ARE 213 Problem Set 2b

Nick Depsky, Will Gorman, Peter Worley

November 9, 2018

Import Data

```
#setwd("~/Dropbox/Berkeley_tings/Fall 2018/ARE213/Problem Sets/SharedFiles/are213/PS1b")
#setwd("C:\\Users\\will-\\Desktop\\are213\\PS2b")
setwd("C:\\Users\\Will\\Desktop\\are213\\PS2b")

dat <- read.dta("traffic_safety2.dta")
```

1a - Aggregate treatment analysis

In the first tibble printed below, we see that the average pre-period dependent variable for the TU site is -1.38 while that for the control states is -1.71 suggesting that the treatment states have a higher per capita fatality rate than the control states.

```
dat$fatal_pc <- log(dat$fatalities/dat$population)
dat$control <- ifelse(dat$state %in% c(6,10,30,41),1,ifelse(dat$state == 99,2,0))

dat %>% group_by(control,primary) %>% summarize(avg = mean(fatal_pc))
```

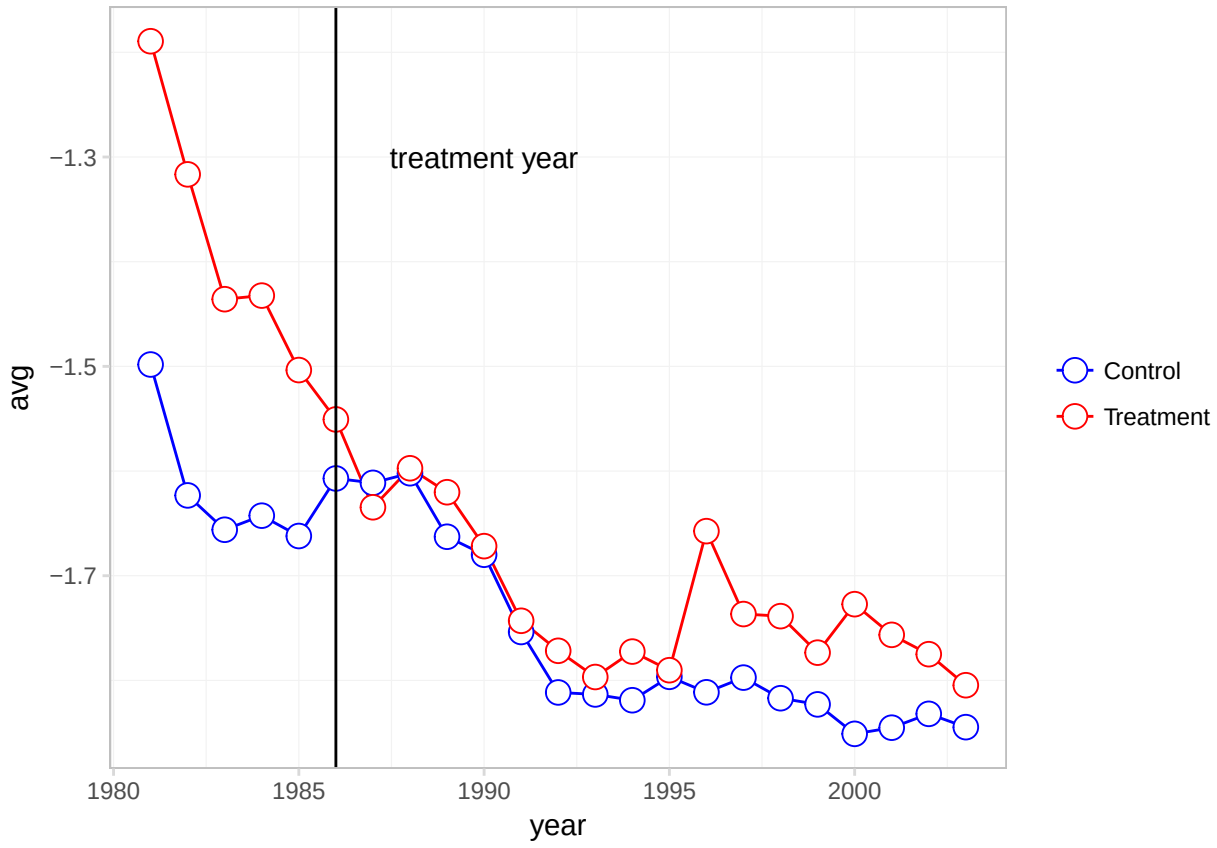
```
## Warning: package 'bindrcpp' was built under R version 3.4.4
```

```
## # A tibble: 6 x 3
## # Groups:   control [?]
##   control primary   avg
##     <dbl>   <int> <dbl>
## 1       0       0 -1.71
## 2       0       1 -1.92
## 3       1       0 -1.49
## 4       1       1 -1.76
## 5       2       0 -1.38
## 6       2       1 -1.72
```

```
## In the above table,  the pre-treatment average are when primary = 0
## The treatment aggregate is equal to 2 and
## the control average is equal to 0

summary_yr <- dat %>% group_by(control,year) %>% summarize(avg = mean(fatal_pc))

ggplot(summary_yr[summary_yr$control != 1,],
       aes(x=year, y=avg, group = factor(control), color = factor(control))) +
  geom_line() +geom_point( size=4, shape=21, fill="white") + theme_plot +
  scale_color_manual(labels = c("Control", "Treatment"), values = c("blue","red")) +
  geom_vline(xintercept = 1986) + annotate("text", x=1990, y=-1.3, label= "treatment year")
```



Looking at the TU pre-period (red line before 1986), we see that there was a significantly different pre-period trend that we would worry about when running our econometric analysis. The graph also shows that the treatment states had a noticeably higher pre-treatment per capita fatality rate than the control states.

```
rm_treat <- dat[dat$control != 1,]

result <- rm_treat %>%
             group_by(control, primary, state) %>%
             filter(year == max(year))

result <- result[result$primary == 0,c(1,2,15,16)]
result <- result[result$year != 2003,]
print(result[order(result$fatal_pc),])
```

```
## # A tibble: 15 x 4
##    state  year fatal_pc control
##    <int> <dbl>    <dbl>   <dbl>
##  1    29  1999    -2.44       0
##  2    45  2001    -2.22       0
##  3    32  1983    -2.14       0
##  4    18  1997    -2.14       0
##  5     4  1992    -2.00       0
##  6    20  1999    -1.97       0
##  7     7  2002    -1.87       0
##  8    13  1997    -1.85       0
```

```
##  9    16  1995    -1.60       0
## 10    35  1990    -1.60       0
## 11     9  1995    -1.59       0
## 12    99  1985    -1.50       2
## 13    34  1997    -1.39       0
## 14     1  1999    -1.36       0
## 15    25  1986    -1.35       0
```

```
rm_treat <- dat[dat$control != 1,]
```

Georgia appears to be the closest state.

1b - Synthetic control method

1c - Graphical interpretation and treatment

1d - Comparison to fixed effects