PhD Econometrics (ECON50580)
Replication and Practice Exercise
aka New Adventures in Causal Inference

This exercise has two goals. One is to bring you one step closer to conducting independent empirical research. Here you can pick a paper and work on some well-defined tasks, but you have considerable discretion over how you approach these tasks. Some techniques will be familiar from the lecture or from previous courses, while others are completely new. In your own projects, you will face many similar challenges: cleaning the data, choosing which analysis to run, interpreting the results, etc. Another important goal of this exercise is to familiarize you with the latest refinements of important methods such as Diff-in-Diff, IV, or fixed effect estimation.

Ground Rules

This assignment has three parts:

1. Replication of the main result of a published paper
2. Application of a new methodology to the data and context the published paper
3. A presentation of your findings in the last two weeks of the course

Deliverables:

1. A written report. The report should i) explain how you go about the replication, and discuss differences between your findings and those in the original paper, as well as limitations of the analysis in the original paper; ii) the application of the new method, which should include a brief description of the method and a discussion of the application of the new method. Code should be attached. Approximate length: 15 pages.
2. A presentation of your findings.

How it works

- Teams of 2-3 students were randomly assigned (see below). Each team will receive joint grades for the written work and presentation.
- You have to pick a paper from an approved list of journals (see below for a list) that uses the basic method indicated below. If you want to replicate a paper that is not from the list of approved journals, please contact me first.
- Criteria for picking the paper: 1) it has to use a given method, 2) you have to be able to get the data to replicate the paper, 3) the paper has not been used as an example to demonstrate the new methodology (most papers showcase their method based on an existing study; DO NOT use that study).
- Replicate the main findings of the paper using software that is not Stata. Most authors offer Stata codes, so replicating the paper in Stata would be trivial. Instead use R, Python, Julia or SAS. Check that your summary statistics match with those in the paper (or if they don't, report and discuss it). Similarly, you may not be able to replicate the results. If that's the case, don't worry, this is very common. Just explain what you did and why,

and discuss why you think there is a difference between the authors' results and yours. You do not have to replicate all findings in a paper. One main finding plus descriptive stats (and potentially one or two graphs) are fine.

- Apply the new methodology to the original paper.
- You will not be assessed based on the correctness of the results but rather on your ability to i) solve problems, ii) critically discuss econometric methods and iii) apply methods to a new context.

Deadline for Submission through Brightspace: April 10. The same submission guidelines as for the problem sets apply to this assignment (i.e. one pdf, show evidence of using version control and soft coding, attach the code, etc).

Groups and Topics

Group 1:

- Basic method: Estimation of Treatment Effects
- New method to apply: Using causal forests to estimate heterogeneous treatment effects (Wager & Athey, 2018)
- Members: Haochi, Sam
- Presentation date: April 18

Group 2:

- Basic method: Estimation of Treatment Effects
- New method to apply: Contamination Bias in Linear Regression (Goldsmith-Pinkham *et al.* , 2021)
- Members: Nadiya, Andy
- Presentation date: April 18

Group 3:

- Basic method: IV
- New method to apply: Doubly Robust LATE (Sloczynski *et al.* , 2022)
- Members: Shreya, Glenn
- Presentation date: April 18

Group 4:

- Basic method: Diff-in-Diff
- New method to apply: Synthetic Difference-in-Differences (Arkhangelsky *et al.* , 2021)
- Members: Aditi, Xuejing

- Presentation date: April 25

Group 5:

- Basic method: Diff-in-Diff
- New method to apply: Doubly robust DiD (Sant'Anna & Zhao, 2020)
- Members: Yasemine, Darragh
- Presentation date: April 25

Approved Journals

Quarterly Journal of Economics, American Economic Review, Journal of Political Economy, Review of Economic Studies, AEJ:Applied, AEJ:Policy, Journal of the European Economic Association, Economic Journal, Review of Economics & Statistics, Journal of Labor Economics, Journal of Human Resources, Journal of Development Economics, Journal of Public Economics, Journal of Health Economics.

References

Arkhangelsky, Dmitry, Athey, Susan, Hirshberg, David A., Imbens, Guido W., & Wager, Stefan. 2021. Synthetic Difference-in-Differences. *American Economic Review*, **111**(12), 4088–4118.

Goldsmith-Pinkham, Paul, Hull, Peter, & Kolesar, Michal. 2021. *Contamination Bias in Linear Regressions.*

Sant'Anna, Pedro H.C., & Zhao, Jun. 2020. Doubly robust difference-in-differences estimators. *Journal of Econometrics*, **219**(1), 101–122.

Sloczynski, Tymon, Uysal, Derya, & Wooldridge, Jeffrey M. 2022 (Nov.). *Doubly Robust Estimation of Local Average Treatment Effects Using Inverse Probability Weighted Regression Adjustment.* IZA Discussion Papers 15727. Institute of Labor Economics (IZA).

Wager, Stefan, & Athey, Susan. 2018. Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. *Journal of the American Statistical Association*, **113**(523), 1228–1242.