Reproducible Research Repository

Does Automatic Loan Approval Reduce Gender Bias in SME Lending?

First Submission: RR_PER_2026_603
Ruyuan Xiao, Mahin Tariq
reproducibility@worldbank.org
April 3, 2026

This review verifies the reproducibility of the exhibits included in the paper *"Does Automatic Loan Approval Reduce Gender Bias in SME Lending?"*.

Contents in this review:

Main findings

- **Every exhibit in the paper has been reproduced accurately.**
- As the package relies on restricted data, the replicators did not have direct access to it. The verification was therefore conducted through virtual verification on March 17, 2026, following the Virtual Reproducibility Verification Protocol. The process was executed in a clean environment to manage dependencies. The verification steps were as follows:
  (a) The authors executed the full workflow by running `master.do`.
  (b) The authors then shared the outputs generated from two consecutive runs.
  (c) The replicators checked whether the outputs from both runs matched and then verified each output against the exhibits in the paper.
- The output demonstrates consistent stability across multiple runs. Specifically, executing the code two times consecutively yielded identical results.
- The code takes approximately 30 minutes to run.
- We conducted our reproducibility analysis based on the paper shared by the authors in the reproducibility package.
- **Verification Process and Data Handling**:
  - The reproducibility package relies on 1 type of data:restricted data (available only from the authors).
  - 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 restricted data can use this file to verify that the data has not been altered.

- **Reproducibility Summary**:
  - **Data**: All 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 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.
- **Figure 1B** Reproduced.
- **Table 1** Reproduced.
- **Table 2** Reproduced.
- **Table 3** Reproduced.
- **Table 4** Reproduced.

Results in the Annex

- **Figure A1** Reproduced.
- **Table A1** Reproduced.
- **Table A2** Reproduced.
- **Table A3** Reproduced.
- **Table A4** Reproduced.

Reproduction Environment

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

- OS: Windows 11 Enterprise
- Processor: Intel(R) Core(TM) i5-1145G7 CPU @ 2.60GHz
- Memory available: 15.7 GB
- Software version: Stata 18.0 MP