Read Me "Occupations and Import Competition"

Sharon Traiberman

July 6, 2019

Data Cleaning/Reduced Form Code

This code is only available through access to Statistics Denmark. Upon receiving access, one can receive copies of my code by contacting the Data Manager at Aarhus University. All of my code is freely available for any replication, or future use. This documentation outlines the files used, with some detail since not all code and data are public. The is a file `main.do` which runs ALL files INCLUDING MATLAB (through shell). This file MUST be run as it contains PATH DEFINITIONS. The code must be run on a machine with access to 4 cores for bootstrapping (or the parfor loop must be manually changed in Matlab). File headers and the code proper contain additional comments. Email me at sharon.traiberman@gmail.com for troubleshooting.

Cleaning:

1. `sasExtraction.zip` – a zip folder of SAS files that extracts SAS views to dta files. All cleaning is done in Stata.

2. `createLEEDPanel.do` – code that reads in all raw data after converting from SAS and does cleaning WITHOUT changing the sample. Aside from reading the data, the key tasks performed are: [1] linking employees to employers and extracting industry affiliation; [2] constructing education categories; [3] concording industry classifications over time; [4] deflating all nominal values to 2000 DKK. Final product here is [1] worker panel (for bulk of analysis), [2] trade data (for calculating offshoring), [3] firm level data (for calculating labor shares in the counterfactual)

3. `newOccCleaning.do` — main code for all definitions, cleaning and sample selection. Everything this code does is contained in the data appendix. Result of this file is the final panel for analysis containing information on occupation/sector/employment, tenure, income, age, and education.

Summary Stats:

1. `tradeExercises.do` – Figure 4 and 6

2. `summaryStatTables.do` – Figure 1-3, Table 1-3

3. `acmEstimation.do` – columns 2-5 in Table 5

Estimation Code [Supply Side]

Stata:

1. `makeMatlabFiles_FullLL.do` – This file converts the stata files to matlab CSVs for doing the first stage (EM algorithm)

2. `makeAfterMatlab.do` – Table 5-6, all appendix tables

3. `postEstimation.do` (PUBLICLY RELEASED) – Extracts parameter estimates and converts them into format for Matlab. Also makes tex format for results tables (appendix A) and Table 5-6.

4. `makeCAtable.do` (PUBLICLY RELEASED) – Table 4 and Figure 5

Matlab:

Bootstrapping:

1. `makeBootstrapNew.do` – Creates bootstrap files. WARNING: Stata seed is set, but not sortseed. Exact numerical duplication is not guaranteed, but discrepancies are insignificant. See: https://www.stata.com/help13.cgi?sortseed

2. `postBootstrapNew.do` – reads Matlab output back to Stata to calculate SEs

Calibration Code [Demand Side]

Run files in this order. Documentation for each file is in the header of each file. The file `allCalibration.do` contains [1] path declarations (YOU MUST RUN THIS BEFORE RUNNING THE FILES!) and [2] runs all the files in the correct order. Documentation is sparser in the Readme as these files are publicly available.

1. `makeEDF.do` – uses a smoothed (due to data censoring issues) EDF of the characteristics of the Danish population in 1996 to construct a weighted grid of initial points for simulation

2. `makeIOTable.do` – converts Danish IO tables to input-output matrices for Matlab

3. `prodcom` & `makeForeignPrices_wits` – constructs import price indices using two different sources (for comparison)

4. `demandParameters.do` & `makeIncomeLevels.do` – uses [publicly available] IO tables AND [proprietary but released] aggregated data on occupational labor shares to construct Cobb-Douglas shares, export demand, total revenues, and production coefficients.

5. `makeMatlabMatrices.do` – Takes output above and rewrites everything for ease of use by Matlab. Importantly: recodes industries to be indexed 1-N, same with occupations, etc.

Counterfactuals

Run files in this order. Documentation for each file is in the header for each file. Documentation is sparser in the Readme as these files are publicly available.
Stata:

1. `postEstimation.do` – Extracts parameter estimates and converts them into format for Matlab. Also makes all results tables (appendix A), Table 5 (column 1) and table 6. This file is previously mentioned in the cleaning/estimation section.

Matlab:
The file `runAllFiles.m` will run all files. This file ALSO runs counterfactuals for appendix I and for the Offshoring Counterfactual (both available online).

1. `steadyStateCapital.m` - Myopic steady state for initial guess

2. `steadyState_pf_cap.m` - Iterates from the initial equilibrium (1996) to a steady state assuming a fixed stock of capital and perfect foresight

3. `actualTransitionCapital.m` - Solves the transition from the initial steady state to the post-trade steady state forward assuming myopia (this is used as an initial guess for the perfect foresight equilibrium as it is very fast to compute this)

4. `cfTransitionCapital.m` - Solves the transition from the initial steady state onward for 40 more periods

5. `actualTransition_pf_cap.m` - Uses the myopic guess to solve for the transition dynamics from the initial steady state to a new equilibrium

6. `counterfactualTransition_pf_cap.m` - Solves from the initial steady state going forward assuming no change in variables[1]

7. `postCounterfactualSim.m` - Figure 7 + Table 8, also creates simulated histories for creation of Figure 8 and Table 9

Auxiliary code called:

- `readInDemandData.m` - Reads in calibrated parameters for demand side. This code MUST BE EDITED BY USERS BEFORE USING. It contains PATH DECLARATIONS.

- `readInParameters.m` - Reads in estimated parameters from supply side. This code MUST BE EDITED BY USERS BEFORE USING. It contains PATH DECLARATIONS.

- `getContinuationValues_n` - Given a guess of wages and price of capital one can calculate from cost minimization the full path of real prices which, paired with initial conditions of labor, allows for calculation of the full path of labor supply decisions and all continuation values

- `excessDemandCapital` - Calculates labor excess demand for myopic model (period-by-period)

- `excessDemandCapital_pf` - Calculates labor excess demand for perfect foresight model (all periods simultaneously)

---

[1]See paper discussing why this is necessary. Also useful for moving other parameters - productivity, capital prices, etc.