EAE6029 - Econometria I
Lista 3

Prof: Pedro Forquesato

Faculdade de Economia, Administração e Contabilidade
Universidade de São Paulo

2023/1

Exercise list 3

Analytical problems

1. Is $\mathbb{E}[\epsilon_{it} \mid X_{it}] = 0$ enough for the fixed effects estimator to be consistent for β? Show it.
2. Show that $\text{Var}(\ddot{X}_{it}) \leq \text{Var}(X_{it})$.
3. Use (2) to show that $\text{Var}(\widehat{\beta}_{\text{fe}}) \geq \text{Var}(\widehat{\beta}_{\text{pols}})$. Assume spherical errors for simplicity.
4. Show that if $T = 2$ the first-difference and within estimators coincide.
5. Consider a difference in differences model with $T = 2$ and $N = 2$, but suppose the time variable is ommited. Namely, we have the following model:

$$Y_{it} = \theta \text{Treated}_i \times \text{After}_t + \text{Treated}_i + \epsilon_{it}$$

Find an expression for the OLS estimator $\hat{\theta}$ and show it is a function only of the treated sample. Under what assumption it is an appropriate estimator of the treatment effect?

Computational/interpretative problems

For this list, we will use the `DS2004.dta` file provided by the textbook author. You should **provide the code** and the results *together*.

```
library(haven)
library(kableExtra)
library(knitr)
```

```r
# setwd() <- this might help

#  you can download the file directly in R, or just do it manually
# url <- "https://www.ssc.wisc.edu/~bhansen/econometrics/Econometrics%20Data.zip"
# download.file(url, "./econ_data.zip")
# unzip("./econ_data.zip")

ds2004 <- read_dta("./DS2004.dta")
knitr::kable(head(ds2004[, c("block", "barrio", "altura", "sameblock", "distance",
                             "public", "bank", "thefts", "gasstation", "month")], 10))
```

block	barrio	altura	sameblock	distance	public	bank	thefts	gasstation	month
870	Once	a2300	0	1	1	0	0	0	4
851	Once	a2400	0	2	0	0	0	0	4
843	Once	a2300	1	0	0	1	0	0	4
796	Once	a500	0	1	0	0	0	0	4
790	Once	a700	1	0	0	0	0	0	4
789	Once	a600	0	1	0	0	0	0	4
844	Once	a2400	1	0	0	0	0	0	4
858	Once	a2400	0	1	0	1	0	0	4
787	Once	a400	0	1	0	1	0	0	4
850	Once	a2300	0	1	0	0	0	0	4

Here we will take a look at DiTella and Schargrodsky (2004) “Do Police Reduce Crime? Estimates Using the Allocation of Police Forces After a Terrorist Attack”.

They analyze whether a plausibly exogenous increase in policing after a terrorist attack in Argentina led to a decrease in automobile theft.

The variable “sameblock” points to whether there is a Jewish institution on the block (which were the target of the terrorist attack and received extra policing). The authors used data from April (month 4) to December 1994 (month 12), and the terrorist attack (treatment) ocurred in July 1994 (month 6).

1. Aggregate the data and run a 2-by-2 diff-in-diff. Construct the 2-by-2 table (like in the slides or textbook), and add the standard deviation of the estimates. Interpret.

For the rest of the list, return to the original (disaggregated) data.

2. Run the canonical DiD regression and a 2WFE regression. Are the results the same?

3. What about if you add controls? Which controls you consider reasonable? Why? Discuss the conditional ignorability of the model, with and without controls.
4. Estimate the standard deviation of the model assuming spherical errors, heteroskedasticity, and autocorrelation. Are the results similar? Which would you use?
5. Plot the average car theft in treatment and control blocks. Are the trends parallel? Estimate a dynamic diff-in-diff and plot it. Interpret.
6. Use the dynamic diff-in-diff above to see how the treatment effect changes over time. Is the policing effect long-lasting or short-lived?
7. Now consider the variable "oneblock", which denotes whether there is a Jewish institution 1 block away. Do the effects extend an extra block?

Let's now play a little with synthetic control. The terrorist attack happened at Asociación Mutual Israelita Argentina, in calle Pasteur, 633. (I already found it for you, it is block 797 in the data-set.)

8. Estimate a synthetic control model to check how the terrorist attack affected car theft in the block of the Asociación.
9. Do the same for 100 other blocks and plot the graphs as a placebo inference test, to check whether the result you found is unlikely or not. Interpret.

Now let's look at some more simple panel data models.

10. Run a pooled OLS regression of sameblock interacted with post July and the controls you used above. Is the result different? What does this say about the endogeneity of the treatment? Be precise.
11. Now add time-specific error components. They alter the results? Interpret.
12. Now run a random effects model. Compare the estimatives and their variance of all the estimators above. What do you see and what did you expect? Why?
13. Do a Hausman test on whether the treatment is exogenous regarding individual-specific error components. Interpret.
14. Since we do not have that many observations, and they are spatially correlated, we might be worried about our inference. Do a placebo test, randomly selecting localities as treated and control groups. Plot a histogram with their distribution, and denote the actual treatment effect. What is the probability you estimated like this of our result being purely by chance?
15. Let's pretend the authors had data on police presence (see the code below). It is reasonable to expect police presence to be endogenous, so we can use "sameblock" interacted with "after july" as an instrument in a 2WFE regression. Run that and a normal 2WFE and compare. Present all the important information and discuss in detail the results.

```
library(fixest)

ds2004$post_treat <- ds2004$month > 6
```

```r
u <- feols(thefts ~ 1 | month + block, data = ds2004)

ds2004$police <- 10*round(u$residuals*20) + 50*ds2004$sameblock*ds2004$post_treat +
    70*ds2004$sameblock + 100*ds2004$public + 100*ds2004$bank +
    50*ds2004$gasstation + 100*(ds2004$barrio == "Once") +
    sample(100:500, size = 7884, replace = TRUE)

knitr::kable(head(ds2004[, c("block", "barrio", "altura", "sameblock", "distance",
                             "bank", "thefts",  "month", "police")], 5))
```

block	barrio	altura	sameblock	distance	bank	thefts	month	police
870	Once	a2300	0	1	0	0	4	521
851	Once	a2400	0	2	0	0	4	376
843	Once	a2300	1	0	1	0	4	744
796	Once	a500	0	1	0	0	4	456
790	Once	a700	1	0	0	0	4	569