README

Setup

Software

- The empirical analysis using the SOEP and SIAB data was conducted using Stata 15. Details on how to obtain access to the micro data for replication purposes can be found below in the Data Availability section.
- The remaining computational analysis was conducted using Julia 1.10.7 (estimation) and Julia 1.10.10 (analysis). The Project.toml and Manifest.toml files can be used to work with the same Julia environment and package versions.
- All figures are stored as PDF files and converted from SVG files using the free software inkscape (https://inkscape.org/).

Computational Setup

Large parts of the analysis were conducted on a Linux high-performance cluster of the "Leibniz Rechenzentrum der Bayerischen Akademie der Wissenschaften" (LRZ). The specific partition of the LRZ Linux cluster I used (cm4_std) has 4 nodes with 112 physical cores and 488 GB memory limit per node. Linux version: 6.4.0-150600.23.25-default.

The file `lrz_cluster_linux_info.txt` in this replication folder shows the server specs. The files `lrz_cluster_julia_version.txt`, and `lrz_cluster_julia_version_1_10_X.txt` contain information on the Linux system and Julia versions used.

The tables and figures were created on a notebook (Windows 10).

Data Availability

The analysis uses micro the following micro data:

- German Socio-Economic Panel (SOEP), data for years 1984–2017, SOEP-Core v34, doi:10.5684/soep.v34.
- Sample of Integrated Labour Market Biographies (SIAB) 1975 – 2017, version 7517_v1.

I am not allowed to share these data due to strict provisions of German data protection law. However, both datasets can be accessed through the respective institutions. Information on data access are provided below. All other data sets are included in the replication kit.

SIAB Data

The data analysis of this paper is based on the weakly anonymous Sample of Integrated Labour Market Biographies (SIAB) 1975 - 2017. The SIAB data were accessed on-site at the Institute for Employment Research (IAB) during my time as an employee at the IAB. They can also be accessed through the Research Data Centre (FDZ).

Data documentation: Antoni, Manfred; Schmucker, Alexandra; Seth, Stefan; vom Berge, Philipp (2019): Sample of Integrated Labour Market Biographies (SIAB) 1975-2017. FDZ-Datenreport, 02/2019 (en), Nuremberg.

Data access for replication access purposes can be obtained via the IAB. For further information, see: https://iab.de/en/facts-and-figures-2/data-access-for-replication-purposes/ Please contact the research data center of the IAB: iab.fdz@iab.de

SOEP Data

Data are available from the German Socio-economic Panel Study (SOEP) due to third party restrictions (for requests, please contact soepmail@diw.de). The scientific use file of the SOEP with anonymous microdata is made available free of charge to universities and research institutes for research and teaching purposes. The direct use of SOEP data is subject to the strict provisions of German data protection law. Therefore, signing a data distribution contract is a precondition for working with SOEP data. The data distribution contract can be requested with a form, available at: http://www.diw.de/soepforms. For further information, contact the SOEP Community Management at either soepmail@diw.de or +49-30-89789-292.

Data documentation: https://www.diw.de/en/diw_01.c.738729.en/edition/soep-core_v34__data_1984-2017.html

Replication Guide

Data Analysis using SOEP Data

Use the do-files in the folder `soep/do` to create the files in the folders `soep/results` and `soep/soep_results`. You have to adjust the path to the replication kit in `master_soep.do`, and then execute the do-file. To use the files for calibrating the model, copy the file `results/tax_function.csv` to `SearchMatching.jl/analysis/data/soep/`.

Data Analysis using SIAB Data

Copy the folder `soep/soep_results` to the folder `orig` on the IAB server as inputs to the data analysis using the SIAB data. Also copy the files in `iab/prog` to the `prog` folder on the IAB server. Based on the confidential SIAB data and the SOEP results, these do-files reproduce the files in the folder `iab/results`. Before executing `00_master.do`, you have to adjust the path to adjust the paths depending on your setup at the IAB research data center. To use the files for calibrating the model, copy all csv-files stored in `results` to `SearchMatching.jl/analysis/data/iab/`.

Calibration of Model Parameters

This step was conducted on a high-performance cluster (see above). Execute the batch file `SearchMatching.jl/analysis/script_estimation.sh`. The code generates the files in the folders

- `SearchMatching.jl/analysis/results/`
- `SearchMatching.jl/analysis/results_r75/`
- `SearchMatching.jl/analysis/results_r80/`

Model Analysis

This step was conducted on a high-performance cluster (see above). Execute `SearchMatching.jl/analysis/script_results.sh` on the high-performance cluster. The script generates the files in the folders

- `SearchMatching.jl/analysis/results/results/`
- `SearchMatching.jl/analysis/results_r75/results/`
- `SearchMatching.jl/analysis/results_r80/results/`

Create Figures and Tables

This step can be done on a notebook. Execute the file `julia_model/main.jl` to generate all the numbers, figures and tables found in the paper. The files are stored in `results`.

Folder Structure

- `iab/`
    - `prog/` contains Stata do-files to analyze restricted-access SIAB data and produce the .csv files containing target moments for the calibration
    - `results/` contains the target moments stored in .csv files
- `julia_model/` contains Julia jl-files to reproduce the results in the paper and Online Appendix. The jl-files access the final data in `SearchMatching.jl/analysis/results/`, `SearchMatching.jl/analysis/results_r75/`, and `SearchMatching.jl/analysis/results_r80/`. To replicate the results stored in `results/`, execute `main.jl`.
- `other_data/` contains unristricted-access data
- `results/` contains the figures, tables and selected numbers shown in the paper and Online Appendix
    - `figures\` contains the figures reported in the paper and Online Appendix
    - `numbers\` contains the numbers reported in the paper and Online Appendix
    - `tables\` contains the tables reported in the paper and Online Appendix
- `SearchMatching.jl/`
    - `analysis/` contains batch files to execute Julia scripts that calibrate the model and run counterfactual analyses to produce the final data for the tables and figures
        - `data/iab/` contains the data target csv files generated from the IAB data
        - `data/soep/` contains the tax function parameters computed from the SOEP data
        - `results/` contains the calibrated baseline model and the counterfactual results (in `results/results/`)

        - `results_r75/` contains the re-calibrated model with a piece-rate of 0.75 and the counterfactual results (in `results_r75/results/`)
        - `results_r80/` contains the re-calibrated model with a piece-rate of 0.80 and the counterfactual results (in `results_r80/results/`)
    - `src/` contains the model Julia source code to run and evaluate the search-and-matching model
    - `SimulatedMoments.jl/` contains Julia source code to calibrate the model
- `soep/`
    - `do/` contains Stata do-files to analyze the restricted-access SOEP and produce intermediate results needed for the analysis of the SIAB data and some final data targets to calibrate the model
    - `results/` contains output data that is not needed as input ot the IAB analysis
    - `soep_results/` contains output data that is needed as input ot the IAB analysis. This folder needs to be copied to the IAB server.