

Zarate Marques R Camp Day 3 HW

Zarate Marques

August 21, 2019

```
rm(list=ls())
gc();gc()
```

```
##          used (Mb) gc trigger (Mb) max used (Mb)
## Ncells  523008 28.0   1183481 63.3   621654 33.2
## Vcells 1002191  7.7    8388608 64.0   1600583 12.3
```

```
##          used (Mb) gc trigger (Mb) max used (Mb)
## Ncells  523447 28.0   1183481 63.3   621654 33.2
## Vcells 1003190  7.7    8388608 64.0   1600583 12.3
```

The question that I am asking today is whether or not news about immigration elicits anxiety among conservatives. Anxiety will arise out of uncertainty or threat where to some immigration can lead to uncertain changes in their surroundings or might be threatened by potentially losing their job. The reason that I am not including democrats in this sample is because we know that democrats typically are not threatened by immigration and so they would add noise to the analysis (The real reason is that the US population is too big and so when I tried to create another binary covariate stating whether or not they were democratic my R would crash). I have income as a covariate as someone with a higher income might not be as threatened of immigration. However, I predict that immigration will have a minimal effect.

```
rm(list=ls())

set.seed(8675309)

population = 131566280 # the population of conservatives in the United States

# population parameters
income <- rexp(n=population, rate = 1/59000) #59000 is the average income in US

sim.data <- data.frame(population, income)

state <- sample(1:50, size = nrow(sim.data), replace = TRUE)
sim.data$state <- state

agg.income <- c()

for (i in 1:50) {
  agg.income[i] <- sum(sim.data$income[state==i]) #aggregate income by state
}

# I don't actually need this aggregate variable. Income could be a factor but not aggregate income

a = 0.05 #intercept/baseline for anxiety about immigration
b = -0.0005 #beta for income, I don't expect it to be significant
tau = 10 #treatment effect, I expect the news story about immigration to increase anxiety
error = rnorm(n=population, mean = 0, sd = 1) #error'
```

```

#control
Y_0 <- a + tau*0 + b*income + error
#treatment
Y_1 <- a + tau*1 + b*income + error

pop_dat <- data.frame(Y_0, Y_1, income, population)

#Sampling from the population
N_samp <- 5000

sample_ind <- sample(1:nrow(pop_dat), size = N_samp) #sample index
sample_dat <- pop_dat[sample_ind,]

d <- ifelse(runif(N_samp)<=.5, 1, 0)
sample_dat$Y_obs <- d*sample_dat$Y_1 + (1-d)*sample_dat$Y_0
head(sample_dat$Y_obs)

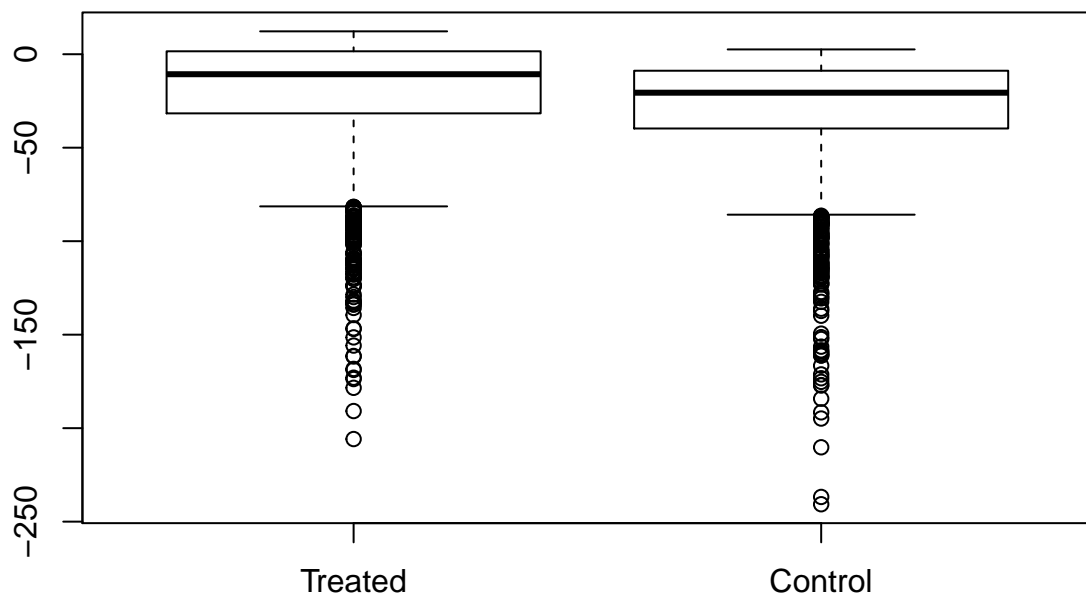
## [1] -4.367446 -42.906675 -10.960068  2.703413 -9.322092  6.423206

sample_dat$status <- ifelse(d==1, "Treated", "Control")
head(sample_dat)

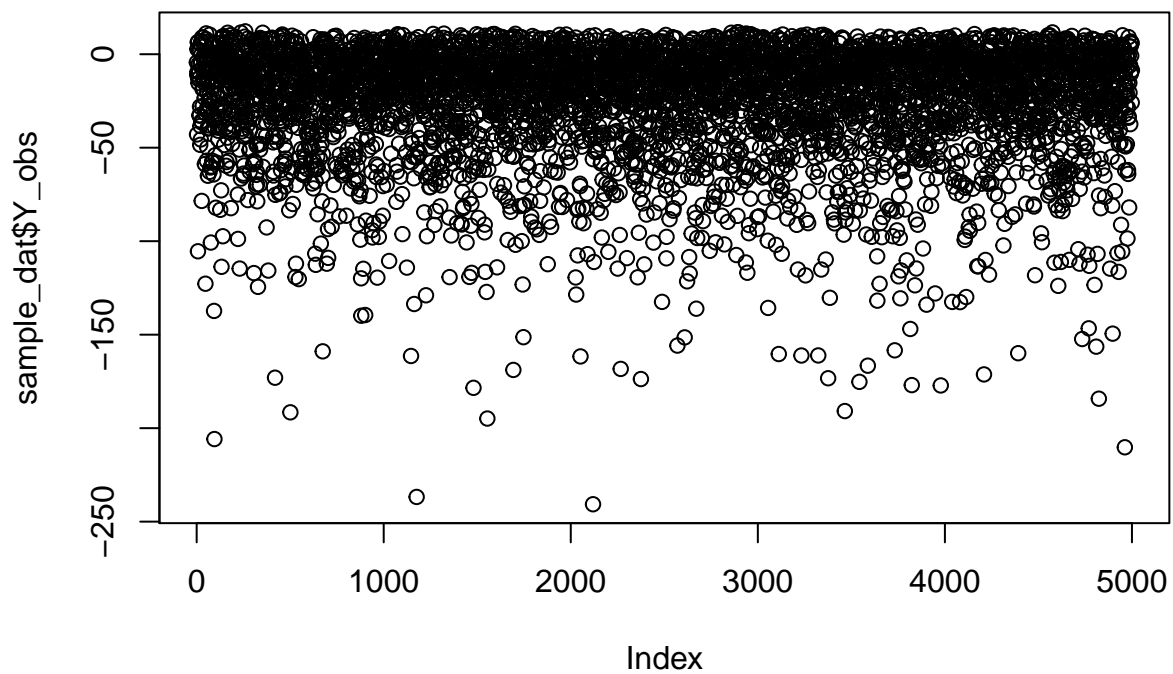
##           Y_0           Y_1    income population    Y_obs status
## 88047548 -14.367446 -4.3674460 26675.511  131566280 -4.367446 Treated
## 44027575 -42.906675 -32.9066746 81485.927  131566280 -42.906675 Control
## 102225841 -20.960068 -10.9600677 43937.467  131566280 -10.960068 Treated
## 131503747 -7.296587  2.7034132 15713.890  131566280  2.703413 Treated
## 2951324  -9.322092  0.6779083 24583.317  131566280 -9.322092 Control
## 75695247 -3.576794  6.4232060  6957.343  131566280  6.423206 Treated

#Visualization
boxplot(sample_dat$Y_obs[sample_dat$status=="Treated"],
        sample_dat$Y_obs[sample_dat$status=="Control"],
        names = c("Treated", "Control"))

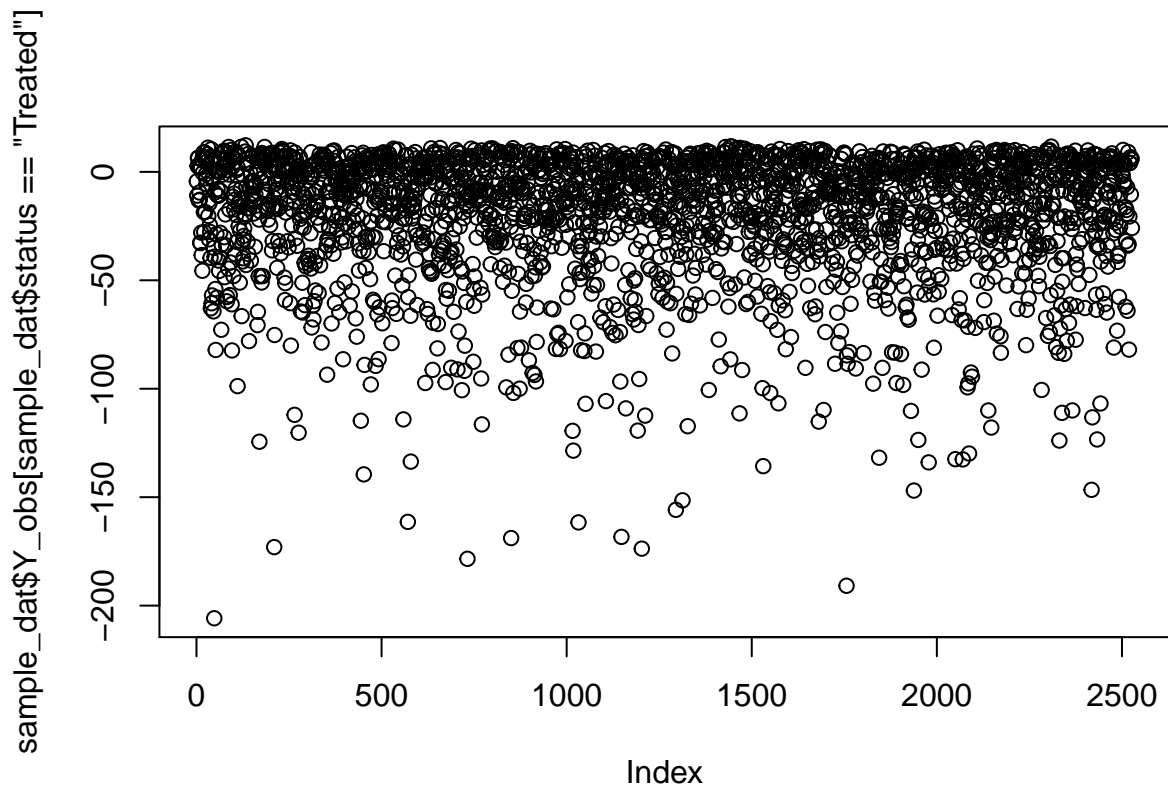
```



```
plot(sample_dat$Y_obs)
```



```
plot(sample_dat$Y_obs[sample_dat$status=="Treated"])
```



```
plot(sample_dat$Y_obs[sample_dat$status=="Control"])

t <- c(0,1)

y_mean <- c(mean(sample_dat$Y_obs[sample_dat$status=="Control"]),
             mean(sample_dat$Y_obs[sample_dat$status=="Treated"]))

plot(y_mean ~ t, xlab = "Treatment Status", ylab = "Levels of Anxiety",
     title("Simulated Result"))
```

sample_dat\$Y_obs[sample_dat\$status == "Control"]

Simulated Result

