



Labor Demand in the Age of Generative AI: Early Evidence from the U.S. Job Posting Data

First Submission: RR_USA_2025_436

*Luis Eduardo San Martin
reproducibility@worldbank.org
October 18, 2025*

This review verifies the reproducibility of the exhibits included in the paper “*a really cool title*”.

Contents in this review:

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

Main findings

- **Every exhibit has been reproduced accurately in the authors’ Databricks workspace.**
- The reproducibility verification was performed via a virtual meeting on October 3, 2025, as some of the restricted-access data could not be shared with reviewers. The code outputs and logs were sent to the reviewers afterward to complete the reproducibility verification.
- The code takes approximately 24 hours to run.
- We conducted our reproducibility analysis based on the paper shared by the authors via OneDrive on October 3, 2025.
- **Verification Process and Data Handling:**
 - The reproducibility package relies on 3 types of data: open, accessible (public but not included in the reproducibility package), and restricted data (available only to the authors).
 - 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 has not been included in the reproducibility package. For more details, please refer to the README file.
 - **Code:** All code files (from cleaning to analysis) are included in the reproducibility package.

- **Outputs:** One output is not generated by code and is created manually in Excel. Instructions for producing this output are included in the README file.
- **Dependencies environment:**
 - * **Stata:** The reviewers created a new environment using the latest versions of dependencies available at the moment of the review.
 - * **R and Python:** The virtual reproducibility verification used the same versions of dependencies installed in the authors' Databricks workspace by the date of the review.

List of exhibits and reproducibility status

Results in the Main Section of the Paper

- **Figure 1** Reproduced.
- **Figure 2** Reproduced.
- **Figure 3** Reproduced.
- **Table 1** Reproduced.
- **Figure 4** Reproduced.
- **Figure 5** Reproduced.
- **Figure 6** Reproduced.
- **Figure 7 (a)** Reproduced.
- **Figure 7 (b)** Reproduced.
- **Figure 7 (c)** Reproduced.
- **Figure 7 (d)** Reproduced.
- **Figure 7 (e)** Reproduced.
- **Figure 7 (f)** Reproduced.
- **Figure 8** Reproduced.
- **Figure 9 (a)** Reproduced.
- **Figure 9 (b)** Reproduced.
- **Figure 9 (c)** Reproduced.
- **Figure 10** Reproduced.
- **Figure 11** Reproduced.

Results in the Annex

- **Figure A1** Reproduced.
- **Figure A2** Reproduced.
- **Figure A3** Reproduced.
- **Figure A4** Reproduced.
- **Figure A5** Reproduced.
- **Figure A6** Reproduced.
- **Table A1 - By Industry** Reproduced
- **Table A1 - By Education Attainment Required** Reproduced after adding the rows *Master* and *Doctorate* to conform the row *Advanced* of the paper exhibit.
- **Table A1 - By Years of Experience Required** Reproduced after adding the rows *6-9 years* and *10+ years* to conform the row *6+ years* of the paper exhibit.
- **Table A2** Reproduced
- **Table A3** Reproduced

Reproduction Environment

Paper exhibits were reproduced on two systems with the following specifications:

- Computer 1 - Stata:
 - OS: Windows 11 Enterprise, version 24H2
 - Processor: Intel(R) Core(TM) Ultra 7 165U (2.10 GHz)
 - Memory available: 32 GB
 - Software version: Stata 18.5 MP-Parallel Edition
- Databricks:
 - Runtime: 16.4, Spark 3.5.2, Scala 2.12
 - Driver: Standard_D16ds_v4, 64GB, 16 Cores
 - Worker: Standard_D16ds_v4, 64GB, 16Cores
 - Software version:
 - * R: 4.4.0
 - * Python: 3.12.3