

# Problem Set 6

Tate Mason

## Question 1

```
library(Matrix)  
library(dplyr)
```

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

filter, lag

The following objects are masked from 'package:base':

intersect, setdiff, setequal, union

```
library(tidyr)
```

Attaching package: 'tidyr'

The following objects are masked from 'package:Matrix':

expand, pack, unpack

```
library(magrittr)
```

Attaching package: 'magrittr'

The following object is masked from 'package:tidyr':

extract

```
df <- read.csv('metrics.csv')
```

## Part A

```
# Creating Variables
id <- df$id
year <- df$year

n <- length(unique(id))
t <- length(unique(year))
nt <- n*t

df %<>% mutate(D=1*(first.displaced <= year & first.displaced !=0), yearf = factor(year))

Y <- Matrix(df$learn)
X <- model.matrix(~ yearf + D-1, data = df)
X <- Matrix(X)

I <- diag(nt)
D <- bdiag(replicate(n,Matrix(rep(1,t)),simplify=FALSE))
M <- I - D%*%solve(t(D)%*%D)%*%t(D)
alpha_hat <- solve(t(X)%*%M%*%X)%*%t(X)%*%M%*%Y
e_hat <- Y - X%*%alpha_hat
var_error = as.numeric(t(e_hat) %*% e_hat) / (nt - ncol(X))
var_alpha = var_error * solve(t(X) %*% M %*% X)
se_alpha = sqrt(diag(var_alpha))
```

Warning in sqrt(diag(var\_alpha)): NaNs produced

```
cat("Part A: Report of Estimates\n")
```

Part A: Report of Estimates

```
alpha_hat
```

```
8 x 1 Matrix of class "dgeMatrix"
      [,1]
yearf2001 -0.23890499
yearf2003 -0.15619951
yearf2005 -0.07474901
yearf2007  0.02087565
yearf2009  0.03020829
yearf2011  0.10327769
yearf2013  0.12895513
D          -0.23172643
```

```
se_alpha
```

```
yearf2001 yearf2003 yearf2005 yearf2007 yearf2009 yearf2011 yearf2013      D
      NaN      NaN      NaN      NaN      NaN      NaN      NaN 0.4422551
```

## Part B

```
calc_group_time_att <- function(df, group, time) {

  if (time < group) {
    return(NA)
  }
  treat <- subset(df, first.displaced == group)
  control <- subset(df, first.displaced == 0)

  pre_period <- group - 2

  treat_pre <- subset(treat, year == pre_period)$learn
  treat_post <- subset(treat, year == time)$learn
  treat_diff <- mean(treat_post) - mean(treat_pre)

  control_pre <- subset(control, year == pre_period)$learn
  control_post <- subset(control, year == time)$learn
  control_diff <- mean(control_post) - mean(control_pre)

  att <- treat_diff - control_diff
```

```

    return(att)
}

groups <- c(2001, 2003, 2005, 2007, 2009, 2011, 2013)
years <- unique(df$year)

att_results <- matrix(NA, nrow = length(groups), ncol = length(years))
rownames(att_results) <- paste0("Group_", groups)
colnames(att_results) <- paste0("Year_", years)

for (i in 1:length(groups)) {
  for (j in 1:length(years)) {
    att_results[i, j] <- calc_group_time_att(df, groups[i], years[j])
  }
}

cat("Part B: Report of ATT Estimates\n")

```

Part B: Report of ATT Estimates

```
print(att_results)
```

	Year_2001	Year_2003	Year_2005	Year_2007	Year_2009	Year_2011
Group_2001	NaN	NaN	NaN	NaN	NaN	NaN
Group_2003	NA	-0.2091116	-0.1561571	-0.1774589	-0.2375154	-0.2347393
Group_2005	NA	NA	-0.1138004	-0.1123821	-0.1677389	-0.1698090
Group_2007	NA	NA	NA	-0.1988995	-0.3070383	-0.2000041
Group_2009	NA	NA	NA	NA	-0.3183560	-0.3114597
Group_2011	NA	NA	NA	NA	NA	-0.2504762
Group_2013	NA	NA	NA	NA	NA	NA
	Year_2013					
Group_2001	NaN					
Group_2003	-0.27811713					
Group_2005	-0.03129936					
Group_2007	-0.25059895					
Group_2009	-0.22098254					
Group_2011	-0.16332055					
Group_2013	-0.20561341					

```
cat("\n")
```