Reproducible Research Repository

Beyond green jobs: Advancing metrics and modelling approaches for a changing labor market

First Submission: RR_WLD_2025_477
Luis Eduardo San Martin
reproducibility@worldbank.org
February 20, 2026

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

Contents in this review:

Main findings

- **Every exhibit of the paper has been reproduced accurately.**
- The code was successfully executed on a new computer after:
  1. Reproducing the Python environment with the file "*envir_bg.yml*".
  2. Running the code `python -m ipykernel install -user -name envir_bg` in the new environment.
- The output demonstrates consistent stability across multiple runs. Specifically, executing the code two times consecutively yielded identical results.
- The code takes approximately 20 minutes to run.
- We conducted our reproducibility analysis based on the paper shared by the authors via OneDrive on October 30, 2025.
- **Verification Process and Data Handling**:
  - The reproducibility package relies on four types of data: open, limited-access (available after account approval), forthcoming (to be published by the dataset authors in the future), and restricted data (available only to the authors).
  - Open data are included in the public reproducibility package.
  - Reviewers used limited-access, forthcoming, and restricted data provided directly by the authors to conduct the reproducibility verification, and these are not included in the public reproducibility package. The forthcoming data will be published in the World Bank Development Data Hub.
  - *data_hash_report.csv* lists the SHA256 hashes of all files in the Data folder to support data integrity checks. Users who acquire the accessible and limited-access 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.
  - **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 reproduced an existing environment for dependencies using an environment metadata file provided by the authors.

List of exhibits and reproducibility status

Results in the Main Section of the Paper

- **Table 1** Does not apply. The exhibit does not show analysis results.
- **Table 2** Does not apply. The exhibit does not show analysis results.
- **Table 3 - Different methodologies for calculating green job numbers** Does not apply. The exhibit does not show analysis results.
- **Figure 1** Reproduced.
- **Figure 2** Reproduced.
- **Figure 3** Does not apply. The exhibit reports findings from a different paper.
- **Figure 4 - Regional and occupational worker mobility patterns in Brazil** Does not apply. The exhibit reports findings from a different paper.
- **Table 3: Datasets that include information on occupational tasks and skills** Does not apply. The exhibit does not show analysis results.
- **Figure B1** Does not apply. The exhibit does not show analysis results.
- **Figure 4: Correlation in occupational similarity scores across a standardized set of ISCO occupations for different datasets and methodologies** Reproduced.
- **Figure 6 - O*NET** Reproduced.
- **Figure 6 - ESCO** Reproduced.
- **Figure B2 - Panel a** Reproduced.
- **Figure B2 - Panel b** Reproduced.
- **Figure 7 - Panel 1** Reproduced.
- **Figure 7 - Panel 2** Reproduced.

Reproduction Environment

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

- OS: Windows 11 Enterprise, version 25H2
- Processor: Intel(R) Core(TM) Ultra 7 165U (2.10 GHz)
- Memory available: 31.5 GB
- Software version:
  - Python 3.8.17
  - Microsoft Excel for Microsoft 365 MSO