | 2903 | NA | NA | NA | NA |
| ## | 2904 | NA | NA | NA | NA |
| ## | 2905 | -0.0830780 | 0.5779918 | 0.311769 | 0.5127885 |
| ## | 2906 | NA | NA | NA | NA |
| ## | 2907 | -0.1294376 | 0.6080585 | 0.160602 | 0.5555172 |
| ## | 2908 | NA | NA | NA | NA |
| ## | 2909 | 0.3154391 | -0.2462574 | 0.193281 | -0.0698249 |
| ## | 2910 | NA | NA | NA | NA |
| ## | 2911 | 0.4398550 | -0.2728843 | 0.222814 | -0.0717245 |
| ## | 2912 | NA | NA | NA | NA |
| ## | 2913 | 1.2930319 | -0.6269838 | 0.852210 | -0.0439651 |
| ## | 2914 | 0.3875846 | -0.0123517 | 0.378307 | 0.2039627 |
| ## | 2915 | NA | NA | NA | NA |
| ## | 2916 | 0.1692178 | 0.0340589 | 0.177782 | 0.0602684 |
| ## | 2917 | 0.0742494 | 0.8431457 | 0.523786 | NA |
| ## | 2918 | NA | NA | NA | NA |
| ## | 2919 | -0.3129331 | 0.6423656 | 0.075539 | 0.4435849 |
| ## | 2920 | NA | NA | NA | NA |
| ## | 2921 | -0.0908668 | 0.3902997 | -0.072490 | 0.3820702 |
| ## | 2922 | NA | NA | NA | NA |
| ## | 2923 | NA | NA | NA | NA |
| ## | 3000 | NA | NA | NA | NA |
| ## | 3001 | NA | NA | NA | NA |
| ## | 3002 | 0.0198581 | 0.0338172 | 0.021948 | 0.0349293 |
| ## | 3003 | 0.0585024 | 0.0479407 | 0.059973 | 0.0471866 |
| ## | 3004 | 0.0028474 | 0.0065235 | 0.002707 | 0.0067516 |
| ## | 3005 | -0.1754000 | 0.0161798 | -0.170461 | 0.0419178 |
| ## | 3100 | 0.0420446 | 0.0355219 | 0.047242 | 0.0461240 |
| ## | 3101 | 0.1296552 | 0.1375260 | 0.197846 | 0.2363269 |
| ## | 3102 | 0.3482705 | 0.2096306 | 0.380775 | 0.2118026 |
| ## | 3103 | -0.0624603 | -0.2836171 | -0.168904 | -0.3249133 |
| ## | 3104 | 0.1266709 | -0.0173824 | 0.124877 | 0.0021770 |
| ## | 3105 | 0.0234134 | 0.2657749 | 0.036782 | 0.2677329 |
| ## | 3106 | -0.2035743 | 0.2749909 | -0.065023 | 0.1507707 |
| ## | 3107 | -0.0346940 | -0.0672712 | -0.033883 | -0.0664353 |
| ## | 3108 | NA | NA | NA | NA |
| ## | 3109 | NA | NA | NA | NA |
| ## | 3110 | 0.1023372 | 0.1473082 | 0.111488 | 0.1610575 |
| ## | 3200 | -0.1151969 | 0.0232880 | -0.105750 | -0.0247129 |
| ## | 3201 | NA | NA | NA | NA |
| ## | 3202 | 0.0087730 | 0.0317319 | 0.015369 | 0.0341048 |
| ## | 3203 | 0.1460878 | -0.0001677 | 0.145773 | 0.0525174 |
| ## | 3204 | 0.0341234 | -0.0773542 | 0.006380 | -0.0657534 |
| ## | 3205 | 0.1613798 | -0.0675955 | 0.132824 | 0.0123980 |
| ## | 3206 | 0.1818404 | -0.4125860 | 0.110137 | -0.3682116 |

| | | | | | |
|----|------|------------|------------|-----------|------------|
| ## | 3207 | -0.2030113 | 0.3776687 | -0.007050 | 0.2671973 |
| ## | 3300 | 0.2342865 | -0.1573256 | 0.096841 | 0.0451315 |
| ## | 3301 | 0.0554174 | -0.1292441 | -0.046119 | -0.0801575 |
| ## | 3302 | 0.2447164 | -0.1663063 | 0.090722 | 0.0477160 |
| ## | 3303 | 0.0041309 | 0.0144961 | 0.013687 | 0.0173802 |
| ## | 3304 | 0.1079770 | 0.0379692 | 0.131732 | 0.1033837 |
| ## | 3305 | -0.0205263 | 0.0401952 | 0.005459 | 0.0265658 |
| ## | 3306 | -0.1203439 | 0.1175814 | -0.061232 | 0.0602755 |
| ## | 3307 | NA | NA | NA | NA |
| ## | 3308 | -0.0059054 | -0.1960912 | -0.147890 | -0.2011308 |
| ## | 3309 | 0.2067416 | -0.2474249 | 0.057438 | -0.1254081 |
| ## | 3310 | -0.0965100 | 0.2528212 | 0.116474 | 0.1694165 |
| ## | 3311 | NA | NA | NA | NA |
| ## | 3312 | 0.0224262 | 0.0155679 | 0.034568 | 0.0330235 |
| ## | 3313 | -0.0901519 | 0.2397138 | NA | 0.2235381 |
| ## | 3314 | 0.6415000 | -0.5352686 | 0.144294 | 0.1069689 |
| ## | 3315 | -0.5102912 | 0.4270597 | -0.152519 | 0.0136147 |
| ## | 3316 | 0.1360968 | -0.2034685 | -0.045697 | -0.0816170 |
| ## | 3317 | 0.1409438 | -0.2259312 | 0.002206 | -0.1451679 |
| ## | 3318 | 0.3028836 | -0.3249334 | NA | -0.1500059 |
| ## | 3319 | NA | NA | NA | NA |
| ## | 3320 | -0.2375782 | 0.2905329 | 0.035809 | 0.0736830 |
| ## | 3321 | 0.1850191 | -0.0381336 | 0.153169 | 0.0978067 |
| ## | 3322 | 0.1819205 | -0.2854528 | 0.028966 | -0.1888169 |
| ## | 3400 | 0.0688593 | -0.1232640 | 0.032665 | -0.0985049 |
| ## | 3401 | NA | NA | NA | NA |
| ## | 3402 | NA | NA | NA | NA |
| ## | 3403 | NA | NA | NA | NA |
| ## | 3404 | NA | NA | NA | NA |
| ## | 3500 | -0.0124877 | -0.3659140 | -0.098547 | -0.3697874 |
| ## | 3501 | NA | NA | NA | NA |
| ## | 3502 | NA | NA | NA | NA |
| ## | 3503 | NA | NA | NA | NA |
| ## | 3504 | NA | NA | NA | NA |
| ## | 3505 | NA | NA | NA | NA |
| ## | 3506 | NA | NA | NA | NA |
| ## | 3600 | NA | NA | NA | NA |
| ## | 3601 | 0.2448428 | 0.2807959 | 0.379860 | 0.3737262 |
| ## | 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 | NA | NA | NA | NA |
| ## | 3612 | -0.0368788 | -0.1605434 | -0.114717 | -0.1792011 |
| ## | 3613 | NA | NA | NA | NA |

```
## 3614 -0.0776235  0.1279182 -0.036858  0.0963945
## 3615          NA          NA          NA          NA
## 3616 -0.1490815  0.3952168  0.198725  0.3250286
```

```
print(RegP)
```

```
##          FFA1p      FLA1p          2p          3p          4p          5+p      FFA2p
## 1000          NA          NA          NA          NA          NA          NA          NA
## 1100 9.139e-01 4.955e-01 4.745e-03 8.735e-07 1.564e-04 8.991e-07 3.736e-01
## 1101          NA          NA          NA          NA          NA          NA 9.947e-02
## 1102 9.056e-01 5.911e-01 3.632e-02 1.399e-03 1.502e-02 2.092e-04 6.733e-01
## 1103 4.581e-01 7.565e-02 3.629e-02 9.088e-03 1.972e-02 8.546e-04 9.155e-01
## 1104 5.983e-02 8.502e-01 3.048e-01 7.651e-01 3.507e-01 8.229e-01 7.642e-02
## 1105 6.369e-02 3.071e-01 1.717e-05 3.663e-05 4.357e-03 2.437e-03 2.462e-01
## 1106 8.002e-01 6.461e-01 1.578e-14 9.574e-18 4.482e-19 1.020e-21 2.235e-01
## 1107 8.432e-01 5.569e-01 1.041e-01 9.653e-01 7.652e-01 3.521e-01 9.289e-01
## 1108          NA          NA          NA          NA          NA          NA          NA
## 1109 1.540e-01 3.299e-01 2.016e-01 6.837e-01 5.906e-02 1.599e-01 4.885e-03
## 1110 6.679e-01 5.718e-01 1.597e-06 2.535e-02 1.660e-03 6.466e-04 3.345e-01
## 1111 7.819e-01 9.627e-02 8.107e-01 6.467e-01 8.690e-01 4.941e-01 7.546e-01
## 1200 1.257e-01 2.011e-02 1.748e-04 2.558e-01 4.516e-04 1.000e+00 6.131e-01
## 1201 5.655e-01 9.583e-01 3.376e-02 3.340e-01 1.171e-05 1.571e-04 3.967e-01
## 1202 5.423e-01 3.600e-01 2.599e-04 6.014e-01 4.398e-04 5.226e-01 2.345e-01
## 1203 2.309e-01 7.613e-02 1.313e-02 8.173e-02 2.036e-01 9.712e-01 3.105e-01
## 1204 6.007e-01 3.756e-01 7.421e-02 7.531e-01 3.538e-01 9.191e-01 9.245e-01
## 1205          NA          NA          NA          NA          NA          NA          NA
## 1206          NA          NA          NA          NA          NA          NA          NA
## 1207 4.175e-02 3.119e-01 1.693e-01 7.184e-03 1.286e-01 4.104e-01 1.355e-01
## 1208 4.159e-01 3.945e-01 4.537e-02 8.286e-01 4.078e-01 6.126e-02 8.775e-01
## 1209          NA          NA          NA          NA          NA          NA          NA
## 1210 5.598e-11 2.911e-04 9.714e-01 1.485e-06 3.329e-07 1.497e-08 6.076e-12
## 1211 6.463e-01 4.394e-01 1.100e-01 4.475e-01 9.911e-01 2.963e-01 3.082e-01
## 1212 3.267e-04 0.000e+00 1.469e-01 2.587e-06 0.000e+00 2.601e-06 6.346e-10
## 1213          NA          NA          NA          NA          NA          NA          NA
## 1300 9.592e-02 8.910e-01 7.217e-02 4.078e-01 2.587e-01 1.093e-01 1.724e-02
## 1301          NA          NA          NA          NA          NA          NA          NA
## 1302          NA          NA          NA          NA          NA          NA 7.348e-02
## 1303 5.648e-02 8.251e-01 2.109e-01 1.663e-01 1.163e-01 6.169e-02 1.092e-01
## 1304 6.647e-01 6.152e-01 8.106e-01 7.615e-01 3.682e-01 5.247e-01 7.535e-01
## 1305 6.312e-01 8.824e-01 5.706e-04 2.700e-05 4.986e-05 1.401e-06 5.787e-01
## 1306 6.234e-01 6.350e-02 2.076e-01 3.603e-03 4.378e-03 1.271e-04 1.818e-01
## 1307 1.854e-01 5.709e-01 8.065e-02 2.486e-01 7.542e-02 6.462e-02 3.322e-01
## 1308 1.850e-01 1.489e-01 1.792e-02 6.408e-02 2.781e-02 7.923e-03 2.030e-01
## 1309 6.210e-01 3.446e-02 1.555e-01 2.475e-01 2.052e-01 3.582e-03 6.007e-01
## 1310 5.415e-01 5.493e-01 2.993e-01 1.333e-01 1.445e-01 1.007e-01 6.976e-01
## 1311 4.507e-02 9.498e-01 1.540e-04 1.309e-04 7.032e-05 6.592e-07 1.625e-01
## 1312 6.450e-01 7.127e-01 2.728e-06 2.133e-08 3.144e-09 2.183e-11 7.580e-01
## 1313 6.421e-01 1.054e-01 4.491e-01 1.558e-01 1.425e-01 8.422e-02 8.946e-01
## 1314 8.060e-01 8.401e-02 4.595e-01 9.407e-01 3.680e-01 6.224e-02 7.193e-01
## 1315 8.233e-02 4.580e-01 3.305e-01 3.750e-01 2.353e-01 7.099e-01 2.246e-01
```

| | | | | | | | | |
|----|------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| ## | 1400 | 7.555e-01 | 7.286e-01 | 1.005e-01 | 6.524e-01 | 6.136e-02 | 3.694e-03 | 7.536e-01 |
| ## | 1401 | 1.134e-01 | 8.583e-01 | 3.708e-01 | 7.814e-01 | 1.885e-01 | 7.027e-01 | 4.193e-02 |
| ## | 1402 | 9.809e-01 | 2.153e-01 | 1.139e-02 | 9.174e-03 | 4.417e-01 | 2.897e-01 | 3.402e-01 |
| ## | 1403 | 6.796e-01 | 9.173e-01 | 9.491e-01 | 6.245e-01 | 3.388e-01 | 2.208e-01 | 6.603e-01 |
| ## | 1404 | NA | NA | NA | NA | NA | NA | 2.181e-03 |
| ## | 1405 | 6.736e-01 | 7.335e-01 | 6.392e-01 | 3.825e-01 | 7.044e-01 | 7.811e-01 | 7.552e-01 |
| ## | 1406 | 5.621e-01 | 6.604e-01 | 3.549e-02 | 2.675e-01 | 7.282e-02 | 4.890e-03 | 3.671e-01 |
| ## | 1407 | 8.183e-01 | 5.602e-01 | 8.612e-01 | 2.675e-01 | 7.127e-01 | 3.318e-02 | 9.720e-01 |
| ## | 1408 | 2.313e-02 | 5.930e-01 | 5.288e-01 | 2.682e-01 | 3.482e-01 | 8.680e-01 | 1.994e-02 |
| ## | 1409 | 6.701e-01 | 4.894e-01 | 9.750e-01 | 5.319e-01 | 1.385e-05 | 0.000e+00 | 4.021e-01 |
| ## | 1410 | NA | NA | NA | NA | NA | NA | NA |
| ## | 1500 | 2.652e-01 | 7.668e-01 | 5.653e-04 | 1.910e-03 | 4.257e-03 | 3.643e-03 | 1.132e-01 |
| ## | 1501 | NA | NA | NA | NA | NA | NA | NA |
| ## | 1502 | 9.823e-02 | 7.938e-01 | 9.171e-05 | 3.465e-06 | 9.419e-07 | 3.287e-10 | 2.974e-01 |
| ## | 1503 | 6.141e-01 | 3.314e-01 | 2.579e-01 | 3.220e-01 | 2.282e-01 | 9.645e-01 | 6.003e-01 |
| ## | 1504 | NA | NA | NA | NA | NA | NA | NA |
| ## | 1505 | NA | NA | NA | NA | NA | NA | 3.312e-01 |
| ## | 1506 | NA | NA | NA | NA | NA | NA | NA |
| ## | 1507 | 4.321e-07 | 9.194e-01 | 1.005e-02 | 4.155e-06 | 5.513e-06 | 3.302e-03 | 4.573e-02 |
| ## | 1508 | 1.235e-01 | 2.293e-01 | 7.047e-03 | 7.827e-06 | 4.289e-04 | 7.074e-06 | 4.125e-01 |
| ## | 1600 | 3.336e-02 | 3.236e-04 | 7.325e-03 | 1.450e-03 | 5.515e-03 | 1.428e-03 | 1.085e-02 |
| ## | 1601 | NA | NA | NA | NA | NA | NA | NA |
| ## | 1602 | 3.910e-01 | 4.741e-01 | 6.242e-02 | 1.244e-01 | 8.636e-02 | 5.531e-02 | 4.087e-01 |
| ## | 1603 | 7.884e-01 | 9.018e-01 | 7.824e-02 | 3.548e-02 | 1.408e-01 | 3.156e-02 | 7.819e-01 |
| ## | 1604 | 3.947e-01 | 6.137e-01 | 6.938e-01 | 9.235e-01 | 5.060e-01 | 5.917e-01 | 3.608e-01 |
| ## | 1605 | 3.435e-01 | 7.809e-01 | 3.192e-01 | 3.292e-01 | 4.316e-01 | 1.188e-01 | 5.047e-01 |
| ## | 1606 | 4.073e-01 | 1.098e-01 | 2.946e-01 | 1.577e-01 | 4.000e-02 | 4.558e-01 | 4.852e-01 |
| ## | 1607 | 7.113e-01 | 5.743e-01 | 7.964e-01 | 2.228e-01 | 1.975e-01 | 1.900e-01 | 7.958e-01 |
| ## | 1700 | 1.227e-03 | 6.289e-01 | 4.811e-01 | 9.834e-02 | 2.892e-02 | 1.732e-01 | 4.729e-04 |
| ## | 1701 | NA | NA | NA | NA | NA | NA | NA |
| ## | 1702 | 1.105e-01 | 5.605e-02 | 2.809e-01 | 9.890e-01 | 1.276e-01 | 5.199e-01 | 8.833e-02 |
| ## | 1703 | 3.523e-02 | 7.130e-01 | 3.231e-01 | 2.203e-02 | 7.906e-03 | 1.849e-02 | 1.211e-01 |
| ## | 1704 | 2.008e-01 | 3.323e-01 | 5.635e-01 | 8.863e-02 | 1.665e-01 | 2.933e-06 | 6.512e-01 |
| ## | 1705 | 7.455e-01 | 2.368e-01 | 3.085e-01 | 5.350e-01 | 5.756e-02 | 9.081e-01 | 5.876e-01 |
| ## | 1706 | 1.896e-01 | 4.169e-01 | 1.244e-02 | 1.634e-03 | 5.500e-03 | 9.992e-03 | 1.853e-01 |
| ## | 1707 | 5.983e-02 | 1.135e-02 | 2.620e-01 | 4.586e-01 | 7.653e-01 | 2.107e-01 | 2.862e-02 |
| ## | 1708 | 6.274e-02 | 9.148e-02 | 6.506e-01 | 6.000e-01 | 8.107e-01 | 2.582e-02 | 3.621e-02 |
| ## | 1709 | 9.966e-01 | 3.646e-01 | 7.088e-01 | 3.520e-01 | 9.323e-02 | 1.558e-03 | 6.643e-01 |
| ## | 1710 | 4.495e-02 | 4.954e-01 | 5.652e-01 | 5.968e-01 | 2.877e-02 | 1.393e-01 | 8.114e-02 |
| ## | 1711 | 1.519e-01 | 9.706e-01 | 6.186e-02 | 4.687e-02 | 2.102e-02 | 4.383e-04 | 2.956e-01 |
| ## | 1712 | 3.043e-01 | 6.556e-01 | 5.794e-01 | 3.580e-01 | 2.920e-02 | 6.993e-01 | 2.971e-01 |
| ## | 1800 | NA | NA | NA | NA | NA | NA | NA |
| ## | 1801 | NA | NA | NA | NA | NA | NA | NA |
| ## | 1802 | 6.770e-01 | 5.344e-01 | 9.266e-01 | 3.456e-01 | 4.516e-01 | 2.006e-01 | 7.273e-01 |
| ## | 1803 | 3.026e-01 | 1.257e-01 | 6.134e-01 | 8.700e-02 | 1.760e-01 | 8.126e-01 | 4.286e-01 |
| ## | 1804 | 1.275e-01 | 6.130e-03 | 7.182e-02 | 9.067e-01 | 3.012e-01 | 5.610e-03 | 4.058e-01 |
| ## | 1900 | 2.653e-01 | 6.879e-01 | 3.827e-04 | 1.452e-02 | 5.029e-01 | 3.152e-01 | 1.300e-01 |
| ## | 1901 | 7.172e-02 | 6.082e-01 | 2.251e-01 | 3.191e-01 | 1.242e-01 | 1.877e-06 | 5.199e-02 |
| ## | 1902 | 5.002e-01 | 4.094e-04 | 4.201e-03 | 6.410e-04 | 2.435e-05 | 4.894e-05 | 1.674e-01 |
| ## | 1903 | NA | NA | NA | NA | NA | NA | NA |

| | | | | | | | | |
|----|------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| ## | 1904 | 1.551e-01 | 1.980e-02 | 4.211e-01 | 8.595e-01 | 7.036e-01 | 4.346e-02 | 1.251e-01 |
| ## | 1905 | NA | NA | NA | NA | NA | NA | NA |
| ## | 1906 | NA | NA | NA | NA | NA | NA | 3.487e-01 |
| ## | 1907 | 3.837e-01 | 4.268e-01 | 7.579e-01 | 6.837e-03 | 1.873e-02 | 1.925e-01 | 3.776e-01 |
| ## | 1908 | NA | NA | NA | NA | NA | NA | 7.043e-02 |
| ## | 1909 | 6.763e-01 | 7.333e-02 | 7.709e-01 | 8.714e-01 | 1.434e-01 | 7.709e-01 | 4.954e-01 |
| ## | 1910 | 4.752e-02 | 6.165e-01 | 7.886e-01 | 4.809e-02 | 1.450e-01 | 2.388e-01 | 6.647e-02 |
| ## | 1911 | NA | NA | NA | NA | NA | NA | 3.368e-01 |
| ## | 1912 | 7.760e-01 | 3.924e-01 | 5.265e-01 | 7.240e-01 | 2.516e-01 | 5.514e-01 | 9.025e-01 |
| ## | 1913 | NA | NA | NA | NA | NA | NA | NA |
| ## | 2000 | 6.783e-01 | 9.719e-01 | 6.638e-03 | 4.198e-02 | 2.179e-02 | 1.375e-01 | 4.932e-01 |
| ## | 2001 | NA | NA | NA | NA | NA | NA | 6.454e-01 |
| ## | 2002 | 5.552e-01 | 8.503e-01 | 1.486e-03 | 7.809e-03 | 9.232e-02 | 3.002e-02 | 4.895e-01 |
| ## | 2003 | 5.456e-01 | 9.434e-01 | 8.930e-03 | 2.426e-01 | 8.216e-01 | 2.942e-02 | 3.773e-01 |
| ## | 2100 | 7.112e-01 | 8.144e-01 | 1.034e-01 | 1.190e-01 | 3.533e-01 | 1.220e-01 | 5.985e-01 |
| ## | 2101 | NA | NA | NA | NA | NA | NA | NA |
| ## | 2102 | 5.458e-01 | 8.359e-01 | 2.916e-03 | 1.609e-02 | 4.679e-04 | 6.649e-01 | 2.842e-01 |
| ## | 2103 | NA | NA | NA | NA | NA | NA | NA |
| ## | 2104 | NA | NA | NA | NA | NA | NA | NA |
| ## | 2105 | 9.909e-01 | 3.330e-02 | 2.775e-02 | 3.151e-03 | 1.189e-02 | 8.483e-01 | 6.339e-01 |
| ## | 2200 | 2.665e-01 | 4.743e-01 | 6.516e-01 | 9.035e-01 | 2.169e-01 | 1.075e-02 | 9.704e-02 |
| ## | 2201 | NA | NA | NA | NA | NA | NA | 7.546e-01 |
| ## | 2202 | NA | NA | NA | NA | NA | NA | 3.850e-01 |
| ## | 2203 | NA | NA | NA | NA | NA | NA | NA |
| ## | 2204 | NA | NA | NA | NA | NA | NA | 9.263e-01 |
| ## | 2205 | 2.916e-01 | 4.998e-03 | 1.030e-01 | 8.270e-03 | 9.970e-03 | 6.476e-02 | 1.319e-01 |
| ## | 2206 | NA | NA | NA | NA | NA | NA | NA |
| ## | 2207 | 2.445e-01 | 9.122e-01 | 1.082e-01 | 3.251e-01 | 1.737e-01 | 2.341e-01 | 3.671e-01 |
| ## | 2208 | 6.648e-01 | 8.077e-01 | 1.361e-03 | 7.425e-06 | 7.086e-04 | 6.410e-04 | 2.254e-01 |
| ## | 2209 | 8.387e-01 | 2.302e-01 | 4.177e-01 | 4.061e-02 | 1.330e-01 | 1.693e-01 | 7.142e-01 |
| ## | 2210 | 2.199e-01 | 2.312e-01 | 2.078e-01 | 7.621e-03 | 9.021e-03 | 1.552e-02 | 6.710e-01 |
| ## | 2211 | 5.334e-01 | 6.578e-01 | 2.584e-01 | 1.343e-01 | 1.415e-02 | 1.556e-01 | 6.065e-01 |
| ## | 2212 | NA | NA | NA | NA | NA | NA | NA |
| ## | 2213 | 8.757e-01 | 3.746e-02 | 7.851e-01 | 9.490e-01 | 2.088e-02 | 2.021e-01 | 8.608e-01 |
| ## | 2214 | 1.674e-02 | 4.379e-01 | 4.799e-04 | 1.847e-04 | 5.680e-02 | 8.675e-03 | 4.845e-05 |
| ## | 2215 | NA | NA | NA | NA | NA | NA | 7.168e-02 |
| ## | 2216 | NA | NA | NA | NA | NA | NA | NA |
| ## | 2300 | 6.677e-01 | 1.671e-01 | 7.471e-01 | 2.578e-02 | 5.674e-01 | 2.383e-01 | 6.373e-01 |
| ## | 2301 | 9.754e-02 | 7.789e-01 | 5.340e-01 | 5.988e-01 | 7.870e-01 | 6.207e-07 | 4.089e-02 |
| ## | 2302 | NA | NA | NA | NA | NA | NA | NA |
| ## | 2303 | 9.280e-01 | 8.107e-01 | 9.140e-03 | 2.070e-04 | 5.832e-03 | 3.093e-05 | 3.342e-01 |
| ## | 2304 | 3.210e-01 | 1.956e-01 | 6.993e-01 | 4.449e-01 | 6.576e-02 | 4.656e-02 | 4.323e-01 |
| ## | 2305 | 4.213e-01 | 5.174e-02 | 6.585e-02 | 3.213e-01 | 4.370e-01 | 1.673e-01 | 2.306e-01 |
| ## | 2306 | NA | NA | NA | NA | NA | NA | NA |
| ## | 2307 | 7.460e-01 | 2.973e-01 | 5.974e-01 | 8.307e-01 | 4.538e-01 | 3.673e-01 | 5.135e-01 |
| ## | 2308 | 3.693e-01 | 9.145e-01 | 4.510e-01 | 1.340e-01 | 6.107e-01 | 3.559e-01 | 4.021e-01 |
| ## | 2309 | 8.787e-01 | 4.338e-01 | 5.398e-01 | 7.555e-01 | 5.153e-01 | 9.821e-01 | 8.019e-01 |
| ## | 2310 | NA | NA | NA | NA | NA | NA | 3.833e-01 |
| ## | 2311 | 8.212e-01 | 9.950e-01 | 6.271e-01 | 6.867e-01 | 9.310e-01 | 4.292e-01 | 9.985e-01 |
| ## | 2312 | 2.455e-01 | 1.149e-01 | 8.795e-01 | 9.952e-01 | 3.645e-01 | 9.392e-02 | 1.510e-01 |

| | | | | | | | |
|---------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| ## 2400 | NA | NA | NA | NA | NA | NA | 2.727e-01 |
| ## 2401 | NA | NA | NA | NA | NA | NA | NA |
| ## 2402 | 6.294e-01 | 3.421e-02 | 1.361e-02 | 1.514e-03 | 7.553e-04 | 2.882e-05 | 4.633e-01 |
| ## 2403 | 3.694e-01 | 2.747e-01 | 4.959e-01 | 4.218e-01 | 9.195e-02 | 3.821e-02 | 5.736e-02 |
| ## 2404 | 1.410e-01 | 7.534e-01 | 3.521e-02 | 2.288e-03 | 1.966e-03 | 1.544e-03 | 4.415e-01 |
| ## 2405 | 6.641e-01 | 4.211e-01 | 5.213e-01 | 5.867e-01 | 5.860e-01 | 8.248e-01 | 9.823e-01 |
| ## 2406 | 6.484e-01 | 4.338e-01 | 1.923e-03 | 2.872e-02 | 4.171e-01 | 3.218e-02 | 8.079e-01 |
| ## 2500 | NA | NA | NA | NA | NA | NA | 4.059e-04 |
| ## 2501 | NA | NA | NA | NA | NA | NA | NA |
| ## 2502 | 2.110e-03 | 2.179e-01 | 4.979e-08 | 3.410e-13 | 1.862e-11 | 4.532e-14 | 1.618e-02 |
| ## 2503 | 6.715e-01 | 7.871e-01 | 6.099e-01 | 7.327e-01 | 7.729e-01 | 9.896e-01 | 5.852e-01 |
| ## 2504 | 4.137e-01 | 6.819e-01 | 4.305e-01 | 5.322e-01 | 6.301e-01 | 7.530e-01 | 4.640e-01 |
| ## 2505 | 1.077e-02 | 8.748e-01 | 1.957e-01 | 2.023e-01 | 1.729e-01 | 4.195e-01 | 9.575e-03 |
| ## 2506 | 7.490e-01 | 7.599e-02 | 4.667e-02 | 4.160e-02 | 2.452e-02 | 3.582e-02 | 9.843e-01 |
| ## 2507 | 8.614e-01 | 7.888e-01 | 1.941e-12 | 5.699e-09 | 1.972e-07 | 6.704e-09 | 6.225e-01 |
| ## 2508 | 7.533e-02 | 9.892e-01 | 2.463e-03 | 3.822e-03 | 5.567e-02 | 4.932e-02 | 1.029e-01 |
| ## 2600 | 8.941e-02 | 3.106e-01 | 2.841e-02 | 7.031e-06 | 4.703e-01 | 9.208e-01 | 5.028e-02 |
| ## 2601 | NA | NA | NA | NA | NA | NA | NA |
| ## 2602 | 3.543e-02 | 1.364e-01 | 7.537e-02 | 4.526e-01 | 3.019e-01 | 9.677e-01 | NA |
| ## 2603 | 1.806e-01 | 1.629e-02 | 9.336e-04 | 8.460e-01 | 5.301e-01 | 8.441e-01 | 9.653e-01 |
| ## 2604 | 3.705e-01 | 3.660e-01 | 1.743e-03 | 6.982e-02 | 1.144e-05 | 1.837e-03 | 2.985e-02 |
| ## 2605 | 9.019e-01 | 2.101e-01 | 5.574e-01 | 2.381e-01 | 3.397e-01 | 4.272e-01 | 6.992e-01 |
| ## 2606 | NA | NA | NA | NA | NA | NA | NA |
| ## 2607 | 2.474e-01 | 1.139e-01 | 3.116e-01 | 2.430e-01 | 2.560e-01 | 9.490e-03 | 2.541e-01 |
| ## 2608 | NA | NA | NA | NA | NA | NA | 1.007e-03 |
| ## 2609 | NA | NA | NA | NA | NA | NA | NA |
| ## 2610 | NA | NA | NA | NA | NA | NA | NA |
| ## 2611 | 8.226e-01 | 4.140e-01 | 1.381e-01 | 3.284e-01 | 6.377e-02 | 3.796e-01 | 8.018e-01 |
| ## 2612 | NA | NA | NA | NA | NA | NA | 4.179e-01 |
| ## 2613 | 6.994e-01 | 3.101e-01 | 8.088e-02 | 1.381e-01 | 1.921e-01 | 7.175e-02 | 8.903e-01 |
| ## 2614 | 1.334e-02 | 7.144e-01 | 8.285e-01 | 9.753e-01 | 2.678e-01 | 6.825e-01 | 9.401e-03 |
| ## 2700 | NA | NA | NA | NA | NA | NA | 1.863e-01 |
| ## 2701 | 7.037e-01 | 1.317e-01 | 7.808e-08 | 5.205e-06 | 2.223e-03 | 2.635e-08 | 2.655e-01 |
| ## 2702 | NA | NA | NA | NA | NA | NA | 4.981e-01 |
| ## 2703 | 1.549e-01 | 2.956e-01 | 4.479e-02 | 2.205e-01 | 7.864e-02 | 1.055e-02 | 1.571e-01 |
| ## 2704 | NA | NA | NA | NA | NA | NA | 3.914e-01 |
| ## 2705 | 4.670e-03 | 7.068e-01 | 9.826e-01 | 6.611e-01 | 4.993e-01 | 1.952e-02 | 4.683e-03 |
| ## 2706 | 9.662e-01 | 8.632e-01 | 4.912e-01 | 7.218e-01 | 4.260e-01 | 4.330e-01 | 5.367e-01 |
| ## 2707 | NA | NA | NA | NA | NA | NA | NA |
| ## 2708 | 2.892e-03 | 1.946e-02 | 3.658e-06 | 1.413e-04 | 1.781e-05 | 8.134e-06 | 2.418e-01 |
| ## 2709 | NA | NA | NA | NA | NA | NA | NA |
| ## 2710 | NA | NA | NA | NA | NA | NA | NA |
| ## 2711 | 6.017e-01 | 3.134e-02 | 7.086e-02 | 1.624e-01 | 8.548e-01 | 5.590e-01 | 3.476e-01 |
| ## 2712 | 7.450e-01 | 2.001e-03 | 6.822e-01 | 4.981e-03 | 1.428e-06 | 3.272e-08 | 8.703e-01 |
| ## 2713 | 9.378e-02 | 8.483e-01 | 2.298e-01 | 3.781e-01 | 7.225e-01 | 3.793e-01 | 4.307e-02 |
| ## 2714 | 4.107e-03 | 0.000e+00 | 2.121e-01 | 9.642e-01 | 7.411e-01 | 4.645e-01 | 2.002e-02 |
| ## 2715 | 2.896e-01 | 3.953e-01 | 3.046e-01 | 1.187e-02 | 5.098e-02 | 7.620e-01 | 5.533e-01 |
| ## 2716 | 1.613e-01 | 8.313e-02 | 2.077e-01 | 1.515e-01 | 5.527e-01 | 4.939e-01 | 6.243e-02 |
| ## 2717 | 9.301e-01 | 2.612e-01 | 3.696e-01 | 2.048e-01 | 2.638e-01 | 2.150e-02 | 8.110e-01 |
| ## 2718 | 6.298e-01 | 8.771e-01 | 1.328e-01 | 1.826e-02 | 2.513e-01 | 6.119e-02 | 1.521e-01 |

| | | | | | | | | |
|----|------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| ## | 2719 | 7.536e-02 | 1.934e-01 | 9.192e-01 | 2.315e-01 | 8.365e-02 | 5.880e-01 | 2.684e-01 |
| ## | 2720 | 3.785e-01 | 8.965e-01 | 2.162e-02 | 6.104e-03 | 4.974e-03 | 1.248e-02 | 2.233e-01 |
| ## | 2721 | 5.792e-01 | 1.064e-01 | 8.810e-01 | 4.415e-01 | 5.860e-01 | 8.767e-01 | 3.093e-01 |
| ## | 2722 | 6.912e-01 | 2.140e-01 | 3.685e-01 | 0.000e+00 | 4.590e-01 | 8.834e-01 | 2.501e-01 |
| ## | 2723 | 2.954e-01 | 2.962e-01 | 2.115e-01 | 2.949e-01 | 5.278e-01 | 9.834e-01 | 1.884e-01 |
| ## | 2724 | 3.854e-02 | 1.276e-01 | 4.953e-01 | 4.388e-01 | 7.733e-01 | 1.141e-01 | 3.105e-02 |
| ## | 2725 | 4.328e-01 | 7.536e-01 | 3.652e-02 | 3.690e-01 | 9.798e-01 | 6.949e-01 | 7.982e-01 |
| ## | 2726 | 2.462e-01 | 3.077e-01 | 1.116e-01 | 2.608e-02 | 5.271e-02 | 3.445e-02 | 2.896e-01 |
| ## | 2727 | 4.211e-01 | 3.201e-01 | 4.698e-01 | 8.738e-01 | 7.366e-01 | 2.216e-01 | 4.880e-01 |
| ## | 2728 | 7.582e-01 | 3.329e-01 | 3.113e-02 | 3.069e-02 | 2.863e-02 | 2.756e-03 | 9.461e-01 |
| ## | 2729 | 7.944e-02 | 3.885e-01 | 3.476e-01 | 1.766e-01 | 7.430e-02 | 2.653e-01 | 1.194e-01 |
| ## | 2730 | 2.847e-01 | 5.421e-03 | 1.818e-02 | 9.984e-03 | 1.275e-03 | 1.047e-04 | 1.650e-01 |
| ## | 2731 | 8.220e-01 | 8.968e-01 | 3.013e-01 | 5.737e-01 | 1.427e-01 | 2.723e-01 | NA |
| ## | 2732 | 7.324e-01 | 8.341e-01 | 9.487e-01 | 4.255e-01 | 5.019e-01 | 2.255e-01 | 5.080e-01 |
| ## | 2733 | 7.347e-01 | 2.279e-01 | 8.064e-01 | 8.915e-01 | 7.238e-01 | 6.855e-01 | 5.853e-01 |
| ## | 2734 | 6.404e-03 | 9.686e-01 | 8.798e-01 | 7.313e-01 | 4.138e-01 | 2.618e-01 | 1.482e-02 |
| ## | 2735 | 1.518e-01 | 2.267e-02 | 7.892e-01 | 8.697e-01 | 8.015e-01 | 5.328e-01 | 2.586e-01 |
| ## | 2736 | 4.466e-01 | 5.843e-01 | 5.768e-01 | 6.555e-01 | 8.313e-01 | 9.869e-02 | 8.933e-01 |
| ## | 2737 | 9.591e-01 | 3.492e-01 | 9.661e-01 | 9.362e-01 | 5.547e-01 | 4.639e-01 | 8.865e-01 |
| ## | 2738 | 4.314e-01 | 8.896e-01 | 2.639e-02 | 4.454e-02 | 6.764e-05 | 1.931e-10 | 9.243e-01 |
| ## | 2739 | 5.000e-01 | 3.853e-01 | 6.949e-03 | 1.071e-03 | 1.329e-04 | 3.453e-03 | 3.662e-01 |
| ## | 2740 | 2.975e-04 | 9.938e-01 | 6.042e-02 | 3.320e-02 | 6.313e-01 | 6.013e-01 | 1.508e-04 |
| ## | 2741 | 2.886e-01 | 4.403e-01 | 8.560e-02 | 6.462e-02 | 1.253e-03 | 1.844e-06 | 1.180e-01 |
| ## | 2742 | 2.209e-02 | 8.631e-01 | 3.785e-03 | 9.514e-04 | 6.253e-04 | 6.902e-04 | 9.217e-02 |
| ## | 2743 | 2.042e-01 | 6.442e-01 | 1.831e-01 | 1.514e-01 | 3.632e-02 | 2.897e-01 | 1.263e-01 |
| ## | 2744 | NA | NA | NA | NA | NA | NA | NA |
| ## | 2745 | 3.175e-02 | 8.762e-01 | 5.518e-01 | 4.994e-01 | 9.036e-01 | 6.706e-02 | 2.403e-03 |
| ## | 2746 | 2.972e-02 | 1.045e-01 | 6.320e-01 | 1.981e-01 | 6.076e-02 | 7.879e-03 | 9.662e-02 |
| ## | 2747 | 4.912e-02 | 2.460e-01 | 5.140e-01 | 1.983e-01 | 1.112e-01 | 4.436e-03 | 9.843e-02 |
| ## | 2748 | 1.957e-01 | 1.550e-01 | 3.271e-01 | 6.684e-02 | 1.374e-01 | 3.898e-02 | 1.955e-01 |
| ## | 2800 | 3.905e-02 | 4.561e-02 | 6.088e-01 | 2.286e-01 | 3.121e-01 | 2.651e-01 | 4.710e-02 |
| ## | 2801 | NA | NA | NA | NA | NA | NA | NA |
| ## | 2802 | NA | NA | NA | NA | NA | NA | NA |
| ## | 2803 | 7.409e-01 | 5.256e-01 | 8.230e-01 | 8.764e-01 | 5.502e-01 | 7.381e-01 | 9.976e-01 |
| ## | 2804 | NA | NA | NA | NA | NA | NA | 5.412e-01 |
| ## | 2805 | 6.162e-01 | 8.375e-01 | 3.536e-01 | 3.288e-01 | 7.011e-03 | 3.995e-02 | 1.930e-01 |
| ## | 2806 | NA | NA | NA | NA | NA | NA | NA |
| ## | 2807 | NA | NA | NA | NA | NA | NA | 7.391e-01 |
| ## | 2808 | 9.364e-01 | 3.121e-01 | 4.959e-02 | 1.005e-01 | 3.841e-01 | 3.384e-02 | 9.763e-01 |
| ## | 2809 | 2.005e-01 | 6.607e-01 | 7.639e-01 | 4.453e-01 | 1.200e-01 | 3.590e-01 | 1.611e-01 |
| ## | 2900 | 6.047e-03 | 9.093e-01 | 3.833e-01 | 1.542e-02 | 4.845e-01 | 2.334e-01 | 5.655e-03 |
| ## | 2901 | NA | NA | NA | NA | NA | NA | NA |
| ## | 2902 | 7.395e-01 | 2.603e-01 | 6.611e-01 | 1.301e-01 | 2.002e-06 | 4.194e-03 | 6.918e-01 |
| ## | 2903 | NA | NA | NA | NA | NA | NA | NA |
| ## | 2904 | NA | NA | NA | NA | NA | NA | NA |
| ## | 2905 | NA | NA | NA | NA | NA | NA | 8.542e-01 |
| ## | 2906 | NA | NA | NA | NA | NA | NA | NA |
| ## | 2907 | NA | NA | NA | NA | NA | NA | 8.034e-01 |
| ## | 2908 | NA | NA | NA | NA | NA | NA | NA |
| ## | 2909 | 1.548e-01 | 3.947e-01 | 3.931e-01 | 4.983e-01 | 3.854e-01 | 9.706e-01 | 1.997e-01 |

| | | | | | | | |
|---------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| ## 2910 | NA | NA | NA | NA | NA | NA | NA |
| ## 2911 | NA | NA | NA | NA | NA | NA | 1.999e-01 |
| ## 2912 | NA | NA | NA | NA | NA | NA | NA |
| ## 2913 | 4.590e-01 | 5.347e-01 | 4.186e-03 | 2.638e-03 | 2.965e-02 | 4.322e-02 | 5.012e-02 |
| ## 2914 | NA | NA | NA | NA | NA | NA | 1.439e-01 |
| ## 2915 | NA | NA | NA | NA | NA | NA | NA |
| ## 2916 | 6.187e-01 | 7.886e-01 | 8.623e-02 | 2.826e-02 | 2.833e-01 | 4.294e-02 | 8.562e-02 |
| ## 2917 | 8.934e-01 | 1.606e-09 | 3.550e-06 | 1.122e-04 | 4.118e-02 | 5.890e-04 | 6.697e-01 |
| ## 2918 | NA | NA | NA | NA | NA | NA | NA |
| ## 2919 | 5.580e-02 | 7.044e-02 | 5.294e-01 | 6.374e-03 | 9.147e-03 | 1.938e-07 | 3.704e-01 |
| ## 2920 | NA | NA | NA | NA | NA | NA | NA |
| ## 2921 | NA | NA | NA | NA | NA | NA | 5.961e-01 |
| ## 2922 | NA | NA | NA | NA | NA | NA | NA |
| ## 2923 | NA | NA | NA | NA | NA | NA | NA |
| ## 3000 | NA | NA | NA | NA | NA | NA | NA |
| ## 3001 | NA | NA | NA | NA | NA | NA | NA |
| ## 3002 | 8.770e-01 | 6.131e-01 | 3.051e-01 | 3.263e-01 | 1.855e-01 | 1.520e-01 | 6.669e-01 |
| ## 3003 | 6.195e-01 | 4.914e-01 | 3.666e-03 | 5.918e-03 | 7.924e-03 | 4.134e-03 | 4.473e-01 |
| ## 3004 | 1.641e-01 | 6.647e-01 | 9.008e-04 | 1.875e-03 | 2.172e-04 | 2.248e-05 | 9.544e-01 |
| ## 3005 | 1.229e-01 | 9.953e-01 | 8.445e-02 | 7.152e-01 | 1.531e-03 | 4.236e-01 | 4.994e-02 |
| ## 3100 | 8.654e-01 | 9.004e-01 | 3.938e-01 | 1.955e-01 | 1.449e-01 | 2.843e-03 | 7.455e-01 |
| ## 3101 | 8.091e-01 | 7.709e-01 | 9.891e-01 | 1.459e-01 | 6.345e-01 | 2.163e-02 | 6.898e-01 |
| ## 3102 | NA | NA | NA | NA | NA | NA | 2.670e-01 |
| ## 3103 | 6.181e-01 | 2.422e-01 | 2.845e-01 | 8.384e-01 | 7.696e-02 | 6.164e-01 | 7.078e-01 |
| ## 3104 | 9.091e-02 | 8.026e-01 | 2.148e-01 | 1.261e-01 | 3.806e-02 | 4.424e-02 | 3.418e-02 |
| ## 3105 | 8.706e-01 | 1.600e-01 | 7.817e-01 | 2.091e-01 | 5.297e-01 | 1.799e-01 | 8.376e-01 |
| ## 3106 | 1.426e-01 | 1.036e-01 | 2.927e-02 | 1.133e-02 | 3.553e-01 | 4.681e-01 | 4.045e-01 |
| ## 3107 | 4.591e-01 | 3.388e-01 | 1.462e-02 | 5.390e-04 | 8.949e-03 | 7.866e-03 | 6.987e-01 |
| ## 3108 | NA | NA | NA | NA | NA | NA | NA |
| ## 3109 | NA | NA | NA | NA | NA | NA | NA |
| ## 3110 | 3.347e-01 | 1.035e-01 | 1.279e-01 | 1.171e-01 | 1.023e-01 | 8.899e-02 | 4.150e-01 |
| ## 3200 | 2.094e-01 | 6.468e-01 | 6.944e-03 | 6.004e-01 | 6.649e-04 | 1.754e-01 | 1.415e-01 |
| ## 3201 | NA | NA | NA | NA | NA | NA | NA |
| ## 3202 | 7.601e-01 | 7.283e-01 | 7.179e-01 | 1.508e-01 | 8.645e-01 | 3.135e-01 | 8.969e-01 |
| ## 3203 | 2.067e-01 | 9.621e-01 | 8.220e-02 | 1.644e-02 | 2.939e-03 | 7.210e-03 | 1.172e-01 |
| ## 3204 | 5.124e-01 | 3.010e-01 | 6.706e-01 | 1.863e-03 | 4.581e-02 | 5.149e-01 | 6.817e-01 |
| ## 3205 | 3.771e-01 | 9.287e-01 | 5.565e-01 | 2.960e-01 | 2.745e-03 | 1.722e-01 | 1.120e-01 |
| ## 3206 | 1.311e-01 | 3.970e-05 | 6.001e-02 | 9.131e-02 | 6.143e-04 | 4.480e-01 | 8.164e-02 |
| ## 3207 | 7.525e-02 | 4.339e-02 | 5.922e-02 | 1.135e-01 | 6.390e-01 | 5.423e-01 | 1.252e-01 |
| ## 3300 | 2.453e-01 | 3.940e-01 | 5.887e-01 | 6.274e-01 | 6.830e-01 | 6.624e-03 | 2.938e-01 |
| ## 3301 | 6.247e-01 | 4.606e-01 | 5.973e-02 | 5.667e-02 | 2.257e-01 | 9.107e-01 | 6.076e-01 |
| ## 3302 | 3.696e-02 | 6.240e-02 | 1.540e-01 | 1.451e-01 | 8.149e-01 | 3.098e-01 | 5.136e-01 |
| ## 3303 | 8.279e-01 | 8.961e-01 | 5.334e-01 | 3.409e-01 | 2.736e-03 | 1.540e-02 | 9.687e-01 |
| ## 3304 | 1.419e-01 | 2.822e-01 | 1.347e-03 | 1.945e-03 | 1.142e-02 | 1.482e-02 | 5.902e-02 |
| ## 3305 | 7.131e-01 | 4.644e-01 | 3.386e-04 | 2.595e-02 | 4.072e-02 | 7.125e-01 | 7.547e-01 |
| ## 3306 | 9.410e-02 | 1.483e-01 | 5.022e-01 | 2.961e-01 | 7.056e-01 | 3.498e-02 | 1.531e-01 |
| ## 3307 | NA | NA | NA | NA | NA | NA | NA |
| ## 3308 | 4.582e-01 | 3.699e-01 | 7.805e-03 | 8.235e-01 | 2.564e-01 | 9.262e-01 | 9.582e-01 |
| ## 3309 | 1.536e-01 | 1.064e-01 | 3.783e-01 | 2.661e-01 | 9.668e-01 | 2.440e-01 | 9.992e-02 |
| ## 3310 | 4.167e-01 | 8.003e-02 | 2.006e-02 | 6.864e-02 | 1.849e-03 | 9.678e-01 | 5.158e-01 |

| | | | | | | | |
|---------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| ## 3311 | NA | NA | NA | NA | NA | NA | NA |
| ## 3312 | 7.883e-01 | 7.909e-01 | 4.345e-02 | 7.248e-01 | 5.824e-01 | 9.637e-01 | 7.365e-01 |
| ## 3313 | 5.432e-01 | 1.229e-01 | 5.433e-01 | 7.709e-01 | 8.397e-01 | 2.910e-01 | 5.487e-01 |
| ## 3314 | 5.737e-06 | 4.023e-03 | 3.837e-02 | 3.170e-05 | 0.000e+00 | 1.527e-01 | 8.986e-55 |
| ## 3315 | 1.023e-01 | 1.531e-01 | 3.587e-01 | 7.052e-02 | 7.953e-01 | 1.537e-02 | 5.810e-02 |
| ## 3316 | 8.913e-01 | 6.337e-01 | 4.025e-02 | 4.294e-01 | 3.095e-02 | 3.759e-01 | 5.358e-01 |
| ## 3317 | 3.283e-01 | 8.434e-02 | 3.083e-01 | 9.789e-01 | 8.117e-02 | 1.385e-03 | 5.101e-01 |
| ## 3318 | 7.332e-02 | 1.251e-01 | 1.759e-03 | 3.998e-01 | 6.444e-01 | 5.623e-08 | 2.746e-01 |
| ## 3319 | NA | NA | NA | NA | NA | NA | NA |
| ## 3320 | 1.393e-01 | 6.378e-02 | 2.680e-01 | 7.257e-01 | 5.360e-01 | 1.137e-02 | 1.116e-01 |
| ## 3321 | 3.957e-01 | 8.895e-01 | 6.864e-01 | 7.495e-01 | 4.048e-01 | 6.420e-01 | 3.441e-01 |
| ## 3322 | 2.635e-01 | 1.633e-01 | 4.506e-01 | 3.770e-01 | 6.534e-03 | 5.599e-01 | 3.478e-01 |
| ## 3400 | 8.129e-01 | 7.556e-01 | 5.283e-02 | 4.583e-05 | 3.611e-04 | 1.255e-06 | 5.362e-01 |
| ## 3401 | NA | NA | NA | NA | NA | NA | NA |
| ## 3402 | NA | NA | NA | NA | NA | NA | NA |
| ## 3403 | NA | NA | NA | NA | NA | NA | NA |
| ## 3404 | NA | NA | NA | NA | NA | NA | NA |
| ## 3500 | 6.787e-01 | 7.471e-03 | 5.824e-03 | 7.331e-03 | 1.500e-01 | 5.769e-01 | 9.240e-01 |
| ## 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 | NA | NA | NA | NA | 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 | 2.496e-01 |
| ## 3602 | NA | NA | NA | NA | NA | NA | NA |
| ## 3603 | NA | NA | NA | NA | 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 | NA | NA | NA | NA | NA | NA | NA |
| ## 3611 | NA | NA | NA | NA | NA | NA | NA |
| ## 3612 | 8.233e-01 | 2.248e-01 | 1.805e-04 | 7.600e-04 | 2.229e-02 | 1.756e-03 | 7.285e-01 |
| ## 3613 | NA | NA | NA | NA | NA | NA | NA |
| ## 3614 | 1.204e-01 | 5.697e-02 | 5.224e-01 | 5.248e-03 | 8.013e-02 | 3.504e-03 | 6.116e-01 |
| ## 3615 | NA | NA | NA | NA | NA | NA | NA |
| ## 3616 | 8.489e-01 | 1.679e-01 | 6.468e-02 | 3.287e-01 | 5.590e-01 | 3.390e-04 | 7.214e-01 |
| ## | FLA2p | FFA3p | FLA4p | | | | |
| ## 1000 | NA | NA | NA | | | | |
| ## 1100 | 8.131e-01 | 3.976e-01 | 8.603e-01 | | | | |
| ## 1101 | 3.414e-01 | 1.181e-01 | NA | | | | |
| ## 1102 | 3.662e-01 | 5.251e-01 | 3.100e-01 | | | | |
| ## 1103 | 1.912e-01 | 8.143e-01 | 1.962e-01 | | | | |
| ## 1104 | 7.949e-01 | 7.870e-02 | 9.340e-01 | | | | |
| ## 1105 | 2.447e-01 | 3.448e-01 | 3.747e-01 | | | | |
| ## 1106 | 4.548e-01 | 2.738e-01 | 5.890e-01 | | | | |


```
## 1107 1.224e-01 5.596e-01 1.032e-01
## 1108      NA      NA      NA
## 1109 1.367e-01 8.822e-03 9.132e-02
## 1110 2.353e-01 3.923e-01 2.760e-01
## 1111 7.566e-02 7.722e-01 7.878e-02
## 1200 2.471e-01 1.406e-01 6.076e-02
## 1201 4.275e-01 6.158e-01 7.052e-01
## 1202 1.374e-01 6.466e-01 2.938e-01
## 1203 1.016e-01 4.644e-01 1.321e-01
## 1204 7.082e-01 6.228e-01 5.024e-01
## 1205      NA      NA      NA
## 1206      NA      NA      NA
## 1207 2.317e-01 6.127e-02 1.124e-01
## 1208 9.012e-01 8.680e-01 8.956e-01
## 1209      NA      NA      NA
## 1210 1.302e-02 2.072e-01 8.462e-01
## 1211 6.654e-01 1.138e-02 2.274e-02
## 1212 2.225e-05 1.710e-01 2.752e-01
## 1213      NA 6.958e-01 6.958e-01
## 1300 5.117e-01 1.701e-02 6.972e-01
## 1301      NA      NA      NA
## 1302 5.404e-01 8.466e-02 9.553e-01
## 1303 8.894e-01 1.018e-01 7.167e-01
## 1304 7.512e-01 6.949e-01 6.984e-01
## 1305 4.959e-01 5.639e-01 4.845e-01
## 1306 4.768e-02 8.911e-02 1.981e-02
## 1307 6.200e-01 2.992e-01 5.292e-01
## 1308 1.618e-01 2.670e-01 2.143e-01
## 1309 3.152e-02 5.434e-01 3.349e-02
## 1310 8.014e-01 6.461e-01 7.053e-01
## 1311 8.560e-01 1.548e-01 7.058e-01
## 1312 5.512e-01 6.936e-01 5.256e-01
## 1313 8.799e-02 9.660e-01 9.065e-02
## 1314 2.080e-01 4.487e-01 1.605e-01
## 1315 3.818e-01 2.854e-01 8.317e-01
## 1400 4.403e-01 9.624e-01 4.956e-01
## 1401 8.732e-01 8.711e-03 4.243e-02
## 1402 2.469e-01 9.550e-01 3.881e-01
## 1403 9.179e-01 6.154e-01 8.348e-01
## 1404 7.128e-01 7.987e-05 3.045e-02
## 1405 8.027e-01 7.984e-01 9.453e-01
## 1406 4.270e-01 1.807e-01 1.756e-01
## 1407 8.281e-01 9.073e-01 8.279e-01
## 1408 5.574e-01 1.979e-02 3.150e-01
## 1409 2.284e-01 8.121e-01 4.288e-01
## 1410      NA      NA      NA
## 1500 7.416e-01 1.037e-01 5.525e-01
## 1501      NA      NA      NA
## 1502 9.964e-01 2.903e-01 9.631e-01
## 1503 1.760e-01 7.195e-01 1.910e-01
```

| | | | |
|---------|-----------|-----------|-----------|
| ## 1504 | NA | NA | NA |
| ## 1505 | 8.526e-01 | 2.675e-01 | 9.966e-01 |
| ## 1506 | NA | NA | NA |
| ## 1507 | 6.745e-01 | 2.463e-02 | 3.939e-01 |
| ## 1508 | 5.946e-01 | 1.727e-01 | 1.487e-01 |
| ## 1600 | 1.626e-04 | 2.682e-02 | 4.650e-04 |
| ## 1601 | NA | NA | NA |
| ## 1602 | 6.237e-01 | 4.370e-01 | 6.670e-01 |
| ## 1603 | 7.774e-01 | 7.739e-01 | 7.608e-01 |
| ## 1604 | 3.302e-01 | 3.330e-01 | 2.949e-01 |
| ## 1605 | 7.678e-01 | 4.868e-01 | 7.245e-01 |
| ## 1606 | 7.245e-02 | 5.498e-01 | 8.199e-02 |
| ## 1607 | 7.814e-01 | 8.373e-01 | 8.011e-01 |
| ## 1700 | 6.664e-01 | 3.756e-04 | 1.577e-01 |
| ## 1701 | NA | NA | NA |
| ## 1702 | 7.391e-02 | 3.084e-01 | NA |
| ## 1703 | 6.022e-01 | 1.581e-01 | 7.499e-01 |
| ## 1704 | 4.255e-01 | 8.494e-01 | 5.494e-01 |
| ## 1705 | 2.382e-01 | 1.679e-01 | 5.813e-02 |
| ## 1706 | 6.604e-01 | 2.067e-01 | 9.092e-01 |
| ## 1707 | 1.007e-02 | 1.332e-01 | 2.015e-02 |
| ## 1708 | 9.756e-02 | 3.198e-01 | 2.548e-01 |
| ## 1709 | 2.509e-01 | 8.723e-01 | 2.353e-01 |
| ## 1710 | 4.730e-01 | 2.317e-02 | 1.002e-01 |
| ## 1711 | 9.611e-01 | 3.151e-01 | 9.443e-01 |
| ## 1712 | 9.497e-01 | 2.997e-01 | 8.943e-01 |
| ## 1800 | NA | NA | NA |
| ## 1801 | NA | NA | NA |
| ## 1802 | 3.896e-01 | 1.397e-01 | 1.493e-01 |
| ## 1803 | 9.331e-02 | 1.851e-01 | 6.039e-02 |
| ## 1804 | 1.121e-03 | 7.880e-01 | 5.749e-04 |
| ## 1900 | 5.151e-01 | 2.119e-01 | 6.748e-01 |
| ## 1901 | 1.822e-01 | 8.517e-02 | 8.693e-01 |
| ## 1902 | 6.716e-03 | 2.257e-02 | 1.445e-05 |
| ## 1903 | NA | NA | NA |
| ## 1904 | 6.153e-03 | 6.156e-01 | 5.496e-02 |
| ## 1905 | NA | NA | NA |
| ## 1906 | 5.716e-01 | 3.333e-01 | 5.691e-01 |
| ## 1907 | 3.527e-01 | 3.919e-01 | 3.976e-01 |
| ## 1908 | 8.520e-02 | 7.390e-02 | 2.378e-01 |
| ## 1909 | 3.866e-02 | 4.946e-01 | NA |
| ## 1910 | 9.591e-01 | 5.418e-02 | 4.937e-01 |
| ## 1911 | 3.794e-02 | 6.400e-01 | 6.565e-02 |
| ## 1912 | 3.927e-01 | 7.514e-01 | 4.342e-01 |
| ## 1913 | NA | NA | NA |
| ## 2000 | 7.829e-01 | NA | NA |
| ## 2001 | 1.999e-06 | 5.627e-01 | 4.960e-06 |
| ## 2002 | 7.513e-01 | 5.161e-01 | 8.676e-01 |
| ## 2003 | 8.342e-01 | 2.926e-01 | 5.814e-01 |
| ## 2100 | 6.044e-01 | 6.908e-01 | 6.825e-01 |

| | | | |
|---------|-----------|-----------|-----------|
| ## 2101 | NA | NA | NA |
| ## 2102 | 4.592e-01 | 2.804e-01 | 4.497e-01 |
| ## 2103 | NA | NA | NA |
| ## 2104 | NA | NA | NA |
| ## 2105 | 1.867e-01 | 9.727e-01 | 2.062e-01 |
| ## 2200 | 9.262e-01 | 1.248e-01 | 9.226e-01 |
| ## 2201 | 5.024e-01 | 5.866e-01 | 4.672e-01 |
| ## 2202 | 2.961e-01 | 5.625e-01 | 4.345e-01 |
| ## 2203 | NA | NA | NA |
| ## 2204 | 7.336e-01 | 9.583e-01 | 7.411e-01 |
| ## 2205 | 2.096e-02 | 7.234e-01 | 5.852e-02 |
| ## 2206 | NA | NA | NA |
| ## 2207 | 6.362e-01 | 2.036e-01 | 2.604e-01 |
| ## 2208 | 7.911e-01 | 2.225e-01 | 7.594e-01 |
| ## 2209 | 2.791e-01 | 8.743e-01 | 3.018e-01 |
| ## 2210 | 1.633e-01 | 6.096e-01 | 1.578e-01 |
| ## 2211 | 9.610e-01 | 6.132e-01 | 9.817e-01 |
| ## 2212 | NA | NA | NA |
| ## 2213 | 6.201e-02 | 6.249e-01 | 4.010e-02 |
| ## 2214 | 1.161e-02 | 1.712e-05 | 4.621e-01 |
| ## 2215 | 1.107e-01 | 4.178e-02 | 2.871e-01 |
| ## 2216 | NA | NA | NA |
| ## 2300 | 1.438e-01 | 8.058e-01 | 1.604e-01 |
| ## 2301 | 4.896e-01 | 9.988e-02 | 6.132e-01 |
| ## 2302 | NA | NA | NA |
| ## 2303 | 7.355e-01 | 3.518e-01 | 7.749e-01 |
| ## 2304 | 6.294e-02 | 7.612e-01 | 7.318e-02 |
| ## 2305 | 1.525e-01 | 3.465e-01 | 2.294e-01 |
| ## 2306 | NA | NA | NA |
| ## 2307 | 3.479e-01 | 3.408e-01 | 2.581e-01 |
| ## 2308 | 8.983e-01 | 3.279e-01 | 6.486e-01 |
| ## 2309 | 1.112e-01 | 5.922e-01 | 1.087e-01 |
| ## 2310 | 6.691e-02 | 4.077e-01 | 8.442e-02 |
| ## 2311 | 9.162e-01 | 9.924e-01 | 9.297e-01 |
| ## 2312 | 5.962e-02 | 1.714e-01 | 6.778e-02 |
| ## 2400 | 3.509e-01 | 6.065e-01 | 5.116e-01 |
| ## 2401 | NA | NA | NA |
| ## 2402 | 3.606e-01 | 4.967e-01 | 3.900e-01 |
| ## 2403 | 3.980e-01 | 6.206e-02 | 4.439e-01 |
| ## 2404 | 9.451e-01 | 4.420e-01 | 9.981e-01 |
| ## 2405 | 3.897e-01 | 8.586e-01 | 3.844e-01 |
| ## 2406 | 3.633e-01 | 7.012e-01 | 3.336e-01 |
| ## 2500 | 4.474e-02 | 5.990e-04 | 7.934e-02 |
| ## 2501 | NA | NA | NA |
| ## 2502 | 3.395e-01 | 1.909e-02 | 5.228e-01 |
| ## 2503 | 6.517e-01 | 5.618e-01 | 6.250e-01 |
| ## 2504 | 6.881e-01 | 4.492e-01 | 6.518e-01 |
| ## 2505 | 6.584e-01 | 9.775e-03 | 6.848e-01 |
| ## 2506 | 2.097e-01 | 9.285e-01 | 2.073e-01 |
| ## 2507 | 3.862e-01 | 7.207e-01 | 4.610e-01 |

```
## 2508 7.471e-01 1.045e-01 8.142e-01
## 2600 9.057e-01 1.366e-02 1.200e-01
## 2601      NA      NA      NA
## 2602      NA 1.177e-01 2.720e-01
## 2603 3.718e-01 9.694e-01 3.565e-01
## 2604 5.533e-01 9.637e-03 1.147e-01
## 2605 1.846e-01 5.267e-01 1.829e-01
## 2606      NA      NA      NA
## 2607 2.540e-01 1.697e-01 2.144e-01
## 2608 4.312e-02 4.272e-02 2.813e-01
## 2609      NA      NA      NA
## 2610      NA 1.739e-01 1.739e-01
## 2611 3.752e-01 7.044e-01 3.470e-01
## 2612 3.879e-05 1.621e-02 1.726e-05
## 2613 3.991e-01 6.384e-01 3.458e-01
## 2614 7.866e-01 2.030e-02 4.372e-01
## 2700 5.265e-01 2.325e-01 9.927e-01
## 2701 7.124e-01 2.048e-01 4.929e-01
## 2702 5.721e-01 5.097e-01 9.175e-01
## 2703 1.152e-01 5.725e-02 5.350e-02
## 2704 8.942e-01 4.061e-01 7.238e-01
## 2705 4.918e-01 6.615e-03 6.842e-01
## 2706 9.260e-02 9.206e-01 9.802e-02
## 2707      NA      NA      NA
## 2708 3.969e-01 3.201e-01 5.500e-01
## 2709      NA      NA      NA
## 2710      NA      NA      NA
## 2711 5.428e-02 6.838e-01 5.229e-02
## 2712 8.864e-02 7.273e-01 9.991e-02
## 2713 5.979e-01 6.748e-02 9.105e-01
## 2714 4.189e-03 1.000e-02 1.443e-01
## 2715 4.065e-01 4.715e-01 3.518e-01
## 2716 5.060e-02 5.603e-02 7.091e-02
## 2717 5.514e-01 8.552e-01 5.802e-01
## 2718 9.874e-01 9.563e-02 4.545e-01
## 2719 3.897e-01 3.780e-01 7.174e-01
## 2720 8.647e-01 2.165e-01 8.078e-01
## 2721 1.252e-01 2.581e-01 1.040e-01
## 2722 3.260e-01 2.920e-01 4.205e-01
## 2723 1.478e-01 2.483e-01 2.105e-01
## 2724 1.422e-01 3.693e-03 2.159e-02
## 2725 7.661e-01 8.304e-01 8.055e-01
## 2726 1.378e-01 3.172e-01 1.453e-01
## 2727 4.875e-01 4.581e-01 4.354e-01
## 2728 2.822e-01 8.776e-01 2.829e-01
## 2729 8.060e-01 1.153e-01 7.562e-01
## 2730 2.490e-02 2.428e-02 2.839e-03
## 2731      NA      NA 8.882e-01
## 2732 6.432e-01 2.674e-01 3.279e-01
## 2733 2.568e-01 8.410e-01 3.119e-01
```

| | | | | |
|----|------|-----------|-----------|-----------|
| ## | 2734 | 8.738e-01 | 1.736e-02 | 8.002e-01 |
| ## | 2735 | 1.733e-02 | 4.385e-01 | 2.485e-02 |
| ## | 2736 | 4.462e-01 | 8.142e-01 | 4.410e-01 |
| ## | 2737 | 5.716e-01 | 6.844e-01 | 4.859e-01 |
| ## | 2738 | 4.968e-01 | 8.359e-01 | 4.839e-01 |
| ## | 2739 | 5.629e-01 | 2.826e-01 | 3.861e-01 |
| ## | 2740 | 8.899e-01 | 1.317e-04 | 9.170e-01 |
| ## | 2741 | 7.215e-01 | 1.431e-01 | 9.284e-01 |
| ## | 2742 | 6.138e-01 | 5.775e-02 | 2.766e-01 |
| ## | 2743 | 7.419e-01 | 1.392e-01 | 6.923e-01 |
| ## | 2744 | NA | NA | NA |
| ## | 2745 | 6.826e-01 | 3.692e-03 | 8.187e-01 |
| ## | 2746 | 7.521e-02 | 1.247e-01 | 9.275e-02 |
| ## | 2747 | 4.363e-01 | 6.050e-02 | 3.701e-02 |
| ## | 2748 | 1.189e-01 | 4.169e-01 | 2.066e-01 |
| ## | 2800 | 2.756e-02 | 1.069e-01 | 7.127e-02 |
| ## | 2801 | NA | NA | NA |
| ## | 2802 | NA | 4.332e-01 | NA |
| ## | 2803 | 8.655e-01 | 9.892e-01 | 8.641e-01 |
| ## | 2804 | 6.765e-02 | 3.786e-01 | 5.687e-02 |
| ## | 2805 | 3.399e-01 | 3.250e-01 | 6.214e-01 |
| ## | 2806 | NA | NA | NA |
| ## | 2807 | 9.525e-02 | 6.284e-01 | 3.081e-02 |
| ## | 2808 | 3.917e-01 | 9.065e-01 | 3.974e-01 |
| ## | 2809 | 7.340e-01 | 1.563e-01 | 9.392e-01 |
| ## | 2900 | 9.158e-01 | 3.392e-03 | 1.224e-01 |
| ## | 2901 | NA | NA | NA |
| ## | 2902 | 4.204e-01 | 9.330e-01 | 4.937e-01 |
| ## | 2903 | NA | NA | NA |
| ## | 2904 | NA | NA | NA |
| ## | 2905 | 1.674e-01 | 1.685e-01 | 1.605e-02 |
| ## | 2906 | NA | NA | NA |
| ## | 2907 | 2.412e-01 | 7.562e-01 | 2.121e-01 |
| ## | 2908 | NA | NA | NA |
| ## | 2909 | 3.988e-01 | 2.919e-01 | 7.528e-01 |
| ## | 2910 | NA | NA | NA |
| ## | 2911 | 2.166e-01 | 4.537e-01 | 7.955e-01 |
| ## | 2912 | NA | NA | NA |
| ## | 2913 | 3.266e-01 | 3.911e-02 | 9.346e-01 |
| ## | 2914 | 9.467e-01 | 5.996e-02 | 1.319e-01 |
| ## | 2915 | NA | NA | NA |
| ## | 2916 | 7.774e-01 | 8.501e-02 | 6.268e-01 |
| ## | 2917 | 1.666e-04 | 7.347e-02 | NA |
| ## | 2918 | NA | NA | NA |
| ## | 2919 | 1.148e-01 | 7.942e-01 | 1.764e-01 |
| ## | 2920 | NA | NA | NA |
| ## | 2921 | 7.707e-02 | 6.855e-01 | 8.532e-02 |
| ## | 2922 | NA | NA | NA |
| ## | 2923 | NA | NA | NA |
| ## | 3000 | NA | NA | NA |

| | | | |
|---------|-----------|-----------|-----------|
| ## 3001 | NA | NA | NA |
| ## 3002 | 4.913e-01 | 6.325e-01 | 4.732e-01 |
| ## 3003 | 5.317e-01 | 4.314e-01 | 5.363e-01 |
| ## 3004 | 9.271e-01 | 9.565e-01 | 9.245e-01 |
| ## 3005 | 8.942e-01 | 6.815e-02 | 7.024e-01 |
| ## 3100 | 8.338e-01 | 7.171e-01 | 7.889e-01 |
| ## 3101 | 7.764e-01 | 2.588e-01 | 4.236e-01 |
| ## 3102 | 5.182e-01 | 2.693e-01 | 6.013e-01 |
| ## 3103 | 2.825e-01 | 4.201e-01 | 2.434e-01 |
| ## 3104 | 8.269e-01 | 3.524e-02 | 9.787e-01 |
| ## 3105 | 5.269e-02 | 7.550e-01 | 5.016e-02 |
| ## 3106 | 2.657e-01 | 7.620e-01 | 5.327e-01 |
| ## 3107 | 6.397e-01 | 7.040e-01 | 6.427e-01 |
| ## 3108 | NA | NA | NA |
| ## 3109 | NA | NA | NA |
| ## 3110 | 1.147e-01 | 3.896e-01 | 1.119e-01 |
| ## 3200 | 7.632e-01 | 1.796e-01 | 7.553e-01 |
| ## 3201 | NA | NA | NA |
| ## 3202 | 6.821e-01 | 8.142e-01 | 6.491e-01 |
| ## 3203 | 9.985e-01 | 8.190e-02 | 5.194e-01 |
| ## 3204 | 3.902e-01 | 9.352e-01 | 4.400e-01 |
| ## 3205 | 5.046e-01 | 1.321e-01 | 8.855e-01 |
| ## 3206 | 1.643e-03 | 3.748e-01 | 2.595e-03 |
| ## 3207 | 7.947e-03 | 9.555e-01 | 4.053e-02 |
| ## 3300 | 4.784e-01 | 2.955e-01 | 6.318e-01 |
| ## 3301 | 2.883e-01 | 6.052e-01 | 4.110e-01 |
| ## 3302 | 6.315e-01 | 5.655e-01 | 7.421e-01 |
| ## 3303 | 8.909e-01 | 8.763e-01 | 8.448e-01 |
| ## 3304 | 4.924e-01 | 3.879e-03 | 1.925e-02 |
| ## 3305 | 5.436e-01 | 9.212e-01 | 6.344e-01 |
| ## 3306 | 1.569e-01 | 4.767e-01 | 4.746e-01 |
| ## 3307 | NA | NA | NA |
| ## 3308 | 9.994e-02 | 3.893e-02 | 7.500e-03 |
| ## 3309 | 3.874e-02 | 6.228e-01 | 2.753e-01 |
| ## 3310 | 9.244e-02 | 1.231e-01 | 2.514e-02 |
| ## 3311 | NA | NA | NA |
| ## 3312 | 8.097e-01 | 4.451e-01 | 4.540e-01 |
| ## 3313 | 1.866e-01 | NA | 2.275e-01 |
| ## 3314 | 2.020e-03 | 3.511e-01 | 5.127e-01 |
| ## 3315 | 1.072e-01 | 1.679e-01 | 9.081e-01 |
| ## 3316 | 3.530e-01 | 6.455e-01 | 4.067e-01 |
| ## 3317 | 2.000e-01 | 9.899e-01 | 3.758e-01 |
| ## 3318 | 8.710e-02 | NA | 4.679e-01 |
| ## 3319 | NA | NA | NA |
| ## 3320 | 4.422e-02 | 6.216e-01 | 3.005e-01 |
| ## 3321 | 8.533e-01 | 2.440e-01 | 4.834e-01 |
| ## 3322 | 2.324e-01 | 8.747e-01 | 4.281e-01 |
| ## 3400 | 3.541e-01 | 7.674e-01 | 4.402e-01 |
| ## 3401 | NA | NA | NA |
| ## 3402 | NA | NA | NA |

```

## 3403      NA      NA      NA
## 3404      NA      NA      NA
## 3500 4.750e-02 5.268e-01 5.597e-02
## 3501      NA      NA      NA
## 3502      NA      NA      NA
## 3503      NA      NA      NA
## 3504      NA      NA      NA
## 3505      NA      NA      NA
## 3506      NA      NA      NA
## 3600      NA      NA      NA
## 3601 5.688e-02 8.585e-02 3.248e-02
## 3602      NA      NA      NA
## 3603      NA      NA      NA
## 3604      NA      NA      NA
## 3605      NA      NA      NA
## 3606      NA      NA      NA
## 3607      NA      NA      NA
## 3608      NA      NA      NA
## 3609      NA      NA      NA
## 3610      NA      NA      NA
## 3611      NA      NA      NA
## 3612 1.123e-01 2.112e-01 3.494e-02
## 3613      NA      NA      NA
## 3614 5.113e-01 8.045e-01 6.127e-01
## 3615      NA      NA      NA
## 3616 1.593e-01 5.536e-01 1.731e-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 ""      ""      ""      ""      "*"      ""      ""      ""      ""      ""
```


[illegible]

[illegible]

| | | | | | | | | | | |
|----|------|-------|-------|-------|-------|-------|-------|-------|------|-------|
| ## | 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 | " " | " " | "*" | "*" | " " | " " | " " | " " | " " | " " |

[illegible]

```
## 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          23         11          10
## 1100          27         14           5
## 1101           3           0           0
## 1102          12           4           3
## 1103          25          11           9
## 1104          19           7          10
## 1105          19           7           7
## 1106          77          42          28
## 1107           8           3           2
## 1108           1           0           0
## 1109           2           2           0
## 1110          16           2           3
## 1111           4           2           0
## 1200           8           3           3
## 1201          28          18          18
## 1202          63          25          24
## 1203          27          19          19
## 1204          14           6           5
## 1205           7           1           1
## 1206           2           2           0
## 1207          20           6           5
## 1208          40          18          18
## 1209           1           1           0
## 1210           7           3           3
## 1211          42          10          10
## 1212          20           5           6
## 1213          15           8           8
## 1300          41          19           9
## 1301           1           1           0
```

| | | | |
|---------|----|----|----|
| ## 1302 | 8 | 5 | 4 |
| ## 1303 | 37 | 13 | 9 |
| ## 1304 | 14 | 8 | 5 |
| ## 1305 | 32 | 20 | 12 |
| ## 1306 | 17 | 12 | 7 |
| ## 1307 | 30 | 14 | 6 |
| ## 1308 | 9 | 3 | 1 |
| ## 1309 | 8 | 1 | 3 |
| ## 1310 | 5 | 3 | 1 |
| ## 1311 | 39 | 18 | 14 |
| ## 1312 | 44 | 24 | 16 |
| ## 1313 | 17 | 7 | 6 |
| ## 1314 | 25 | 3 | 5 |
| ## 1315 | 2 | 0 | 0 |
| ## 1400 | 9 | 2 | 2 |
| ## 1401 | 9 | 3 | 4 |
| ## 1402 | 6 | 1 | 1 |
| ## 1403 | 25 | 5 | 4 |
| ## 1404 | 4 | 2 | 0 |
| ## 1405 | 19 | 7 | 8 |
| ## 1406 | 8 | 4 | 3 |
| ## 1407 | 18 | 6 | 10 |
| ## 1408 | 34 | 9 | 16 |
| ## 1409 | 4 | 1 | 1 |
| ## 1410 | 3 | 0 | 0 |
| ## 1500 | 22 | 10 | 8 |
| ## 1501 | 2 | 1 | 1 |
| ## 1502 | 13 | 4 | 1 |
| ## 1503 | 11 | 3 | 3 |
| ## 1504 | 0 | 0 | 0 |
| ## 1505 | 3 | 1 | 1 |
| ## 1506 | 0 | 0 | 0 |
| ## 1507 | 2 | 0 | 0 |
| ## 1508 | 6 | 2 | 0 |
| ## 1600 | 54 | 16 | 12 |
| ## 1601 | 2 | 0 | 0 |
| ## 1602 | 12 | 5 | 4 |
| ## 1603 | 11 | 8 | 2 |
| ## 1604 | 10 | 5 | 3 |
| ## 1605 | 19 | 6 | 3 |
| ## 1606 | 22 | 5 | 6 |
| ## 1607 | 9 | 4 | 3 |
| ## 1700 | 9 | 4 | 2 |
| ## 1701 | 2 | 0 | 0 |
| ## 1702 | 3 | 0 | 0 |
| ## 1703 | 3 | 0 | 0 |
| ## 1704 | 4 | 1 | 1 |
| ## 1705 | 12 | 4 | 4 |
| ## 1706 | 17 | 4 | 4 |
| ## 1707 | 5 | 2 | 3 |

| | | | |
|---------|----|----|----|
| ## 1708 | 6 | 1 | 1 |
| ## 1709 | 8 | 2 | 1 |
| ## 1710 | 19 | 5 | 5 |
| ## 1711 | 8 | 1 | 2 |
| ## 1712 | 16 | 3 | 2 |
| ## 1800 | 2 | 2 | 0 |
| ## 1801 | 0 | 0 | 0 |
| ## 1802 | 2 | 1 | 2 |
| ## 1803 | 7 | 1 | 1 |
| ## 1804 | 1 | 0 | 0 |
| ## 1900 | 11 | 4 | 5 |
| ## 1901 | 2 | 1 | 1 |
| ## 1902 | 6 | 2 | 3 |
| ## 1903 | 0 | 0 | 0 |
| ## 1904 | 9 | 6 | 4 |
| ## 1905 | 1 | 0 | 0 |
| ## 1906 | 2 | 0 | 0 |
| ## 1907 | 4 | 0 | 0 |
| ## 1908 | 3 | 1 | 0 |
| ## 1909 | 1 | 0 | 0 |
| ## 1910 | 2 | 1 | 1 |
| ## 1911 | 1 | 1 | 0 |
| ## 1912 | 4 | 1 | 2 |
| ## 1913 | 1 | 0 | 0 |
| ## 2000 | 10 | 2 | 1 |
| ## 2001 | 7 | 2 | 3 |
| ## 2002 | 54 | 15 | 13 |
| ## 2003 | 17 | 2 | 2 |
| ## 2100 | 17 | 5 | 7 |
| ## 2101 | 3 | 1 | 0 |
| ## 2102 | 12 | 1 | 1 |
| ## 2103 | 2 | 1 | 1 |
| ## 2104 | 1 | 0 | 0 |
| ## 2105 | 34 | 6 | 6 |
| ## 2200 | 10 | 2 | 1 |
| ## 2201 | 4 | 4 | 0 |
| ## 2202 | 6 | 1 | 0 |
| ## 2203 | 5 | 3 | 2 |
| ## 2204 | 32 | 14 | 7 |
| ## 2205 | 18 | 5 | 4 |
| ## 2206 | 0 | 0 | 0 |
| ## 2207 | 5 | 0 | 0 |
| ## 2208 | 17 | 1 | 1 |
| ## 2209 | 24 | 9 | 4 |
| ## 2210 | 24 | 3 | 2 |
| ## 2211 | 18 | 2 | 1 |
| ## 2212 | 9 | 1 | 2 |
| ## 2213 | 9 | 6 | 1 |
| ## 2214 | 1 | 0 | 0 |
| ## 2215 | 9 | 1 | 1 |

| | | | |
|---------|----|----|----|
| ## 2216 | 3 | 0 | 0 |
| ## 2300 | 15 | 6 | 3 |
| ## 2301 | 9 | 1 | 1 |
| ## 2302 | 1 | 0 | 0 |
| ## 2303 | 13 | 8 | 3 |
| ## 2304 | 7 | 2 | 2 |
| ## 2305 | 8 | 2 | 0 |
| ## 2306 | 7 | 4 | 2 |
| ## 2307 | 15 | 11 | 7 |
| ## 2308 | 33 | 12 | 13 |
| ## 2309 | 8 | 2 | 2 |
| ## 2310 | 9 | 6 | 3 |
| ## 2311 | 5 | 2 | 0 |
| ## 2312 | 10 | 1 | 2 |
| ## 2400 | 8 | 4 | 2 |
| ## 2401 | 0 | 0 | 0 |
| ## 2402 | 14 | 6 | 3 |
| ## 2403 | 25 | 10 | 8 |
| ## 2404 | 29 | 10 | 12 |
| ## 2405 | 12 | 4 | 5 |
| ## 2406 | 3 | 2 | 0 |
| ## 2500 | 47 | 16 | 9 |
| ## 2501 | 0 | 0 | 0 |
| ## 2502 | 21 | 10 | 2 |
| ## 2503 | 7 | 2 | 0 |
| ## 2504 | 24 | 4 | 3 |
| ## 2505 | 6 | 2 | 0 |
| ## 2506 | 7 | 2 | 0 |
| ## 2507 | 9 | 3 | 1 |
| ## 2508 | 16 | 3 | 1 |
| ## 2600 | 11 | 1 | 2 |
| ## 2601 | 5 | 3 | 1 |
| ## 2602 | 14 | 2 | 2 |
| ## 2603 | 9 | 0 | 0 |
| ## 2604 | 25 | 5 | 3 |
| ## 2605 | 5 | 0 | 0 |
| ## 2606 | 4 | 0 | 0 |
| ## 2607 | 8 | 2 | 1 |
| ## 2608 | 4 | 1 | 1 |
| ## 2609 | 0 | 0 | 0 |
| ## 2610 | 3 | 0 | 0 |
| ## 2611 | 8 | 1 | 1 |
| ## 2612 | 1 | 1 | 0 |
| ## 2613 | 9 | 2 | 1 |
| ## 2614 | 4 | 1 | 0 |
| ## 2700 | 85 | 41 | 28 |
| ## 2701 | 37 | 27 | 14 |
| ## 2702 | 4 | 3 | 1 |
| ## 2703 | 10 | 5 | 4 |
| ## 2704 | 0 | 0 | 0 |

| | | | |
|---------|----|----|----|
| ## 2705 | 29 | 10 | 5 |
| ## 2706 | 3 | 1 | 1 |
| ## 2707 | 3 | 2 | 0 |
| ## 2708 | 10 | 9 | 1 |
| ## 2709 | 1 | 1 | 0 |
| ## 2710 | 2 | 1 | 1 |
| ## 2711 | 4 | 1 | 1 |
| ## 2712 | 14 | 9 | 6 |
| ## 2713 | 7 | 5 | 4 |
| ## 2714 | 9 | 5 | 4 |
| ## 2715 | 14 | 7 | 3 |
| ## 2716 | 8 | 4 | 3 |
| ## 2717 | 20 | 12 | 8 |
| ## 2718 | 8 | 4 | 2 |
| ## 2719 | 35 | 22 | 18 |
| ## 2720 | 4 | 3 | 1 |
| ## 2721 | 6 | 2 | 1 |
| ## 2722 | 4 | 3 | 3 |
| ## 2723 | 15 | 12 | 4 |
| ## 2724 | 7 | 3 | 3 |
| ## 2725 | 27 | 14 | 13 |
| ## 2726 | 22 | 11 | 5 |
| ## 2727 | 12 | 4 | 4 |
| ## 2728 | 37 | 19 | 12 |
| ## 2729 | 20 | 17 | 12 |
| ## 2730 | 35 | 24 | 19 |
| ## 2731 | 9 | 8 | 4 |
| ## 2732 | 60 | 21 | 11 |
| ## 2733 | 9 | 5 | 5 |
| ## 2734 | 18 | 10 | 7 |
| ## 2735 | 37 | 24 | 13 |
| ## 2736 | 23 | 15 | 9 |
| ## 2737 | 16 | 3 | 5 |
| ## 2738 | 65 | 36 | 30 |
| ## 2739 | 55 | 40 | 31 |
| ## 2740 | 6 | 3 | 1 |
| ## 2741 | 26 | 9 | 10 |
| ## 2742 | 12 | 8 | 6 |
| ## 2743 | 3 | 2 | 1 |
| ## 2744 | 0 | 0 | 0 |
| ## 2745 | 9 | 6 | 4 |
| ## 2746 | 49 | 18 | 9 |
| ## 2747 | 5 | 2 | 0 |
| ## 2748 | 7 | 2 | 2 |
| ## 2800 | 24 | 12 | 10 |
| ## 2801 | 4 | 2 | 2 |
| ## 2802 | 14 | 8 | 7 |
| ## 2803 | 11 | 8 | 6 |
| ## 2804 | 11 | 6 | 5 |
| ## 2805 | 7 | 3 | 4 |

| | | | |
|---------|----|----|----|
| ## 2806 | 0 | 0 | 0 |
| ## 2807 | 4 | 1 | 2 |
| ## 2808 | 23 | 15 | 10 |
| ## 2809 | 6 | 5 | 3 |
| ## 2900 | 29 | 22 | 18 |
| ## 2901 | 6 | 4 | 5 |
| ## 2902 | 6 | 4 | 3 |
| ## 2903 | 1 | 1 | 0 |
| ## 2904 | 1 | 1 | 0 |
| ## 2905 | 7 | 4 | 5 |
| ## 2906 | 0 | 0 | 0 |
| ## 2907 | 6 | 3 | 1 |
| ## 2908 | 0 | 0 | 0 |
| ## 2909 | 4 | 2 | 0 |
| ## 2910 | 5 | 4 | 3 |
| ## 2911 | 2 | 2 | 0 |
| ## 2912 | 4 | 3 | 3 |
| ## 2913 | 14 | 13 | 12 |
| ## 2914 | 2 | 2 | 0 |
| ## 2915 | 0 | 0 | 0 |
| ## 2916 | 25 | 21 | 12 |
| ## 2917 | 3 | 3 | 0 |
| ## 2918 | 0 | 0 | 0 |
| ## 2919 | 11 | 6 | 7 |
| ## 2920 | 0 | 0 | 0 |
| ## 2921 | 16 | 10 | 12 |
| ## 2922 | 3 | 1 | 2 |
| ## 2923 | 0 | 0 | 0 |
| ## 3000 | 0 | 0 | 0 |
| ## 3001 | 1 | 1 | 0 |
| ## 3002 | 16 | 5 | 7 |
| ## 3003 | 29 | 16 | 16 |
| ## 3004 | 17 | 10 | 7 |
| ## 3005 | 8 | 5 | 3 |
| ## 3100 | 10 | 0 | 0 |
| ## 3101 | 5 | 0 | 0 |
| ## 3102 | 5 | 2 | 1 |
| ## 3103 | 6 | 2 | 2 |
| ## 3104 | 53 | 10 | 7 |
| ## 3105 | 10 | 5 | 1 |
| ## 3106 | 2 | 0 | 0 |
| ## 3107 | 40 | 6 | 6 |
| ## 3108 | 1 | 0 | 0 |
| ## 3109 | 3 | 0 | 0 |
| ## 3110 | 5 | 2 | 1 |
| ## 3200 | 16 | 11 | 6 |
| ## 3201 | 3 | 1 | 2 |
| ## 3202 | 25 | 13 | 12 |
| ## 3203 | 29 | 15 | 15 |
| ## 3204 | 47 | 25 | 27 |

| | | | |
|---------|-----|----|----|
| ## 3205 | 18 | 9 | 9 |
| ## 3206 | 4 | 1 | 3 |
| ## 3207 | 17 | 9 | 7 |
| ## 3300 | 19 | 10 | 8 |
| ## 3301 | 27 | 19 | 14 |
| ## 3302 | 16 | 8 | 8 |
| ## 3303 | 19 | 7 | 5 |
| ## 3304 | 157 | 94 | 79 |
| ## 3305 | 63 | 24 | 22 |
| ## 3306 | 36 | 21 | 20 |
| ## 3307 | 6 | 4 | 1 |
| ## 3308 | 35 | 18 | 16 |
| ## 3309 | 13 | 6 | 8 |
| ## 3310 | 32 | 22 | 21 |
| ## 3311 | 2 | 1 | 1 |
| ## 3312 | 117 | 50 | 49 |
| ## 3313 | 7 | 2 | 3 |
| ## 3314 | 17 | 9 | 8 |
| ## 3315 | 23 | 13 | 14 |
| ## 3316 | 79 | 41 | 41 |
| ## 3317 | 3 | 2 | 1 |
| ## 3318 | 12 | 12 | 0 |
| ## 3319 | 6 | 4 | 3 |
| ## 3320 | 45 | 11 | 10 |
| ## 3321 | 15 | 6 | 7 |
| ## 3322 | 13 | 3 | 3 |
| ## 3400 | 8 | 3 | 6 |
| ## 3401 | 1 | 0 | 0 |
| ## 3402 | 2 | 1 | 1 |
| ## 3403 | 6 | 4 | 1 |
| ## 3404 | 2 | 1 | 1 |
| ## 3500 | 0 | 0 | 0 |
| ## 3501 | 0 | 0 | 0 |
| ## 3502 | 0 | 0 | 0 |
| ## 3503 | 0 | 0 | 0 |
| ## 3504 | 4 | 3 | 0 |
| ## 3505 | 0 | 0 | 0 |
| ## 3506 | 2 | 2 | 0 |
| ## 3600 | 0 | 0 | 0 |
| ## 3601 | 3 | 3 | 0 |
| ## 3602 | 3 | 3 | 0 |
| ## 3603 | 2 | 1 | 1 |
| ## 3604 | 0 | 0 | 0 |
| ## 3605 | 1 | 0 | 0 |
| ## 3606 | 0 | 0 | 0 |
| ## 3607 | 1 | 0 | 0 |
| ## 3608 | 0 | 0 | 0 |
| ## 3609 | 2 | 2 | 0 |
| ## 3610 | 4 | 4 | 0 |
| ## 3611 | 0 | 0 | 0 |

```
## 3612      41      19      16
## 3613       0       0       0
## 3614       1       1       0
## 3615       0       0       0
## 3616      10       6       9
```

```
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=""))
```