README for:
Inequality of Opportunity and Income Redistribution

Overview

The code in this replication package constructs the analysis files using data from our experiment on redistribution preferences. The analysis is conducted using Stata. One main code file runs all the codes necessary to clean and prepare the data for analysis, and another code file generates the figures and tables in the paper. The main analysis uses experimental data collected by the authors from participants in the Federal Reserve Bank of New York's Survey of Consumer Expectations (SCE).

Data Availability and Provenance Statements

☐ This paper does not involve analysis of external data (i.e., no data are used or the only data are generated by the authors via simulation in their code).

If box above is checked and if no simulated/synthetic data files are provided by the authors, please skip directly to the section on Computational Requirements. Otherwise, continue.

Statement about Rights

☒ I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

☐ I certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package. Appropriate permission are documented in the LICENSE.txt file.

Summary of Availability

☒ All data **are** publicly available.

☐ Some data **cannot be made** publicly available.

☐ **No data can be made** publicly available.

Details on each Data Source

The analysis in this paper relies on experimental data collected by the authors. The authors provide de-identified versions of the datasets sufficient for replication. Below are details on each of these datasets. The authors will assist with any reasonable attempts to replicate the analysis for two years following publication.

Our experiment consisted of two components: a worker task conducted in October 2021 with workers recruited from Amazon Mechanical Turk, and a spectator decision task conducted in November 2021 with participants recruited from the Federal Reserve Bank of New York's *Survey of Consumer Expectations* (SCE). The replication package includes the raw data files from both components.

1. The worker data files contain information about worker performance in the encryption task, their productivity multipliers, and whether they learned about their multiplier before or after completing the real-effort task:

    - `exante_2021-10-01.csv`
    - `expost_2021-10-01.csv`
    - `instructions_2021-10-01.csv`
    - `PageTimes-2021-10-01.csv`
    - `results_2021-10-01.csv`

2. The spectator data files contain the redistribution decisions made by spectators, their treatment assignment, their responses to our exit survey with political and social views, and demographic information:

    - `all_apps_wide_2021-11-03.csv`
    - `exit_2021-11-03.csv`
    - `experiment_ids_with_demographics.csv`
    - `instructions_2021-11-03.csv`
    - `PageTimes-2021-11-03.csv`
    - `rounds_2021-11-03.csv`
    - `rounds_2021-11-03_cf_implemented.csv`

After running the cleaning scripts (`clean-worker.do` and `clean-spectator.do`), two Stata datasets are generated: `worker.dta` and `spectator.dta`, which are then used for all subsequent analyses.

Computational requirements

Software Requirements

- Stata (code was last run with version 18)
    - `estout` (as of 2023-12-01)
    - `reghdfe` (as of 2023-12-01)
    - `lpoly` (as of 2023-12-01)

Memory and Runtime Requirements

Summary Approximate time needed to reproduce the analyses:

- ☒ <10 minutes
- ☐ 10-60 minutes
- ☐ 1-2 hours
- ☐ 2-8 hours
- ☐ 8-24 hours
- ☐ 1-3 days
- ☐ 3-14 days
- ☐ > 14 days
- ☐ Not feasible to run on a desktop machine, as described below.

Details The code that reproduces the tables and figures (`master.do`) runs in less than 5 minutes on a standard desktop computer with at least 8GB of RAM.

Description of programs/code

- `master.do`: This code executes all the files that clean the raw data and prepare the final analysis datasets. The codes that this file executes are organized in the following folders:
  - `clean`: Contains the files that clean and prepare the raw data
    - `clean-worker.do`: Cleans the raw data from the worker sample
    - `clean-spectator.do`: Cleans the raw data from the spectator sample
  - `results`: Contains the code that produces all figures and tables in the paper
    - `figures.do`: Creates the main figures in the paper
    - `tables.do`: Creates the main tables in the paper
    - `appendix-tables.do`: Creates the tables in the appendix
    - `appendix-figures.do`: Creates the figures in the appendix
    - `appendix-figure-B1`: Creates Appendix Figure B1 (requires R)
    - `appendix-figure-B2`: Creates Appendix Figure B2 (requires Matlab)

List of tables and programs

The provided code reproduces:

- ☐ All numbers provided in text in the paper
- ☒ All tables and figures in the paper

☐ Selected tables and figures in the paper, as explained and justified below.

Figure/Table	Program	Line(s)
Figure 1	results/figures.do	28-46
Figure 2, Panel A	results/figures.do	86-93
Figure 2, Panel B	results/figures.do	146-153
Figure 3, Panel A	results/figures.do	184-191
Figure 3, Panel B	results/figures.do	194-199
Figure 4	results/figures.do	234-250
Figure 5	results/figures.do	289-298
Figure 6, Panel A	results/figures.do	96-110
Figure 6, Panel B	results/figures.do	156-169
Table 1	results/tables.do	5-133
Table 2	results/tables.do	137-209
Table 3	results/tables.do	213-331
Appendix Tables	results/appendix-tables.do	various
Appendix Figures	results/appendix-figures.do	various

Acknowledgements

Some content on this page was copied from Hindawi. Other content was adapted from Fort (2016), Supplementary data, with the author's permission.