



Algorithms and Bureaucrats: Evidence from Tax Audit Selection in Senegal

Third Submission: RR_SEN_2025_468

Maria Reyes Retana, Madalynn Smith

reproducibility@worldbank.org

January 30, 2026

This review verifies the reproducibility of the exhibits included in the paper “*Algorithms and Bureaucrats: Evidence from Tax Audit Selection in Senegal*”.

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 and 10 randomly selected exhibits in the appendix were reproduced accurately**
- The code was successfully executed on a new computer after changing the file paths and running the following codes:
 1. Master Tax.do.
 2. MASTER Prepare Analysis.do.
 3. MASTER Analysis.do.
 4. Construct_Data_RF_For_All_Logs.do.
 5. Senegal_Lit_Graph.do.
 6. R script oo_Main.R
- The output demonstrates consistent stability across multiple runs. Specifically, executing the code two times consecutively yielded identical results. Note: stability was verified virtually on a call with the authors on December 8.
- The code takes approximately 5 hours 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 3 types of data: Accessible, open, and restricted data.
 - 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 accessible files in the Data folder to support data integrity checks. Hashes for restricted data are not shared in this package, as the folder with the data contained many files with personal information as title, making it difficult to generate the hashes for the rest.
- **Reproducibility Summary:**
 - **Data:** Most data sources are restricted and not 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

- **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 Reproduced.** NOTE: A: Reproduced, B: rogue vertical lines on the brackets, C: Reproduced
- **Figure 2 Reproduced.** NOTE: Reviewed using "1 summary results CP and VG.pdf"
- **Figure 3 Reproduced.**
- **Figure 4 Reproduced.**

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 assuming that if all randomly selected exhibits are reproducible, then the rest should be as well. The seed used to generate the random selection was 4668, the Stata version used was 18.5, and the exhibits selected were: 3,5,9,11,16,26,29,33,37,43; equivalent to Figure B1; Table B9; Table C3; Table C5; Table C10; Figure D3; Figure E1; Figure E3; Table F4; Table F5; Figure G2.

- **Figure B1** Reproduced.
- **Table B9** Reproduced.
- **Table C3** Reproduced.
- **Table C5** Reproduced.
- **Table C10** Reproduced.
- **Figure D3** Reproduced.
- **Figure E1** Reproduced.
- **Figure E3** Reproduced.
- **Table F4** Reproduced.
- **Table F5** Reproduced.
- **Figure G2** 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