Problem Set 2: Omitted Variable Bias Key

Claire Duquennois

***Group Member 1:*** Scott Heyman

***Group Member 2:*** Xiang Li

***Group Member 3:*** Aileen Wang

# Empirical Analysis using Data from Washington (2008, AER)

This exercise uses data from Ebonya Washington’s paper, “Female Socialization: How Daughters Affect their Legislator Father’s voting on Women’s Issues,” published in the *American Economic Review* in 2008. This paper studies whether having a daughter affects legislator’s voting on women’s issues.

# Finding the data

The data can be found by following the link on the AER’s website which will take you to the ICPSR’s data repository. You will need to sign in to get access to the data files. Once logged in, you will find the set of files that are typically included in a replication file. These include several datasets, several .do files (which is a STATA command file), and text files with the data descriptions which tell you about the different variables included in the dataset. For this assignment we will be using the basic.dta file.

Download it and save it in a `data’ folder located in the same folder as your project repository. Since some datasets in this course will be big, we want to avoid keeping the data on github so I would recommend not placing the data in the project repository itself.

# Set up and opening the data

Because this is a .dta file, you will need to open it with the read.dta function that is included in the haven packages.

Other packages you will need: dplyr, stargazer and lfe.

Remember, if you have not used a package before you will need to install the package as follows

#install.packages('haven',repos = "http://cran.us.r-project.org")  
#install.packages("dplyr",repos = "http://cran.us.r-project.org")  
#install.packages("stargazer",repos = "http://cran.us.r-project.org")  
#install.packages("lfe",repos = "http://cran.us.r-project.org")

Hint: Once you have run these once, on your machine, you may want to comment them out with a # so that your code runs faster.

This .Rmd file will be opened on different computers. But you don’t want to have to change the filepaths each time you pull a new version off of GitHub. Because of this, I would recommend you avoid using any computer specific filepaths in your code. Instead, make sure you and your groupmates structure your project folders in the same way and only specify filepaths within your project folder. R uses the folder where you are saving your code as it’s default “working directory” (where things will be saved or be searched for unless specified otherwise). You can move up to the parent folder by using .. in the file path. Thus, if your data is not saved in the forked github repository but is saved in a folder called data next to it you can call your data with the following file path: "../data/basic.dta".

## 3.1 Question: Now that the packages are installed, call all your packages and load your data. How many observations are in the original dataset?

**Code:**

basic <- read\_dta("basic.dta")  
count(basic)

## # A tibble: 1 × 1  
## n  
## <int>  
## 1 1740

**Answer:**  
There are 1740 observations in the original dataset.

# Cleaning the data

## 4.1 Question: The original dataset contains data from the 105th to 108th U.S. Congress. We only use the observations from the 105th congress. Refer to the data documentation to find the relevant variable and then use the filter function in the dplyr package to extract observations from the 105th congress.

**Code:**

basic\_105 <- filter(basic, congress == "105")  
count(basic\_105)

## # A tibble: 1 × 1  
## n  
## <int>  
## 1 435

## 

## 4.2 Question:The dataset contains many variables, some of which are not used in this exercise. Keep the following variables in the final dataset (Hint: use the select function in dplyr).

| Name | Description |
| --- | --- |
| aauw | AAUW score |
| totchi | Total number of children |
| ngirls | Number of daughters |
| party | Political party. Democrats if 1, Republicans if 2, and Independent if 3. |
| famale | Female dummy variable |
| white | White dummy variable |
| srvlng | Years of service |
| age | Age |
| demvote | State democratic vote share in most recent presidential election |
| medinc | District median income |
| perf | Female proportion of district voting age population |
| perw | White proportion of total district population |
| perhs | High school graduate proportion of district population age 25 |
| percol | College graduate proportion of district population age 25 |
| perur | Urban proportion of total district population |
| moredef | State proportion who favor more defense spending |
| statabb | State abbreviation |
| district | id for electoral district |
| rgroup | religious group |
| region | region |

**You can find the detailed description of each variable in the original paper. The main variable in this analysis is AAUW, a score created by the American Association of University Women (AAUW). For each congress, AAUW selects pieces of legislation in the areas of education, equality, and reproductive rights. The AAUW keeps track of how each legislator voted on these pieces of legislation and whether their vote aligned with the AAUW’s position. The legislator’s score is equal to the proportion of these votes made in agreement with the AAUW.**

**Code:**

finaldata <- select(basic\_105, aauw,totchi,ngirls,party,female,white,srvlng,age,demvote,medinc,perf,perw,perhs,percol,perur,moredef,statabb,district,rgroup,region)

## 

## 4.3 Question: Make sure your final dataset is a data frame. You can check your data’s format with the command is. If the first element of the returned vector is not “data.frame”, convert your dataset with the function as.data.frame.

**Code:**

is(finaldata)

## [1] "tbl\_df" "tbl" "data.frame" "list" "oldClass"   
## [6] "vector"

# Summary Statistics

## 5.1 Question: Report summary statistics of the following variables in the dataset: political party, age, race, gender, AAUW score, the number of children, and the number of daughters. Present these summary statistics in a formatted table, you can use stargazer or other packages. Make this table as communicative as possible.

Hints: If you want RMarkdown to display your outputted table, include the code results = "asis" in the chunk header. This is true for all chunks that output a formatted table. In the stargazer command, you will want to specify the format of the table by including the code results="html" for html output or results="latex" for a pdf output.

**Code:**

summary(finaldata, results = "asis")

## aauw totchi ngirls party   
## Min. : 0.00 Min. : 0.000 Min. :0.000 Min. :1.000   
## 1st Qu.: 0.00 1st Qu.: 2.000 1st Qu.:0.000 1st Qu.:1.000   
## Median : 38.00 Median : 2.000 Median :1.000 Median :2.000   
## Mean : 47.31 Mean : 2.493 Mean :1.274 Mean :1.529   
## 3rd Qu.:100.00 3rd Qu.: 3.000 3rd Qu.:2.000 3rd Qu.:2.000   
## Max. :100.00 Max. :10.000 Max. :7.000 Max. :3.000   
## NA's :1 NA's :1   
## female white srvlng age   
## Min. :0.0000 Min. :0.000 Min. : 1.000 Min. :26.00   
## 1st Qu.:0.0000 1st Qu.:1.000 1st Qu.: 3.000 1st Qu.:45.00   
## Median :0.0000 Median :1.000 Median : 5.000 Median :51.00   
## Mean :0.1103 Mean :0.869 Mean : 8.678 Mean :51.67   
## 3rd Qu.:0.0000 3rd Qu.:1.000 3rd Qu.:13.000 3rd Qu.:58.00   
## Max. :1.0000 Max. :1.000 Max. :47.000 Max. :87.00   
##   
## demvote medinc perf perw   
## Min. :0.2600 Min. :16683 Min. :0.4651 Min. :0.1866   
## 1st Qu.:0.4200 1st Qu.:29593 1st Qu.:0.5129 1st Qu.:0.7563   
## Median :0.4800 Median :34018 Median :0.5212 Median :0.8712   
## Mean :0.5032 Mean :35923 Mean :0.5208 Mean :0.8043   
## 3rd Qu.:0.5700 3rd Qu.:40683 3rd Qu.:0.5297 3rd Qu.:0.9347   
## Max. :0.9400 Max. :64199 Max. :0.5677 Max. :0.9871   
##   
## perhs percol perur moredef   
## Min. :0.3360 Min. :0.0530 Min. :0.1310 Min. :0.0000   
## 1st Qu.:0.6985 1st Qu.:0.1435 1st Qu.:0.5650 1st Qu.:0.1716   
## Median :0.7640 Median :0.1840 Median :0.7897 Median :0.2056   
## Mean :0.7501 Mean :0.2006 Mean :0.7515 Mean :0.2141   
## 3rd Qu.:0.8125 3rd Qu.:0.2420 3rd Qu.:0.9768 3rd Qu.:0.2524   
## Max. :0.9230 Max. :0.5140 Max. :1.0000 Max. :0.5052   
## NA's :26   
## statabb district rgroup region   
## Length:435 Min. : 1.000 Min. :0.000 Length:435   
## Class :character 1st Qu.: 3.000 1st Qu.:1.000 Class :character   
## Mode :character Median : 6.000 Median :1.000 Mode :character   
## Mean : 9.979 Mean :1.529   
## 3rd Qu.:13.000 3rd Qu.:2.000   
## Max. :52.000 Max. :4.000   
##

# 

# Generate Variables

## 6.1 Question:Construct a variable called , a binary set to 1 if the observation is for a republican.

**Code:**

finaldata$repub\_i <- ifelse(finaldata$party == 2, 1, 0)

# Run Estimations

## 7.1 Question: (2 pages) Estimate the following linear regression models using the felm command (part of the lfe package). Report your regression results in a formatted table using a package such as stargazer. Report robust standard errors in your table (Hint: in stargazer specify se = list(model1$rse, model2$rse, model3$rse)). Make this table as informative as possible by adding needed information and removing superfluous information.

**Code:**

reg.1 <- felm(aauw ~ ngirls, finaldata)  
reg.2 <- felm(aauw ~ ngirls+totchi, finaldata)  
reg.3 <- felm(aauw ~ ngirls+totchi+female+repub\_i, finaldata)  
  
stargazer(reg.1, reg.2, reg.3, header = FALSE, type = "text", se = list(reg.1$rse, reg.2$rse, reg.3$rse))

##   
## =========================================================================  
## Dependent variable:   
## -----------------------------------------------------  
## aauw   
## (1) (2) (3)   
## -------------------------------------------------------------------------  
## ngirls -2.784 5.776\*\* 2.825\*\*   
## (1.750) (2.714) (1.306)   
##   
## totchi -7.992\*\*\* -3.149\*\*\*   
## (1.784) (0.964)   
##   
## female 12.577\*\*\*   
## (3.258)   
##   
## repub\_i -71.783\*\*\*   
## (2.100)   
##   
## Constant 50.964\*\*\* 59.982\*\*\* 87.822\*\*\*   
## (3.036) (3.520) (1.809)   
##   
## -------------------------------------------------------------------------  
## Observations 434 434 434   
## R2 0.006 0.051 0.796   
## Adjusted R2 0.003 0.047 0.794   
## Residual Std. Error 41.939 (df = 432) 41.010 (df = 431) 19.055 (df = 429)  
## =========================================================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

reg.1 <- felm(aauw ~ ngirls, finaldata)  
reg.2 <- felm(aauw ~ ngirls|totchi, finaldata)  
reg.3 <- felm(aauw ~ ngirls|totchi+female+repub\_i, finaldata)  
  
stargazer(reg.1, reg.2, reg.3, header = FALSE, type = "text", se = list(reg.1$rse, reg.2$rse, reg.3$rse))

##   
## =========================================================================  
## Dependent variable:   
## -----------------------------------------------------  
## aauw   
## (1) (2) (3)   
## -------------------------------------------------------------------------  
## ngirls -2.784 5.748\*\* 3.043\*\*   
## (1.750) (2.667) (1.359)   
##   
## Constant 50.964\*\*\*   
## (3.036)   
##   
## -------------------------------------------------------------------------  
## Observations 434 434 434   
## R2 0.006 0.065 0.799   
## Adjusted R2 0.003 0.040 0.793   
## Residual Std. Error 41.939 (df = 432) 41.154 (df = 422) 19.122 (df = 420)  
## =========================================================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

## 

## 7.2 Question: (2 pages) Compare the OLS estimates of across the above three specifications. Discuss what explains the difference (if any) of the estimate across three specifications? Which control variable is particularly important and why?

**Code:**

beta1\_model1 <- coef(reg.1)["ngirls"]  
beta1\_model2 <- coef(reg.2)["ngirls"]  
beta1\_model3 <- coef(reg.3)["ngirls"]  
  
c(beta1\_model1, beta1\_model2, beta1\_model3)

## ngirls ngirls ngirls   
## -2.783935 5.748309 3.042570

**Answer:** Control variables are different from the three models, the first model only considered ngirls (number of daughters), the second one considered ngirls and totchi (total number of children), the third one also included these two variables and added female and repub\_i(two dummy variables). If we just look into the first model, it may seem that the number of girls has negative correlation, however, once we run the other models, it becomes clear that there was likely omitted variable bias and that the number of girls is positvely correlated.

repub\_i is important because it has the largest influence and it’s significant.

## 

## 7.3 Question: Consider the third specification (with 3 controls in addition to . Conditional on the number of children and other variables, do you think is plausibly exogenous? What is the identifying assumption necessary for to be interpreted as a causal estimate? What evidence does Washington give to support this assumption?

**Answer:** is plausibly exogenous. Assumption: conditional on number of children, the number of female children is a random variable. According to the data in the appendix, representatives did not follow the rule of stopping childbearing based on preference for children, and voters did not consider the gender mix of children in their selection of representatives.

## 7.4 Question: (2 pages) It is possible that the effects of having daughters might be different for female and male legislators. Estimate four different models to think about this question: the equivalent of model 3 separately on men and women, model 3 with a single interaction term added, and model 3 with three interaction terms added. Present your results in a table. Is there evidence that the effect of a daughter differs for male and female legislators? Of the four models you estimated, which are equivalent, which are different, and why?

**Code:**

data\_male <- filter(finaldata, female == 0)  
data\_female <- filter(finaldata, female == 1)  
  
reg.m <- felm(aauw ~ ngirls+totchi+female+repub\_i, data\_male)

## Warning in chol.default(mat, pivot = TRUE, tol = tol): the matrix is either  
## rank-deficient or not positive definite

reg.f <- felm(aauw ~ ngirls+totchi+female+repub\_i, data\_female)

## Warning in chol.default(mat, pivot = TRUE, tol = tol): the matrix is either  
## rank-deficient or not positive definite

reg.mn <- felm(aauw ~ ngirls+totchi+female+repub\_i+female\*ngirls, finaldata)  
reg.3n <- felm(aauw ~ ngirls+totchi+female+repub\_i+female\*ngirls+totchi\*ngirls+repub\_i\*ngirls, finaldata)  
  
stargazer(reg.m, reg.f, reg.mn,reg.3n, header = FALSE, type = "text", se = list(reg.m$rse, reg.f$rse,reg.mn$rse,reg.3n$rse))

##   
## ==========================================================================================  
## Dependent variable:   
## ----------------------------------------------------------------------  
## aauw   
## (1) (2) (3) (4)   
## ------------------------------------------------------------------------------------------  
## ngirls 3.021\*\* -0.246 2.706\*\* 4.142\*\*   
## (1.302) (6.281) (1.281) (1.721)   
##   
## totchi -3.417\*\*\* 1.222 -3.133\*\*\* -2.328\*\*   
## (0.986) (3.971) (0.958) (1.042)   
##   
## female 11.103\*\* 10.895\*\*   
## (0.000) (0.000) (4.659) (4.796)   
##   
## repub\_i -71.943\*\*\* -70.881\*\*\* -71.791\*\*\* -74.316\*\*\*   
## (2.102) (9.462) (2.098) (3.021)   
##   
## ngirls:female 1.097 1.036   
## (3.295) (3.463)   
##   
## ngirls:totchi -0.704\*   
## (0.363)   
##   
## ngirls:repub\_i 2.166   
## (1.971)   
##   
## Constant 88.337\*\*\* 94.006\*\*\* 87.936\*\*\* 87.111\*\*\*   
## (1.810) (3.441) (1.831) (2.362)   
##   
## ------------------------------------------------------------------------------------------  
## Observations 386 48 434 434   
## R2 0.791 0.729 0.796 0.798   
## Adjusted R2 0.789 0.711 0.794 0.795   
## Residual Std. Error 18.920 (df = 382) 20.365 (df = 44) 19.074 (df = 428) 19.020 (df = 426)  
## ==========================================================================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

**Answer:** It seems as if there is a difference between male and females with males being effected more by having daughters than females. 3 is the same as 1 and 2 if applied to the male and female dataset respectivly.

# Fixed Effects:

## 8.1 Question: (2 pages) Equation 1 from Washington’s paper is a little bit different from the equations you have estimated so far. Estimate the three models specified below (where is a fixed effect for the number of children). Present your results in a table and explain the difference between the three models.

Hint: you will need to generate the dummy variables for the second equation or code it as factor(). For the third equation, the felm function allows you to specify fixed effects.

**Code:**

reg.1 <- felm(aauw ~ ngirls+totchi, finaldata)  
reg.2 <- felm(aauw ~ ngirls+as.factor(totchi), finaldata)  
reg.3 <- felm(aauw ~ ngirls|totchi, finaldata)  
  
stargazer(reg.1, reg.2, reg.3, header = FALSE, type = "text", se = list(reg.1$rse, reg.2$rse, reg.3$rse))

##   
## =========================================================================  
## Dependent variable:   
## -----------------------------------------------------  
## aauw   
## (1) (2) (3)   
## -------------------------------------------------------------------------  
## ngirls 5.776\*\* 5.748\*\* 5.748\*\*   
## (2.714) (2.667) (2.667)   
##   
## totchi -7.992\*\*\*   
## (1.784)   
##   
## as.factor(totchi)1 7.616   
## (8.816)   
##   
## as.factor(totchi)2 -6.182   
## (7.074)   
##   
## as.factor(totchi)3 -17.186\*\*   
## (7.770)   
##   
## as.factor(totchi)4 -25.833\*\*\*   
## (9.090)   
##   
## as.factor(totchi)5 -28.128\*\*   
## (11.601)   
##   
## as.factor(totchi)6 -34.712   
## (24.334)   
##   
## as.factor(totchi)7 -65.986\*\*\*   
## (11.828)   
##   
## as.factor(totchi)8 -74.859\*\*\*   
## (15.283)   
##   
## as.factor(totchi)9 -81.108\*\*\*   
## (14.386)   
##   
## as.factor(totchi)10 -75.360\*\*\*   
## (11.957)   
##   
## Constant 59.982\*\*\* 52.367\*\*\*   
## (3.520) (5.400)   
##   
## -------------------------------------------------------------------------  
## Observations 434 434 434   
## R2 0.051 0.065 0.065   
## Adjusted R2 0.047 0.040 0.040   
## Residual Std. Error 41.010 (df = 431) 41.154 (df = 422) 41.154 (df = 422)  
## =========================================================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

**Answer:** The first model looks at the number of girls and the total number of childern with no dummy variables. The seccond model has a dummy variable for each total amount of kids you could have (1-10). The third looks at the number of kids as a fixed effect.

## 8.2 Question: (2 pages) Reproduce the results in column 2 of table 2 from Washington’s paper.

**Code:**

# add age square and srvlng to data   
finaldata$age2 <- finaldata$age^2  
finaldata$srvlng2 <- finaldata$srvlng^2  
  
#reg.2 <- felm(aauw ~ ngirls+as.factor(totchi), finaldata)  
  
reg.t2 <- felm(data = finaldata, aauw ~ ngirls+female+white+repub\_i+age+age2+srvlng+srvlng2+as.factor(rgroup)+demvote|totchi+region)  
  
stargazer(reg.t2, header = FALSE, type = "text", se = list(reg.t2$rse))

##   
## ===============================================  
## Dependent variable:   
## ---------------------------  
## aauw   
## -----------------------------------------------  
## ngirls 2.385\*\*   
## (1.198)   
##   
## female 9.194\*\*\*   
## (3.336)   
##   
## white 0.144   
## (3.539)   
##   
## repub\_i -60.468\*\*\*   
## (2.993)   
##   
## age 0.854   
## (0.997)   
##   
## age2 -0.006   
## (0.010)   
##   
## srvlng -0.208   
## (0.320)   
##   
## srvlng2 0.004   
## (0.012)   
##   
## as.factor(rgroup)1 -5.671   
## (3.468)   
##   
## as.factor(rgroup)2 -10.175\*\*\*   
## (3.312)   
##   
## as.factor(rgroup)3 -2.466   
## (5.740)   
##   
## as.factor(rgroup)4 4.012   
## (3.926)   
##   
## demvote 62.148\*\*\*   
## (13.065)   
##   
## -----------------------------------------------  
## Observations 434   
## R2 0.840   
## Adjusted R2 0.828   
## Residual Std. Error 17.441 (df = 402)   
## ===============================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

## 

## 8.3 Question: Explain what the region fixed effects are controlling for?

**Answer:** They are controlling for the differences in regions. Rather than comparing all of the regions it is comparing the different congress members within their respective regions.

## Question: (2 pages) Reload the data and this time we will keep observations from all of the congresses. Generate a variable that creates a unique identifier for region by year. Estimate the following models and present your results in a table.

**is a fixed effect for the total number of children, is a region fixed effect, is a year (congress session) fixed effect and is a region by year fixed effect and is a total children by region by year fixed effect. Explain what the differences between these four different estimation. Is there a downside to a specification like the fourth specification?**

**Code:**

# build year fixed effect  
data8 <- basic  
  
data8$uniqueYearRegion <- paste(data8$year, data8$region, sep="\_")  
data8$uniqueYearRegionChildren <- paste(data8$year, data8$region, data8$totchi, sep="\_")  
  
  
# make regressions   
reg.1 <- felm(data = data8, aauw ~ ngirls|totchi+region)  
reg.2 <- felm(data = data8, aauw ~ ngirls|totchi+region+congress)  
reg.3 <- felm(data = data8, aauw ~ ngirls|totchi+region+uniqueYearRegion)  
reg.4 <- felm(data = data8, aauw ~ ngirls|uniqueYearRegionChildren)  
  
stargazer(reg.1, reg.2, reg.3, reg.4, header = FALSE, type = "text", se = list(reg.1$rse, reg.2$rse, reg.3$rse, reg.4$rse))

##   
## ===============================================================================================  
## Dependent variable:   
## ---------------------------------------------------------------------------  
## aauw   
## (1) (2) (3) (4)   
## -----------------------------------------------------------------------------------------------  
## ngirls 5.058\*\*\* 5.043\*\*\* 5.125\*\*\* 4.987\*\*\*   
## (1.207) (1.208) (1.216) (1.415)   
##   
## -----------------------------------------------------------------------------------------------  
## Observations 1,735 1,735 1,735 1,735   
## R2 0.148 0.151 0.155 0.238   
## Adjusted R2 0.138 0.139 0.127 0.110   
## Residual Std. Error 39.354 (df = 1714) 39.330 (df = 1711) 39.605 (df = 1679) 39.982 (df = 1486)  
## ===============================================================================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

**Answer:** the first 2 have have no combined fixed effects. This means if we wanted to we could see the effects of each variable on the result. The third has a combined year and region. The downside to the fourth method is that, since all of the fixed effects are combined, it may be difficult to tell what the effects of the different variables are. Additionally, the data may become too specific.

## Question: In her paper, Washington chooses not to pool the data for all four congresses and instead estimates her main specification on each year separately. Why do you think she makes this choice?

**Answer:** she examines each year seperatly because she does not want the variance between the different congresses effecting her results. She can come to a more accurate conclusion looking at each individual congress.

## Question: Check to see that names uniquely identify each congress person. If you are not sure if they do, make a unique identifier for each congress person.

**Answer:** They are all unique for each givin congress. If they werent there would be a value that was equal TRUE **Code:**

name\_105 <- basic\_105$name  
basic\_106 <- filter(basic, congress == "106")  
basic\_107 <- filter(basic, congress == "107")  
basic\_108 <- filter(basic, congress == "108")  
name\_106 <- basic\_106$name  
name\_107 <- basic\_107$name  
name\_108 <- basic\_108$name  
duplicated(name\_105)

## [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [13] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [25] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [37] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [49] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [61] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [73] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [85] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [97] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [109] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [121] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [133] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [145] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [157] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [169] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [181] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [193] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [205] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [217] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [229] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [241] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [253] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [265] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [277] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [289] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [301] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [313] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [325] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [337] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [349] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [361] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [373] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [385] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [397] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [409] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [421] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [433] FALSE FALSE FALSE

duplicated(name\_106)

## [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [13] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [25] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [37] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [49] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [61] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [73] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [85] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [97] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [109] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [121] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [133] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [145] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [157] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [169] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [181] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [193] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [205] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [217] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [229] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [241] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [253] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [265] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [277] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [289] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [301] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [313] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [325] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [337] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [349] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [361] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [373] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [385] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [397] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [409] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [421] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [433] FALSE FALSE FALSE

duplicated(name\_107)

## [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [13] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [25] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [37] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [49] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [61] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [73] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [85] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [97] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [109] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [121] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [133] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [145] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [157] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [169] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [181] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [193] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [205] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [217] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [229] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [241] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [253] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [265] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [277] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [289] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [301] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [313] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [325] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [337] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [349] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [361] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [373] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [385] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [397] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [409] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [421] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [433] FALSE FALSE FALSE

duplicated(name\_108)

## [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [13] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [25] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [37] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [49] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [61] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [73] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [85] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [97] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [109] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [121] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [133] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [145] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [157] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [169] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [181] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [193] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [205] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [217] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [229] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [241] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [253] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [265] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [277] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [289] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [301] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [313] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [325] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [337] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [349] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [361] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [373] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [385] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [397] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [409] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [421] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [433] FALSE FALSE FALSE

## Question:(2 pages) Because we have data for four congress sessions, we may be able to see how an individual congress person’s voting patterns change as the number of daughters they have changes. Propose an estimating equation that would allow you to estimate this, run your estimation and present your results. Be sure to define all new variables. What do your results tell you? Why?

**Equation:** We have to build a dataset of congress members whos number of children changed between different congresses then run a regression with the number of girls, total children, female, and republican

**Code:**

# have to be in congress twice   
dataCD <- basic  
# Remove duplicates by single column  
data\_repeat <- dataCD %>%  
 group\_by(name) %>%  
 filter(n()>1)   
  
  
# Number of daughters changed. Looks at people who have had a change in the number of daughters and have served multiple terms   
changed\_daughters <- data\_repeat %>%  
 group\_by(name) %>%  
 filter(length(unique(ngirls)) > 1) %>%  
 ungroup()  
  
# same as before but looks only after someone has had an additional daughter. Gets rid of whatever they were befores having another daughter while in congress  
changed\_daughters2 <- data\_repeat %>%  
 group\_by(name) %>%  
 arrange(congress) %>% # Sort by year to ensure chronological order  
 filter(ngirls > lag(ngirls, default = first(ngirls))) %>%  
 ungroup()  
  
  
reg.1 <- felm(aauw ~ ngirls+totchi+female+repub, changed\_daughters)  
reg.2 <- felm(aauw ~ ngirls+totchi+female+repub, changed\_daughters2)  
print(summary(reg.1),digits=3)

##   
## Call:  
## felm(formula = aauw ~ ngirls + totchi + female + repub, data = changed\_daughters)   
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -27.77 -12.03 -3.62 7.41 88.31   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 92.05 4.47 20.58 <2e-16 \*\*\*  
## ngirls 1.94 4.36 0.45 0.66   
## totchi -1.41 3.37 -0.42 0.68   
## female 9.69 9.16 1.06 0.29   
## repub -80.02 5.31 -15.07 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 21.8 on 71 degrees of freedom  
## Multiple R-squared(full model): 0.775 Adjusted R-squared: 0.762   
## Multiple R-squared(proj model): 0.775 Adjusted R-squared: 0.762   
## F-statistic(full model):61.2 on 4 and 71 DF, p-value: <2e-16   
## F-statistic(proj model): 61.2 on 4 and 71 DF, p-value: <2e-16

print(summary(reg.2),digits=3)

##   
## Call:  
## felm(formula = aauw ~ ngirls + totchi + female + repub, data = changed\_daughters2)   
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -21.44 -12.60 -7.36 9.69 87.50   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 87.689 10.632 8.25 1.1e-07 \*\*\*  
## ngirls 0.917 9.117 0.10 0.92   
## totchi -1.021 6.676 -0.15 0.88   
## female 15.249 21.238 0.72 0.48   
## repub -73.961 10.557 -7.01 1.1e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 23.9 on 19 degrees of freedom  
## Multiple R-squared(full model): 0.731 Adjusted R-squared: 0.675   
## Multiple R-squared(proj model): 0.731 Adjusted R-squared: 0.675   
## F-statistic(full model):12.9 on 4 and 19 DF, p-value: 3e-05   
## F-statistic(proj model): 12.9 on 4 and 19 DF, p-value: 3e-05

**Answer:** The results are pretty inconclusive from this test. The only significant coefficient is republican which makes sense. This has daughters positive but as mentioned earlier it is not significant.

## Question: Can you think of any identification concerns with this approach?

It is likely that the pool of people is simply too small which would make it difficult to have a significant coefficient regarding the number of daughters. Maybe if we used more congresses, a better conclusion could be made

**Answer:**

## Question: (2 pages) Using data from all four congresses, estimate the same specification as that used in column 2 of table 2 with the addition of year and individual fixed effects and report your results. Why aren’t you able to estimate a coefficient for certain covariates?

**Code:**

# add age square and srvlng to data  
basicdata <- basic  
  
basicdata$age2 <- basicdata$age^2  
basicdata$srvlng2 <- basicdata$srvlng^2  
basicdata$repub\_i <- ifelse(basicdata$party == 2, 1, 0)  
  
#reg.2 <- felm(aauw ~ ngirls+as.factor(totchi), finaldata)  
  
reg.89 <- felm(data = basicdata, aauw ~ ngirls+female+white+repub\_i+age+age2+srvlng+srvlng2+as.factor(rgroup)+demvote|totchi+region+congress+name)

## Warning in chol.default(mat, pivot = TRUE, tol = tol): the matrix is either  
## rank-deficient or not positive definite

stargazer(reg.89, header = FALSE, type = "text", se = list(reg.89$rse))

##   
## ===============================================  
## Dependent variable:   
## ---------------------------  
## aauw   
## -----------------------------------------------  
## ngirls 2.010   
## (2.700)   
##   
## female   
## (0.000)   
##   
## white   
## (0.000)   
##   
## repub\_i -3.034   
## (4.593)   
##   
## age 10.393\*\*\*   
## (3.947)   
##   
## age2 -0.003   
## (0.008)   
##   
## srvlng -0.990\*   
## (0.551)   
##   
## srvlng2 0.0004   
## (0.011)   
##   
## as.factor(rgroup)1   
## (0.000)   
##   
## as.factor(rgroup)2   
## (0.000)   
##   
## as.factor(rgroup)3   
## (0.000)   
##   
## as.factor(rgroup)4   
## (0.000)   
##   
## demvote 0.454   
## (8.239)   
##   
## -----------------------------------------------  
## Observations 1,735   
## R2 0.973   
## Adjusted R2 0.958   
## Residual Std. Error 8.716 (df = 1121)   
## ===============================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

**Answer:** since we added all of the years and the added fixed variables we now have to deal with more multicolinearity.

## Question: Which fixed effects from the original specification are now redundant?

**Answer:** female, white, and religion are all redundunt because they are included in the fixed effect

## Question: Can you estimate a coefficient for ? What does this imply?

**Answer:** we can estimate a coefficient for repub which implies that it is not included in the fixed effect.

# Submission instructions:

* Since this is a group assignment only one member of the group will upload it to gradescope.
* Make sure the final version of your assignment is knit in pdf format and uploaded to gradescope. Make sure you have one question response per page (unless otherwise indicated) so that question positions align with the template in gradescope.The final PDF should be 29 pages long.