



Can the Flying Geese Fly Again? Automation and the Pattern of Industrialization

First Submission: RR_WLD_2026_755

Maria Reyes Retana Torre, Nihaa Sajid

reproducibility@worldbank.org

September 23, 2026

This review verifies the reproducibility of the exhibits included in the paper “*Can the Flying Geese Fly Again? Automation and the Pattern of Industrialization*”.

Contents in this review:

1. Main findings
2. List of exhibits and reproducibility status
3. Reproduction Environment

Main findings

- **Every exhibit of the paper has been reproduced accurately.**
- The code was successfully executed on a new computer after:
 1. Running S1_RUN_STATA.do after changing the path in the do file.
 2. Running S2_RUN_MATLAB after changing the path.
 3. Running S3_RUN_STATA.do after changing the path.
 4. Running Opening the R project, restoring the environment, and running S4_RUN_R.R.
- The output demonstrates consistent stability across multiple runs. Specifically, executing the code two times consecutively yielded identical results.
- The code takes approximately 2 hours to run.
- We conducted our reproducibility analysis based on the paper shared by the authors.
- **Verification Process and Data Handling:**
 - The reproducibility package relies on 2 types of data: open and restricted data (available only from the data owners).
 - Open data is included in the public reproducibility package.
 - Reviewers used restricted data provided directly by the authors to conduct the reproducibility verification, and this is not included in the public reproducibility package.
 - *data_hash_report.csv* lists the SHA256 hashes of all files in the Data folder to support data integrity checks. Users who acquire the forthcoming/accessible/limited-access/restricted data can use this file to verify that the data has not been altered.

- **Reproducibility Summary:**

- **Data:** Some data is restricted and not included in the reproducibility package.
- **Code:** All code files (from cleaning to analysis) are included in the reproducibility package.
- **Outputs:** All outputs are generated by code included in the reproducibility package.
- **Dependencies environment:** The reviewers created a new environment using the latest versions of dependencies available at the moment of the review.

List of exhibits and reproducibility status

Results in the Main Section of the Paper

- **Figure 1A Reproduced.** This exhibit was verified against *fig_II12c_keycountries_trend.png*.
- **Figure 1B Reproduced.** This exhibit was verified against *fig_II12d_keycountries_2020.png*.
- **Figure 2 Reproduced.** Differences between the axes in the manuscript and the reproduced output. This exhibit was verified against *Fig2_MatchDataRobotsCountryMean*.
- **Table 1 Reproduced.** This exhibit was verified against *table_1.tex*.
- **Table 2 Reproduced.** This exhibit was verified against *table_2.tex*.
- **Figure 3A Reproduced.** Differences between the x-axis in the manuscript and the reproduced output. This exhibit was verified against *FigA3a_wage_population.png*.
- **Figure 3B Reproduced.** This exhibit was verified against *FigA3b_omega_psi.png*.
- **Figure 4A Reproduced.** This exhibit was verified against *Fig3a_prod_3D.png*.
- **Figure 4B Reproduced.** This exhibit was verified against *Fig3b_prod_sectors.png*.

Results in the Annex

For the Appendix, we did not review every exhibit. Instead, we randomly selected 10 exhibits from the remaining datasets to assess the appendix. Our review was based on those 10 exhibits. Since they were chosen randomly, we are operating under the assumption that if all randomly selected exhibits are reproducible, then the rest should be as well. The seed used to generate the random selection was [494092], the Stata version used was [18.5], and the exhibits selected were: Table A2, Figure A1, Table A3, Figure A2, Table A5, Table A6, Table A7, Table A8, Table A9 and Table A10

- **Table A2 Does not Apply** The exhibit does not show analysis.
- **Figure A1 Reproduced.** Differences between the x-axis in the manuscript and the reproduced output. This exhibit was verified against *FigA1_automated_tasks_elas.png*.
- **Table A3 Does not Apply** The exhibit does not show analysis. Reproduced.
- **Figure A2 Panel A Reproduced.** This exhibit is verified against *fig_II12a_region_trend.png*.

- **Figure A2 Panel B Reproduced.** This exhibit is verified against *fig_II12b_region_2020.png*.
- **Table A5 Reproduced.** This exhibit is verified against *table_1a.tex*.
- **Table A6 Reproduced.** This exhibit is verified against *table_2a.tex*.
- **Table A7 Reproduced.** This exhibit is verified against *table_1_chn.tex*.
- **Table A8 Reproduced.** This exhibit is verified against *table_2_chn.tex*.
- **Table A9 Reproduced.** This exhibit is verified against *table_1a_chn.tex*.
- **Table A10 Reproduced.** This exhibit is verified against *table_2a_chn.tex*.

Reproduction Environment

Paper exhibits were reproduced on a computer with the following specifications:

- OS: Windows 11 Enterprise
- Processor: INTEL(R) XEON(R) PLATINUM 8562Y+ 2.80 GHz (2 processors)
- Memory available: 32.0 GB
- Software version: Stata 18.5 MP, Matlab 2024a, R 4.6.1