

Replication Package ReadMe

1. Overview

This replication package contains all code used to produce the results in the paper. Due to administrative and confidential data restrictions, the datasets are not included. Nevertheless, the full computational workflow is documented to allow verification of the code structure and logic.

2. Data Availability

The data used in this project are administrative and confidential and cannot be shared publicly. Instructions on how to request access are provided in the separate document `Data_Access_Instructions.pdf`.

The dataset used in the paper is constructed from multiple administrative sources and must be reconstructed upon access to the underlying data.

3. Folder Structure

The replication package is organized as follows:

- `stata/`: Stata scripts for data preparation and descriptive analysis
 - `data_preparation.do`: prepares CSV inputs for MATLAB
 - `graphs_descriptive.do`: produces descriptive figures
- `matlab/`: MATLAB scripts for estimation, simulations, and output generation

Required (not included) directories:

- `csvdata/`: input data in CSV format
- `matlabdata/`: intermediate outputs
- `graphs/`: generated figures (including subfolders)
- `tables/`: generated LaTeX tables (including subfolders)

4. Software Requirements

The code requires:

- Stata (version 17 or above)
- MATLAB (with standard toolboxes)

Users should ensure that all directories are correctly set in the working paths before running the code.

5. Replication Workflow

Step 1: Data preparation and descriptive statistics (Stata)

Run `data_preparation.do` to construct the CSV input files used by MATLAB.

To reproduce the descriptive statistics figures, run `graphs_descriptive.do`.

Step 2: Main analysis (MATLAB)

Run `m0_main_all.m` to execute the full workflow, including model estimation, computation of standard errors, simulations, and generation of tables and figures.

6. Description of MATLAB Code

- `m0_main_all.m`: master script that orchestrates the full workflow and calls all main components of the analysis.
- `m1_*`, `m2_*`, etc.: core modules corresponding to major components of the empirical workflow.
- These scripts call a combination of additional `.m` files, including both high-level routines and lower-level functions.

The internal structure and dependencies of each module are documented in the comments within the respective files.

7. Outputs

The code produces:

- Figures, saved in the `graphs/` directory and its subfolders
- Tables (in LaTeX format), saved in the `tables/` directory and its subfolders

These outputs correspond to the results reported in the paper.