README for "Tax Policy and Abnormal Investment Behavior" Replication Package

For more information, please contact Qiping Xu or Eric Zwick.

This research was conducted while Zwick worked under an Intergovernmental Personnel Arrangement in the IRS Office of Research, Analysis, and Statistics under the SOI Joint Statistical Research Program. As a result, micro data is not available for public disclosure. This README describes the contents of the published replication package. A private replication package has been prepared and stored on Treasury servers and is available to researchers with permission to work with the data.

To access the internal replication package,[1] please contact Eric Zwick (`ezwick@chicagobooth.edu`) or John Guyton (`John.Guyton@irs.gov`), who supervised this project.

1 Files Included

The replication package contains:

1. This `README.pdf`
2. Five Stata dofiles: `master.do`, `make_kinky.do`, `programs_kinky.do`, `draft_kinky.do`, `appendix_kinky.do`
3. A note titled `model_replication_note.pdf` found in subdirectory */model replication* explaining the model replication Matlab code
4. Thirteen Matlab scripts in */model replication/code/vfi*: `vfi_full.m`,`vfi_ov.m`,`vfi_hycta.m`, `vfi_full_tax21.m`, `vfi_full_tax46.m`,`vfi_full_taxcredit.m`,`vfi_full_drift.m`, `vfi_full_inflation.m`, `vfi_full_depr06.m`, `vfi_ov_tax21.m`, `vfi_ov_tax46.m`, `vfi_full_depr18.m`,`vfi_full_beta09725.m`
5. Five Matlab scripts in */model replication/code/vfi*: `gendist.m`, `simul_full.m`, `simul_hycta.m`, `simul_ov.m`, `simul_did.m`

[1] Named RFS-XuZwick-2024-rep-package.zip. Available on the Stata server.

2 Code

Start with `master.do`, which is the parent dofile. Running this dofile will build the analysis datasets and reproduce all figures and tables from the paper. This dofile also organizes the other programs.

Each child dofile is structured as a set of programs with a `main` program at the bottom that calls the programs in order.

`make_kinky.do` – Builds main analysis data sets. Defines key analysis variables.

`programs_kinky.do` – Contains helper programs used to produce certain figures.

`draft_kinky.do` – Produces all figures and tables for main draft, excluding appendix.

`appendix_kinky.do` – Produces final graphs and tables for online appendix.

Additionally, the subdirectory */model replication* contains Matlab scripts to replicate the models in this paper and a note providing more details about the model code replication.

`model_replication_note.pdf` – Describes the code that solves the dynamic program with value function iteration and simulates firm investment behavior.

/model replication/code/vfi – Contains Matlab scripts that perform value function iteration to solve three versions of the investment model with alternative parameterizations.

/model replication/code/simul – Contains Matlab scripts that use optimal policy functions to simulate firms' capital paths for each of the three versions of the model.

3 Data

The input datasets utilizing external data are as follows:

1. Compustat data were accessed via Wharton Research Data Services (WRDS). To access these data, one needs to create an account at https://wrds-www.wharton.upenn.edu and have an institutional or personal subscription to the service. These are the Compustat input datasets used to build the main analysis datasets: `anncomp.dta`, `compustat_sp_rating.dta`, `compustat_company.dta`, `compustat_segmerged.dta`, `compustat_seg_customer.dta`, `funda.dta`, `fundq.dta`, `internationala.dta`, `internationalq.dta`
2. IBES data were also sourced from WRDS: `statsum_epsus.dta`
3. Orbis data: `orbis_layer.dta`

The datasets built utilizing the aforementioned inputs are as follows:

1. Preliminary datasets: `sicsegment.dta, exeownership.dta, accountingy.dta, sp_rating.dta, inventory.dta, quarterlydata.dta, capxf1.dta`
2. Merged data from preliminary datasets: `mergeddata.dta`
3. Main analysis datasets: `capxannual.dta, capxannual_p.dta, capxannual_trim5.dta, capxannual_trim5pre83.dta, international_trim5.dta`

The datasets for the cumulative effect regression are as follows:

1. Quarterly ppent data to serve as denominator: `ppentq.dta`
2. Event time: `q4event.dta`
3. Cumulative regression data: `cumulative.dta`

Please see the data section of the paper for documentation regarding the source of raw data inputs used to build the following datasets:

1. BLS_PPI: `industrycode.dta, ppinaics6.dta`
2. M3_Survey: `naics.dta, sic.dta`
3. RateWatch: `interestrate.dta, commequp250k.dta, commoper50k.dta, commreal1mil.dta, personalunsec.dta, buslnsec50k.dta`
4. RD: `rddecomp.dta`
5. leases_and_loans: compiled via `http://www.elfaonline.org/data/MLFI`

All internal data are stored on IRS servers and available to researchers with permission to work with it. For licensed datasets that cannot be shared or model output, we provide header datasets to illustrate the structure.

4 Program

The following STATA programs are needed:

1. *ffind*: You can download it from: https://sites.google.com/site/judsoncaskey/data
2. *binscatter*, *listtex*, *texsave*, *reghdfe*