README

Replication materials for:
"Hours and Wages"
by Alexander Bick, Adam Blandin, and Richard Rogerson
Accepted, Quarterly Journal of Economics

1 Overview

This document describes the data and code used to replicate the results in the paper. Required software programs are Stata and Python. There are also data files in .xls and .csv format that can be viewed using Excel or a similar compatible program. Runtimes correspond to a Dell Precision Desktop with 32 GB of RAM and an Intel Xeon W-2195 CPU 2.3Ghz (12 cores) running Windows 10 Enterprise.

The data output was created using Stata/MP version 15.1.

- To reproduce all empirical output in this paper, execute the file "empirics/master.do" in Stata.
  - Line 4: set the path to the empirics folder
  - Line 6: if run_data is set to 1, the raw data files (e.g., the CPS ORG from IPUMS) are read in and prepared for the final analysis. With this hardware, this part of the code takes about 50 minutes. The folder empirics/data already includes the data files for the final analysis.
  - Line 7: if run_results is set to 1, all figures and tables in the paper and online appendix are reproduced. With this hardware, this part of the code takes about 25 minutes.
- We are not allowed to share the Danish micro data from the European Community Household Panel (ECHP) needed to reproduce Figure 11.
  - The ECHP can be easily obtained for any researcher whose institution has a data sharing agreement with Eurostat. If such a data sharing agreement does not exist, it needs to be set up at the institution level first before individual researchers can request the data.
  - Once the ECHP data are obtained, place the raw data in the folder empirics/datafiles_raw_echp, set run_data to 1 in line 4, replace line 25 with "local dss echp" in empirics/master.do, and run that part of the code. Next run the dofile for figure 11 (dofiles/results/fig_11.do).

Model simulations were conducted using Python 3.9.5. Python is an open source programming language that can be downloaded at no cost from `https://www.python.org/downloads/`.

- To reproduce all results based on model simulations in this paper, first create a Virtual Python Environment in the subfolder "replication/model" using the "requirements.txt" file from that folder. Then execute the file "replication/model/main.py".

- With the hardware described above, "replication/model/main.py" runs in about 8 hours. The vast majority of this runtime is spent on the simulation for models M1 and M2 in order to create Figures 7 and 9 and the paper. These models require a large number of simulations in order to generate consistent estimates.
- If you are not interested in the simulation results for M1 and M2 in Figures 7 and 9 in the paper, you can change the simulation size n_sim from 15 million to a lower value (in line 40 of the file "replication/model/fig_6_7_8_9_tab_3_4.py"), which will substantially speed up the runtime. For example, setting the simulation size to 1 million reduces the runtime to less than one hour.