README for "Labor Market Design Can improve Match Outcomes: Evidence from Matching Teach For America Teachers to Schools"

Overview

The code in this replication package constructs the analysis files from several data sources and output for the paper using Stata. One master file runs all of the code to generate the output for the figures and tables in the main paper and the appendix. This file also contains information about the proprietary data used in this article.

Data Availability and Provenance Statements

The empirical results from "Labor Market Design Can improve Match Outcomes: Evidence from Matching Teach For America Teachers to Schools" are derived from proprietary administrative data from Teach For America.

Researchers may request access to the administrative data required to replicate the results from this paper by contacting Teach For America at research@teachforamerica.org.

Statement about Rights

I certify that the author of the manuscript has legitimate access to and permission to use the data used in this manuscript.

Summary of Availability

These data are proprietary and cannot be made publicly available. See above for instructions on how to request access to the data.

Required Data Files

To replicate the analysis in this paper using the proprietary data, the following datasets must be stored in a subfolder titled "data":

teacher_data.dta: Contains information about Teach For America teachers in the Chicago region and three comparison regions who were in the 2009 through 2014 cohorts. Special education teachers are excluded.

preference_data.dta: Contains information about teachers' and schools' preference reports from the 2014 high school interview days where DA was used. Teacher and school characteristics are included.

all_int_days.dta: Contains information about interview day schedules for all interview days held for the 2013 and 2014 cohorts of TFA teachers in the Chicago region. Schedules are linked to teacher and school characteristics.

Computational requirements

Software Requirements

- All code was run on Stata-MP version 18.0 with up to 4 cores with the following added packages:
  - jarowinkler
  - reclink
  - synth
  - unique

Memory and Runtime Requirements

Summary The code was last run on May 19, 2025 on a **Mac Studio 2022 with an Apple M1 Max Chip nad 64 gb of memory**. It was run using STATA/MP 18.0 for Mac with a license that includes up to 4-cores.

Description of programs/code

The code needed to replicate the results for this project is saved in a folder called "Programs" in the replication package. The program "_run_all_files.do" organizes the analysis and is the only file that needs to be modified.

In order to run the code, update the global macro "dir" on line 8 of _run_all_files.do to the correct project directory on the computer you are using. This directory should include a folder called "Programs" with all of the programs that are called in _run_all_files.do.

When the proprietary data is available, all of the tables and figures in the paper and appendix can be replicated by running _run_all_files.do.

_run_all_files.do calls the following programs:

- 1_teacherpreferences.do – Estimates teacher preferences. Produces teacher estimates in Table 1 and Table A4. Also saves teacher utility draws for simulation.
- 2_schoolpreferences.do – Estimates school preferences. Produces teacher estimates in Table 1. Also saves school position utility draws for simulation.
- 3_prepSimulation.do – Prep simulation data.
- 4_estimateFOthresholds.do – Estimate thresholds for FO.
- 5_simulateFO.do – Simulate FO.
- 6_simulateOtherMechanisms.do – Simulate DA, TP-DA, CP-DA, RSD, and Snake Draft.
- 7_cleanSimulation.do – Compile simulation results into one file.
- 8_anSimulation.do – Saves estimates based on simulation, including Panel A of Table 2, Table 3, and Table 4.
- 9_hiringDiffInDiff.do – Diff-in-diff interview day outcomes. Produces Panel B of Table 2.
- 10_retention_did.do – Diff-in-diff retention impacts. Panel A of Table 5.
- 11_retention_synth.do – Synthetic control retention estimates. Panel B of Table 5.
- 12_satisfaction_did.do – Diff-in-diff satisfaction impacts. Panel A of Table 6.
- 13_satisfaction_synth.do – Synthetic control satisfaction estimates. Panel B of Table 6.
- 14_intDayAttend.do – Creates appendix Tables A1, A2, and A5.
- 15_schedules.do – Create appendix Tables A3 and A10.

- 16_attritionCorrelates.do – Creates appendix Table A6.
- 17_retentionHumanities.do – Creates appendix Table A7.
- 18_satisfactionHumanities.do – Creates appendix Table A8.
- 19_performance_did.do – Creates appendix Table A9 and Figure A2.
- 20_fo_thresholds.do – Creates Appendix Figure A1.