README and Guidance

Overview

The replication package includes all the programs necessary to derive the results in the paper and the online Appendix from raw data. The main code was run in Stata MP18 and in Matlab 2019b.

The replicator should expect the whole package to run for several days depending on their computer specification (main.m takes about 1 day per calibration; the Stata code runs for about 20-30 minutes).

The replication folder is organized as follows:

- *Data* contains: the raw data (Data/Raw); temporary data from the Matlab code (Data/Temp); the output from the Matlab code (Data/Output); and the output from the Stata code (Data/Final).
- *Programs* contains: the Matlab code (Programs/Matlab) with main.m running the whole Matlab code and calling the other files (calibrations with a prefix code_*, simulations with a prefix simulations_*, and IRFs); the Stata code (Programs/Stata) with Master.do running the whole Stata code and calling the other files (one do-file per exhibit, numbered as in the paper and online Appendix). Note that Programs/Matlab contains external functions (CEtools) and internal functions (Functions) in dedicated subfolders.
- *Figures* contains the produced figures in .png and in .eps format, following the numbering of the paper and the online Appendix.
- *Tables* contains the produced tables in .tex, following the numbering of the paper and the online Appendix.

Finally, the present *README* file is in the root folder (in Word format and in pdf).

Data Availability and Provenance Statements

Statement about Rights

We certify that the authors of the manuscript have legitimate access to and permission to use the data used in this manuscript.

Summary of Availability

All data are made publicly available.

Assistance

The authors will assist with any reasonable replication attempts for five years following publication. Please contact Yanos Zylberberg at yanos.zylberberg@gmail.com.

Dataset list

Data file	Source	Notes	Provided
`Data/Raw/` `pwt1001.dta`	Penn World Tables	Public	Yes

`Data/Raw/`
`WDI_Cons_Data.csv`
`WDI_Controls_Data.csv`
`WDI_Debt_Data.csv`
`WDI_GDP_Data.csv`
	World Development Indicators	Public	Yes

`Data/Raw/`
`TRates_IMF.txt`
`LTrates_OECD`
`BRates_IMF`
	IMF, OECD	Public	Yes

`Data/Raw/`
`DGE.txt`
`ILO.txt`
`MIMIC.txt`
	Informality Economy Database	Public	Yes

`Data/Raw/`
`LP_dataset`
`LP_dataset_born`
	Born et al., "Does austerity pay off?," Review of Economics and Statistics, 2020, 102 (2), 323–338.	Public	Yes

`Data/Raw` `Consodata`	OECD Consumption Data	Public	Yes

`Data/Final`
`data_evade`
`data_fc`
`data_final`
	Authors' calculations of tax compliance	Public	Yes

Computational requirements

Software Requirements

We conducted all the analysis using Stata MP Version 18 and Matlab 2019b. The code was last run on a Windows 11 operating system, and runtime was a bit more than one week.

Required packages

Stata: The program `Programs/Stata/Master.do` runs the whole cleaning and analysis procedure. It contains lines to install or verify that the following packages are installed: `outreg2`, `ranktest`, `ivreg2`.

Description of Programs

Programs that calibrate the baseline and counterfactual economies, simulate these economies, and prepare the impulse response functions are in `Programs/Matlab` and called by `main.m`.

Programs that produce all tables and figures in the paper are in `Programs/Stata` and called by `Master.do`.

The file `Programs/Stata/Master.do` sets the paths for all Stata do files, install possible dependencies, runs all the do-files, and generates the results used in all tables and figures.

Instructions to Replicators

- Adjust the main path in `Programs/Matlab/main.m`
- Run `Programs/Matlab/main.m` in Matlab
- Adjust the main path in `Programs/Stata/Master.do` and install dependencies if needed
- Run `Programs/Stata/Master.do`

Note that the Stata program run under `Programs/Stata/Master.do` requires the output of the Matlab program, but we have left the intermediary data in the Output subfolder such that these programs can be run independently.

List of tables and programs

The provided code reproduces all tables (Tables 1, 2, 3, B1, B2, B3, B4) and figures (Figures 1, 4, 5, 6, 7, B1, B2, B3) in the paper and online Appendix. The “theoretical” Figures 2 and 3 are directly produced in the text.