



Selection or More Capital? Experimental Evidence on High-Growth Entrepreneurs in Kenya

First Submission: RR_KEN_2026_673

Ankriti Singh, Eric Sabiti
reproducibility@worldbank.org

July 8, 2026

This review verifies the reproducibility of the exhibits included in the paper “*Selection or More Capital? Experimental Evidence on High-Growth Entrepreneurs in Kenya*”.

Contents in this review:

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

Main findings

- **Every exhibit in the main section of the paper has been reproduced accurately.**
- The code was successfully executed on a new computer after:
 1. Updating the working directory in line 13 of the do-file "master file.do", and running it.
- The code takes approximately 3 hours, 19 minutes, and 20 seconds 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 1 type of data: limited-access (forthcoming as licensed files to World Bank Microdata Library)
 - Reviewers used limited-access data provided directly by the authors to conduct the reproducibility verification, and this is not included in the public reproducibility package. The data is forthcoming on the World Bank Microdata Library.
 - *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:** All data is limited-access 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:** 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**

- **Table 1** 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.
- **Figure 1** Does not apply. Exhibit is not an analytical output.
- **Figure 2** Reproduced.
- **Figure 3** Reproduced. Figure was produced manually in Excel.

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 928383, the Stata version used was 19.5, and the exhibits selected were: Table A1, Table A2, Table A4, Table A6, Table C1, Table C2, Table E1, Figure F1, Table F2, Table F3.

- **Table A1** Reproduced.
- **Table A2** Reproduced.
- **Table A4** Reproduced.
- **Table A6** Reproduced.
- **Table C1** Reproduced.
- **Table C2** Reproduced
- **Table E1** Reproduced.
- **Figure F1** Reproduced.
- **Table F2** Reproduced.
- **Table F3** Reproduced.

Reproduction Environment

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

- OS: Windows (CCSS)
- Processor: AMD EPYC 7763 64-Core Processor, 2445 Mhz, 8 Core(s), 16 Logical Processor(s)
- Memory available: 128 GB
- Software version: Stata 19.5 MP