Replication Data for "Automation and Polarization"

Daron Acemoglu **Jonas Loebbing**

May 29, 2025

In this document, we explain how to replicate the quantitative analysis in our article "Automation and Polarization." The replication produces all figures and tables from Section 6 and online appendix B.5.

For a detailed description of the data sources, see online appendix B.5.2; for an explanation of our numerical approach to solve the model, see online appendix B.5.1; and for an explanation of how we calibrate the model, see online appendix B.5.3.

1 Replication Instructions

The replication proceeds in three steps:

1. Extract all wage-related data moments (log wage levels and log wage changes) for the calibration from the data provided by Acemoglu and Autor (2011) and Acemoglu and Restrepo (2022)
    - Open the Stata do-file "Wage Percentiles and Effects.do" from the folder "Analysis/Wage Data Analysis"
    - Set the working directory in line 8 to the folder "Replication AL 2025" and execute the code
    - This produces:
        - the files "wage.csv" and "wage-data.dta" in the folder "Output/Wage Data"
        - the figure "AR22-fig7D.gph" in the folder "Output/Figures" (note: this figure is not used in the paper, it is just a check for comparison with Figure 7D in Acemoglu and Restrepo 2022)
        - figures "fig-b1.gph" and "fig-b1.pdf" in the folder "Output/Figures" (Figure B-1 of the paper)
2. Calibration
    - Open the MATLAB script "AutomationPolarization_calibration.m" from the folder "Analysis/Calibration"
    - Set the variable "basefolder" in line 11 to the path of the folder "Replication AL 2025" and execute the code

- This produces:
  - the files "calibration.mat", "calibrationgrid.mat", "calibrationlocalsearch.mat", and "calibrationmw.mat" in the folder "Output/Calibration Results"

3. Compute all counterfactuals and produce all figures and tables
   - Open the MATLAB script "AutomationPolarization_main.m" from the folder "Analysis/Simulations"
   - Set the variable "basefolder" in line 9 to the path of the folder "Replication AL 2025" and execute the code
   - This produces:
     - the files "counterfactual-capprod.mat", "counterfactual-curvature.mat", "counterfactual-laborsupply.mat", "counterfactual-mw-new.mat", "robustness-lambda.mat", "robustness-q.mat" in the folder "Output/Simulation Results"
     - all tables of the paper in the folder "Output/Tables"
     - all figures of the paper in the folder "Output/Figures", except for Figure B-1, which is produced in step 1 of the replication (see above)

Each step can be performed in isolation. To only produce the figures and tables of the paper – without replicating the counterfactual simulations – you can perform a shortened version of Step 3:

A. Produce only figures and tables
   - Open the MATLAB script "AutomationPolarization_main.m" from the folder "Analysis/Simulations"
   - Set the variable "basefolder" in line 9 to the path of the folder "Replication AL 2025"
   - Set the variable "option" to "basic" in line 16 and execute the code
   - this produces all tables of the paper (in "Output/Tables") and all figures except for Figure B-1 (in "Output/Figures")

2 Technical Setup

All computations for the paper were performed on a machine with the following specifications:

- Software:
  - OS: Windows Server 2022 Standard, 64-bit
  - MATLAB R2024a, with Parallel Computing Toolbox
  - Stata/MP 19.0 (Revision 08 Apr 2025)

- Hardware:
  - CPU: 2 x Intel Xeon Gold 6346 (32 cores, 3.1 GHz)
  - RAM: 1 TB

Computing times for the replication steps described above were as follows:

- Step 1: < 5 seconds
- Step 2: approx 40 hours
- Step 3: approx 15 minutes
- Step A (shortened version of Step 3, see above): 18 seconds

3 Structure of the Replication Data

1. Analysis
   (a) Calibration
      - contains all MATLAB scripts and functions for the calibration of the model
      - the script "AutomationPolarization_calibration.m" is the main script that performs the entire calibration and saves the results in "Output/Calibration Results"
      - all other functions in the folder are called from within the main script

   (b) Simulations
      - contains all MATLAB scripts and functions for the counterfactual simulations and the production of all figures and tables of the paper (except for Figure B-1)
      - the script "AutomationPolarization_main.m" is the main script that performs all counterfactual simulations, produces all figures and tables, and saves the results in "Output/Simulation Results", "Output/Figures" and "Output/Tables", respectively
      - all other functions in the folder are called from within the main script

   (c) Wage Data Analysis
      - contains the Stata do-file "Wage Percentiles and Effects.do", which extracts the data moments in Table 1.B – except for the income share of equipment and its change from 1980 to 2016 – from the datasets provided by Acemoglu and Autor (2011) and Acemoglu and Restrepo (2022), and saves the results in "Output/Wage Data"

2. Data
   (a) Capital Shares and User Costs

- "Acemoglu Loebbing Capital Shares and User Cost.xlsx" computes the income share of equipment and its change from 1980 to 2016 used as data moments for the calibration (see Tables 1.B and 2) and the user cost indices of equipment, software, and equipment plus software that we use to compute the model moments in our calibration that involve changes between 1980 and 2016 (see Tables 1.B. and 2)
- all other files in the folder contain the raw data as downloaded from the respective sources (Bureau of Economic Analysis and Federal Reserve Economic Data) that we use in "Acemoglu Loebbing Capital Shares and User Cost.xlsx"

(b) Minimum wage

- "Acemoglu Loebbing Minimum Wages.xlsx" derives the minimum wages used to compute the counterfactual minimum wage increase in Figures 9 and B-4 and to perform the calibration in online appendix B.5.7
- the remaining files in the folder contain raw data as downloaded from the Bureau of Economic Analysis that is used in "Acemoglu Loebbing Minimum Wages.xlsx"

(c) Wages

- contains the Stata dataset "morg-pctile-mov-1974-2008.dta" from the replication files of Acemoglu and Autor (2011), with data on the US 1980 wage distribution, from which the do-file "Wage Percentiles and Effects.do" (see above) extracts the data moments for our calibration that pertain to log wage levels in 1980 (see Table 1.B)
- and the Stata dataset "AR22_geresults.dta" from the replication files of Acemoglu and Restrepo (2022), with their estimates of the effects of automation on wages in the US between 1980 and 2016, from which the do-file "Wage Percentiles and Effects.do" (see above) extracts the data moments for our calibration that pertain to changes in log wages from 1980 to 2016 (see Table 1.B)

3. Output

(a) Calibration Results

- contains the MATLAB datasets with the results of our calibration as produced by the script "AutomationPolarization_calibration.m" (see above)
- "calibration.mat" contains the results of the baseline calibration described in Table 1; "calibrationmw.mat" contains the results of the calibration with minimum wage of online appendix B.5.7; "calibrationgrid.mat" contains the results from the grid search in the first step of our baseline calibration; "calibrationlocalsearch.mat" contains the results from the local minimization the starts from the results of the grid search

(b) Figures

- contains all figures of the paper, as produced by the do-file "Wage Percentiles and Effects.do" (AR22-fig7D.gph, fig-b1.gph, fig-b1.pdf) and the script "AutomationPolarization_main.m" (all other figures)

(c) Simulation Results

- contains the MATLAB datasets with the results of our counterfactual analyses and robustness checks as produced by the script "AutomationPolarization_main.m"
- "counterfactual-capprod.mat" contains the results of the counterfactual reduction in capital cost in Figure 7; "counterfactual-curvature.mat" contains the results of the artificial intelligence counterfactual in Figure 8; "counterfactual-laborsupply.mat" contains the results of the counterfactual low-skill supply reduction in Figure 9; "counterfactual-mw-new.mat" contains the results of the counterfactual minimum wage increase in Figure 9; "robustness-lambda.mat" contains the results of the robustness checks with respect to the elasticity of substitution between tasks in Figure 6 and Table 2; "robustness-q.mat" contains the results of the robustness checks with respect to the capital cost index in Figure 6 and Table 2

(d) Tables

- contains all tables of the paper as produced by the script "AutomationPolarization_main.m"

(e) Wage Data

- contains two datasets (Stata and csv) with US log wage levels in 1980 and changes in log wages due to automation between 1980 and 2016, which are extracted from the replication files of Acemoglu and Autor (2011) and Acemoglu and Restrepo (2022) by the do-file "Wage Percentiles and Effects.do" and then used by the script "AutomationPolarization_calibration.m" for our calibration

References

Acemoglu, Daron and David H. Autor (2011) "Skills, Tasks and Technologies: Implications for Employment and Earnings," in *Handbook of Labor Economics*, 4, 1043–1171: Elsevier.

Acemoglu, Daron and Pascual Restrepo (2022) "Tasks, Automation, and the Rise in US Wage Inequality," *Econometrica*, 90 (5), 1973–2016.