Reproducible Research Repository

# *Helping women start up: Understanding the Constraints facing Women led Startups in India*

*Fourth Submission: RR_IND_2026_671*
*Zulekha Shoib, Vaishnavi Dhas*
*reproducibility@worldbank.org*
*August 17, 2026*

This review verifies the reproducibility of the exhibits included in the paper *"Helping women start up: Understanding the Constraints facing Women led Startups in India"*.

**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. Stata: Updating the working directory in line 38 of the do-file "MAIN.do", and running it.
  2. R: Opening the R project "Replication.Rproj". Restoring the environment by running renv::restore() and following the prompts. Opening Main.R and running the code.
- The output demonstrates consistent stability across multiple runs. Specifically, executing the code two times consecutively yielded identical results.
- The code takes approximately 25 minutes to run.
- We conducted our reproducibility analysis based on the paper shared by the authors by email on August 12, 2026.
- **Verification Process and Data Handling**:
  - The reproducibility package relies on 3 types of data: open, accessible and limited-access.
  - Reviewers used limited-acess data provided directly by the authors to conduct the reproducibility verification, and this is not included in the public reproducibility package.
  - Intermediate data is the starting point of the public reproducibility package and and can be used to fully reproduce the results in the paper.
  - Reviewers did not verify that the limited-access data matches what was provided by the authors as a license is required for access.

  - *data_hash_report.csv* lists the SHA256 hashes of all files in the Data folder to support data integrity checks. Users who acquire the limited-access data can use this file to verify that the data has not been altered.

- **Reproducibility Summary**:
  - **Data**: Some data is limited-access and has not been included in the reproducibility package. For more details, please refer to the README file. Intermediate data is included in the reproducibility package and can be used to fully reproduce the results in the paper
  - **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 1** Reproduced.
- **Figure 2** Reproduced.
- **Figure 3** Reproduced.
- **Figure 4** Reproduced.
- **Figure 5** Reproduced.
- **Figure 6** Reproduced.
- **Figure 7** Reproduced.
- **Table 1** Does not apply. This figure does not contain analytical output.
- **Table 2** Reproduced.
- **Table 3** Reproduced.
- **Table 2** Reproduced.
- **Table 3** Reproduced.
- **Table 4** Reproduced.
- **Table 5** Reproduced.
- **Table 6** Reproduced.
- **Table 7** Reproduced.
- **Table 8** Reproduced.

- **Table 9** Reproduced.
- **Table 10** Reproduced.
- **Table 11** Reproduced.
- **Table 12** Reproduced.
- **Table 13** Reproduced.
- **Table 14**Reproduced.
- **Table 15** Reproduced.
- **Table 16** Reproduced.
- **Table 17** Reproduced.

### *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: R 4.5.1, Stata 19.5 MP