ReadMe for Replicating

A Practical Guide to Counterfactual Estimators for Causal Inference with Time-Series Cross-Sectional Data

Licheng Liu (MIT) Ye Wang (UNC) Yiqing Xu (Stanford)

Computational platforms:

1. We tested our code on an iMac5K 2020 with Intel 6-core processor and 64G RAM. The operation system is MacOS 12.0.1 (Monterey).
2. The R version is 4.1.0.

Installing packages

1. Please install "fastplm" (1.0.8), "panelView" (1.1.9) and "fect" (0.4.1) from tar.gz files from the "packages" folder, for example, in Mac/Linux command lines:

>R CMD install fect_0.4.1.tar.gz

or install from files in RStudio.

2. Please install the following packages from CRAN.

- readstata13 (>0.10.0)
- foreach (>1.5.2)
- future (>1.23.0)
- doParallel (>1.0.16)
- ggplot2 (>3.3.5)
- GGally (>2.1.2)
- gridExtra (>2.3)
- abind (>1.4-5)
- doRNG (>1.8.2)
- Rcpp (>1.0.8)
- extraDistr (>1.9.1)

3. Please check this page if you encounter an installation error, such as the "-lgfortran" error.

Folders and their functionalities

code	Stores code files required to simulate and plot data, and summarize results
data	Stores data for three empirical examples
log	Stores log files
packages	Stores required packages
results	Stores intermediate empirical and simulation results (~47M)
graph	Saves graphics

Note that the file structure needs to be recreated if you download the replication materials from Harvard Dataverse. In addition, two files are in the root folder:

ReadMe.pdf	This file
Codebook.pdf	Codebook for the datasets

Notes

1. Please make sure to set path to the root replication folder in *R*.
2. Please execute the code files sequentially based on the name order.
3. All graphics are saved in the "graph" folder. When you run the code, existing files in these folders will be overwritten, which is normal.
4. Figures 1-3, Tables 1-2 in the main text, as well as Figure A1-A5, A12 and Table A1 in SI, present information on the model setup, estimation strategies, diagnostic tests, and comparisons with other methods. Replication is not applicable.
5. Certain results (e.g., runtime and confidence intervals) may not be exactly republicated due to randomness of the simulation processes using parallel computing.
6. Full replication takes a few days to finish. To speed things up, you can use the intermediate files that store empirical and simulation results to make plots and show results.

Mapping file names to Figures appearing in the paper

In Paper	File Name	In SI	File Name
Figure 4	sim_gap.pdf	Figure A6(a)	sim_infer_n50.pdf
Figure 5	sim_placebo.pdf	Figure A6(b)	sim_infer_n100.pdf
Figure 6	sim_equiv.pdf	Figure A7	sim_treat.pdf
Figure 7	sim_carryover.pdf	Figure A8	sim_outcome.pdf
Figure 8(a)	ex_HH2015_gap.pdf	Figure A9	sim_compare.pdf
Figure 8(b)	ex_HH2015_placebo.pdf	Figure A10(a)	sim_tests_n100.pdf
Figure 8(c)	ex_HH2015_equiv.pdf	Figure A10(b)	sim_tests_n300.pdf
Figure 9(a)	ex_FM2015_gap.pdf	Figure A11	ex_HH2015_treat.pdf
Figure 9(b)	ex_FM2015_placebo.pdf	Figure A13	ex_FM2015_treat.pdf
Figure 9(c)	ex_FM2015_carryover.pdf	Figure A14	ex_FM2015_equiv.pdf
		Figure A15(a)	ex_FM2015b_gap.pdf
		Figure A15(b)	ex_FM2015b_placebo.pdf
		Figure A15(c)	ex_FM2015b_carryover.pdf
		Figure A16(a)	ex_FM2015_cohorts0.pdf
		Figure A16(b)	ex_FM2015_cohorts.pdf

Code files

Filename	Description	Output	Execution Time
0_root.R	Execute all code files.		
1_ex_sim0.R	Performs analysis on a simulated dataset.	• Figures 4, 5, 6, 7, A7, A8, • simexample.RData	~1.5 hours
2_ex_HH2015.R	Performs analysis on Hainmueller and Hangartner (2015).	• Figures 8 and A11 • ex_HH2015.RData	< 10 minutes
3_ex_FM2015.R	Performs analysis on Fouirnaies and Mutlu-Eren (2015).	• Figures 9, A13, 14, 15, 16 • ex_FM2015.RData	< 30 minutes
4_sim_tests_n100.R	Conducts simulations comparing the F test and the equivalence test (n = 100).	• sim_tests_n100.RData	< 2 hours
5_sim_tests_n300.R	Conducts simulations comparing the F test and the equivalence test (n = 300).	• sim_tests_n300.RData	< 3 hours
6_sim_ife_mc.R	Conducts simulations comparing IFEct and MC.	• sim_ife_mc.RData	~ 16-24 hours
7_sim_inference.R	Conduct simulations to investigate properties of inferential methods.	• qqplots_N50.RData • qqplots_N100.RData	< 30 minutes
8_plot_sim.R	Plot simulation results.	• Figure A6 • Figure A9 • Figure A10	< 5 minutes
simulateData.R	Code to simulate panel data with general treatment structure.	It will be sourced by other program files. No need to run independently.	N/A
simulateDID.R	Code to simulate panel data with DID or staggered adoption treatment.	It will be sourced by other program files. No need to run independently.	N/A

Note: There is no need to replicate Figures 1-3, Tables 1-2 in the main text, as well as Figure A1-A5, A12 and Table A1 in SI. Code files should be put in a subfolder called “code”.

Other files

Filename	Subfolder	Description
Codebook.pdf	root	Codebook for the datasets.
hh2015.dta	data	Data file for Hainmueller and Hangartner (2015).
fm2015.dta	data	Data file for Fouirnaies and Mutlu-Eren (2015).
fastplm_1.0.8.tar.gz	package	Source file for R package “fastplm”.
panelView_1.1.9.tar.gz	package	Source file for R package “panelView”.
fect_0.4.1.tar.gz	package	Source file for R package “fect”.
simexample.RData	results	Stored results for the simulated example.
ex_HH2015.RData	results	Stored results for Hainmueller and Hangartner (2015).
ex_FM2015.RData	results	Stored results for Fouirnaies and Mutlu-Eren (2015).
qqplots_N50.RData	results	Stored results for the QQ plot (N = 50).
qqplots_N100.RData	results	Stored results for the QQ plot (N = 100).
sim_ife_mc.RData	results	Stored results for the IFE/MC comparison.
sim_tests_n100.RData	results	Stored results for the F/equivalence test comparison (N = 100)
sim_tests_n300.RData	results	Stored results for the F/equivalence test comparison (N = 300)
log_simdata.txt	log	Log file for “1_ex_sim0.R”.
log_hh2015.txt	log	Log file for “2_ex_HH2015.R”.
log_fm2015.txt	log	Log file for “3_ex_FM2015.R”.
log_sim_tests_n100.txt	log	Log file for “4_sim_tests_n100.R”.
log_sim_tests_n300.txt	log	Log file for “5_sim_tests_n300.R”.
log_inference.txt	log	Log file for “7_sim_inference.R”.
log_sim_ife_mc.txt	log	Log file for “6_sim_ife_mc.R”.