ARE 213 Problem Set 2B

Becky Cardinali, Yuen Ho, Sara Johns, and Jacob Lefler

Due 11/16/2020

Question 1

We first estimate an event study specification.

(a) First determine the minimum and maximum event time values that you can estimate in this data set. Code up a separate event time indicator for each possible value of event time in the data set. Estimate an event study regression using all the event time indicators. What happens?

The data set contains data for each year in $[1981, 2003]$. Therefore, the minimum number of event time values we can estimate in this data set is 23 $(2003 - 1981 + 1)$, which would be the case where all states are treated in the same year, all states are treated during all years in the data set, or all states are untreated during all years in the data set.

The maximum event time value we can estimate in this data set is 47?? $(2003 - 1981 + 1) * 2$, which would be the case where at least one state was treated for the entire sample (event time 1 to 23) and at least one state was never treated during the sample (event time -23 to -1) (23 + 23 + 1 for event time 0)

We have a single treated state ($s = 1$ when the state has a primary seat belt law) and a single control state ($s = 0$ when a state does not have a primary seat belt law). So for an event study regression using all the event time indicators, we will estimate the regression:

$$Y_{ist} = \alpha + \sum_{j=-T_0}^{T-T_0} \tau_j D_{jst} + \gamma_s + \delta_t + \epsilon_{st} + u_{ist}$$

where T_0 is the period just prior to treatment, and D_{jst} is an indicator function for period t falling j periods after T_0 in the treated state (i.e., $\mathbf{1}(t - T_0 = j) * \mathbf{1}(s = 1)$), so index j is "event time."

```
# separate event time indicator for each possible value of event time in the data set

# create event time variable
traffic <- traffic[, event_time := NA]

# iterate through all states (this includes state 99)
for(s in unique(traffic$state)){

  # data frame for just state s
  temp <- subset(traffic, state == s)
  # make sure it's in ascending order by year
  temp <- temp[order(year), ]

  # find the row # and year corresponding to the first occurrence of primary == 1 in the dataframe for just t
  row_of_first_occ <- match(1, temp$primary)
  first_treated_year <- traffic[match(1, temp$primary), year]

  # case where a state was never treated in the sample (row_of_first_occ == NA)
  if(is.na(row_of_first_occ)){
    k <- -1
    for(i in 23:1){
```

```r
      # year 2003 corresponds to event time = -1, year 2002 corresponds to event time = -2, etc.
      setDT(traffic)[, event_time := ifelse(state == s & year == (2004 + k), k, event_time)]
      k <- k - 1
    }
  }

  # case where a state was treated in the sample (row_of_first_occ > 1)
  # no cases of a state being treated during the entire sample
  else{
    # for the first year with primary == 1, make event time = 0
    setDT(traffic)[, event_time := ifelse(state == s & year == first_treated_year, 0, event_time)]

    k <- 1
    for(i in (row_of_first_occ+1):23){
      # year after event time 0 corresponds to event time 1, 2 years after event time 0 corresponds to event
      setDT(traffic)[, event_time := ifelse(state == s & year == (first_treated_year + k), k, event_time)]
      k <- k + 1
      }

    k <- -1
    for(i in (row_of_first_occ-1):1){
      # year before event time 0 corresponds to event time -1, 2 years before event time 0 corresponds to eve
      setDT(traffic)[, event_time := ifelse(state == s & year == (first_treated_year + k), k, event_time)]
      k <- k - 1
      }
  }
}

# make separate event time indicator for each possible value of event time in the data set + state and year d
traffic <- traffic[,dummy_cols(traffic, select_columns = c("year", "state", "event_time"))]

# create y variable
traffic[, ln_fat_pc := log((fatalities/population))]

# event study regression using all the event time indicators
# below isn't working because event_time_-23 and all the negative number ones are weird
# even when when I just included event_time_2*primary, it removed event_time_2*primary because of collinearit
# es1 <- feols(ln_fat_pc ~ 1 + event_time_-23*primary + (event_time_2*primary) + event_time_3*primary, fixef
```

Question 2

We now apply the synthetic control methods from Abadie et al (2010).

(a)

i. Compare the average pre-period log traffic fatalities per capita of the TU site to that of the average of all the "control" states. Next, graph the pre-period log traffic fatalities by year for the pre-period for both the TU and the average of the control group. Interpret.

```r
treatment_year <- 1986 # TU has primary=1 starting in 1986
traffic_pre <- traffic[year<1986,]
TU_fatal_pre <- traffic_pre[state==99,fatalities] # treated unit
CU_fatal_pre <- traffic_pre[state!=99 && state!=09 && state!=19 && state!=35&& state!=48,fatalities] # contro
```