



The Transparency Tax: Discouraging Firm Ownership via Tax Havens

First Submission: RR_ECU_2026_719

Eric Sabiti

reproducibility@worldbank.org

September 24, 2026

This review verifies the reproducibility of the exhibits included in the paper *"The Transparency Tax: Discouraging Firm Ownership via Tax Havens"*.

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 partially reproduced on a new computer because the cleaning code required proprietary inputs that could not be shared. See below for details.
- The following steps were executed on the author's computer:
 1. Updating the working directory in "replication_masterdo.do", and running the code.
- The following steps were executed on the replicator's computer:
 1. Updating the working directory in line 56 of the do-file "3_replication_exhibits_construction.do", and running the script.
 2. Updating the working directory in line 34 of the R script "1a_cleaning_subsidiaries.R", and line 30 of the do-file "Orbis_replication/Code ob_master.do", and running the do-file.
- The output demonstrates consistent stability across multiple runs. Specifically, executing the code two times consecutively yielded identical results.
- The code takes approximately 2 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 2 types of data: open, and restricted data (available only from the data owners).
 - Open data is included in the public reproducibility package.

- Due to a proprietary input required for cleaning data, that portion of the code was not executed by the replicators. All steps are documented in the README and package. The authors shared the pre-processed outputs, and the replicators used these to execute and verify the analysis code and outputs.
- *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:** Some data is restricted and has not been included in the reproducibility package. For more details, 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

- **Figure 1** Reproduced.
- **Figure 2** Reproduced.
- **Figure 3** Reproduced.
- **Table 1** Reproduced.
- **Table 2** Reproduced.
- **Table 3** Reproduced.
- **Table 4** Reproduced.
- **Table 5** Reproduced.
- **Table 6** Reproduced.
- **Table 7** 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 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 212088, the Stata version used was 19.5, and the exhibits selected were: Figure A2, Figure A5, Figure A10, Figure 11, Table A4, Table A6, Table A10, Table A11, Table B1, Table B2.

- **Figure A2** Does not apply. The exhibit is not an analytical output.
- **Figure A5** Reproduced.
- **Figure A10** Reproduced.
- **Figure A11** Reproduced.
- **Table A4** Reproduced.
- **Table A6** Reproduced.
- **Table A10** Reproduced.
- **Table A11** Reproduced.
- **Table B1** Reproduced.
- **Table B2** 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: Stata 19.5 MP; R 4.5.3