ARE 213 PS 2b

S. Sung, H. Husain, T. Woolley, A. Watt

2021-11-23

Contents

Problem 1

We first estimate an event study specification.

Part (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?

Let's determine the event time values

Part (b)

Estimate another event study regression using all the event time indicators save one that you choose to omit. Generate a plot of the event study coefficients.

Part (c)

Create minimum and maximum event time indicators that correspond to bins of event time $<$ -5 and event time $>$ 5 respectively. Appropriately specify and estimate an event study regression using these min and max event time indicators. Generate a plot of the event study coefficients. Explain which specification you prefer, this one or the one in part (b).

Part (d)

What happens to your estimates from part (b) if you exclude the "pure control" states from your sample? What about if you exclude the pure controls in part (c)?

Part (e)

Overall, does the event study regression make you more confident or less confident that seat belt laws reduce fatalities (relative to the fixed effects results that you estimated on the last problem set)? Briefly explain.

Part (f*)

Building off the event study regression from part (c), estimate the interaction weighted event study estimator from Sun and Abraham (2020). As a reminder, the interacted event study regression takes the standard event time indicators (without any binning) and interacts each one with a cohort indicator (a cohort refers to a group of states that share the same date on which they were first treated). You then form the estimate for event time coefficient τ_j by averaging the estimates of the cohort-specific τ_j using the weights described in Sun and Abraham (2020).

Problem 2

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

Part (a)

We created an aggregate "treatment" state (state number 99 or "TU") which combines the (population weighted) data from the first 4 states to have a primary seatbelt law (CT, IA, NM, TX). Please use this state as the "treatment" state in the synthetic control analysis.

—— 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.

—— a.ii

Compare the dependent variable between the TU site and each control state for the year before the treatment. Which control state best matches the TU? Now compare this state's covariates with the TU covariates. Do they appear similar? What might this imply for in terms of using this state as the counterfactual state?

Part (b)

Apply the synthetic control method using the available covariates and pre-treatment outcomes to construct a synthetic control group.

—— **b.i**

Discuss the synthetic control method including its benefits and potential drawbacks.

—— **b.ii**

Use the software package provided by Abadie et al to apply the synthetic control method. (You are free to use either Stata, Matlab, or R but answers will be provided in Stata and R only). Please be sure to state precisely what the command is doing and how you determined your preferred specification.

Part (c)

Graphical interpretation and treatment significance.

—— c.i

Generate graphs plotting the gap between the TU and the synthetic control group under both your preferred specification and a few other specifications you tried.

—— c.ii

Compare the graph plotting the gap between the TU and the synthetic control group under your preferred specification with the graphs plotting the gap between each control state and its "placebo" treatment. Do you conclude that the treatment was significant? Why or why not?

—— c.iii

Create a graph of the post-treatment/pre-treatment prediction ratios of the Mean Squared Prediction Errors (MSPE) for the actual and "plecebo" treatment gaps in (ii). [See Abadie et al. for an example]. Do you conclude that the treatment was significant? Why or why not?

Part (d)

How do your synthetic control results compare to your fixed effects results from Question (3) in the last problem set? Interpret any differences.

Appendix A: R Code

```
rm(list=ls())
knitr::opts_chunk$set(echo = F)
# stargazer table type (html, latex, or text)
# Change to latex when outputting to PDF, html when outputting to html
table_type = "latex"

# install.packages("Synth")
library(tidyverse)
library(haven)
library(stargazer)
library(ggplot2)
library(tinytex)
library(Synth)
# library(plm)
# library(lmtest)
# library(sandwich)
# library(gridExtra)
# library(grid)
# library(gtable)
# library(fastDummies)
# library(EnvStats)


# Load data from PS2a with previous log variables
data = read_dta('traffic_safety2.dta') %>%
    mutate(fat_pc = fatalities/population,
           ln_fat_pc = log(fat_pc),
           ln_tvmt_pc = log(totalvmt/population),
           ln_precip = log(precip),
           ln_rspeed = log(rural_speed),
           ln_uspeed = log(urban_speed))

#heres a chunk
```