Readme for replication package of "Efficient Estimation for Staggered Rollout Designs", by Roth and Sant'Anna, published in JPE: Micro

The code in this replication package constructs the analysis file from Roth and Sant'Anna (2023). Our replication package contains two separate zip files: `RS_Staggered_simulations.zip` and `RS_Staggered_application.zip`. `RS_Staggered_simulations.zip` contains all the codes and results of our Monte Carlo simulation exercises, while `RS_Staggered_application.zip` contains all the codes to replicate our empirical application. The entire analysis was conducted in `R`. In what follows, we explain each of these folders separately.

1 Simulation Files

Once unzipped, the `RS_Staggered_simulations` folder contains several sub-folders. We recommend using the `RStudio` project file, `Staggered_Simulations_SR2022.Rproj`, to navigate the replication files for the simulations. These files are also available on Github, at https://github.com/pedrohcgs/RS_Staggered_simulations.

In what follows, we describe each folder and its contents:

1. `Codes`: Contains all the codes to run the simulations.

   (a) `1_simulations-using-wood-et-al-functionalized.R` contains codes to run all the simulations based on Wood, Tyler, Papachristos, Roth and Sant'Anna (2020b) in `Azure`. In order to successfully run this code, you need to set up an `Azure` account. We recommend following the steps provided here: https://github.com/Azure/doAzureParallel. This would require you to set the credential via the `credentials2_anonymous.json` file in the parent folder. Since this can be time-consuming, we have saved all simulation results into the `Temp/Wood-et-al-sims/` folder.

   (b) `2_simulations-using-wood-et-al-graphs-functionalized.R` contains code to analyze all the simulation results generated in `1_simulations-using-wood-et-al-functionalized.R`. It creates the tables and saves them at `Tables/Wood-et-al-sims/` folder. This code can be successfully run using the saved simulation results, without running step 1.

   (c) `3_simulations_two_periods.R` contains codes to run all the simulations based on the simple setup with two periods and two groups in `Azure`. We have saved the simulation results into the `Temp/` folder as `2period-sim-results.rds`.

(d) `aux_fte_theme.R` contains auxiliary code to fix the style of plots.

(e) `aux_simulation_functions.R` contains auxiliary code to run the simulations in the simple case with two periods and two treatment groups.

2. `Data`: Contains the dataset used by Wood, Tyler and Papachristos (2020a) and expanded in Wood et al. (2020b). These files are also available at https://github.com/george-wood/procedural_justice and https://github.com/george-wood/procedural_justice_revisited, respectively. We keep them here so everything is self-contained. We use these files in our simulation based on Wood et al. (2020b).

3. `Tables`: Contains the output tables for all the simulations.

4. `Temp`: Contains "temporary" files, which, in our context, are simulation results generated by the codes in the `Codes` folder.

2 Application

Once unzipped, the `RS_Staggered_application` folder contains several sub-folders. We recommend using the `RStudio` project file, `RS_Application.Rproj`, to navigate the replication files for our empirical application revisiting Wood et al. (2020b). These files are also available on Github, at https://github.com/pedrohcgs/RS_Staggered_application.

In what follows, we describe each folder and its contents:

1. `Codes`: Contains all the codes to replicate the empirical application.

   (a) `00_run_all.R` is the master file that calls all the other codes and replicates the entire analysis.

   (b) `01_main_wood-et-al-no-special_ops-no-pilot.R` contains the main codes to replicate our analysis. It replicates Figure 1 and Table 5 of the main text, Figures C.1, C.2, and C.4 in the Appendix.

   (c) `03_balance_checks_pre_tests.R` contains codes to do the balance tests and reproduces Table 6 of the main text and Table C.1 in the Appendix.

   (d) `04_balance_on_age.R` contains codes to do the balance tests on age and reproduces Figure C.3 in the Appendix.

   (e) `aux_fte_theme.R` contains auxiliary code to fix the style of plots.

2. `Data`: Contains the dataset used by Wood et al. (2020a) and expanded in Wood et al. (2020b). These files are also available at https://github.com/george-wood/procedural_justice and https://github.com/george-wood/procedural_justice_revisited, respectively. We keep them here so everything is self-contained. We use these files in our simulation based on Wood et al. (2020b).

3. `Tables`: Contains the output tables for all the models and specifications.

4. `Figures`: Contains the output plots for all the models and specifications.

5. `Temp`: Contains "temporary" files, which, in our context, are the empirical results generated by the codes in the `Codes` folder.

References

Roth, Jonathan and Pedro H. C. Sant'Anna, "Efficient Estimation for Staggered Rollout Designs," *Journal of Political Economy: Microeconomics*, 2023, (Forthcoming).

Wood, George, Tom R. Tyler, and Andrew V. Papachristos, "Procedural justice training reduces police use of force and complaints against officers," *Proceedings of the National Academy of Sciences*, 2020, *117* (18), 9815–9821.

_ , _ , _ , Jonathan Roth, and Pedro H.C. Sant'Anna, "Revised Findings for "Procedural justice training reduces police use of force and complaints against officers"," *Working Paper*, 2020.