Reproducible Research Repository

Financial Development Data

First Submission: FR_WLD_2026_770
Maria Reyes Retana, Nihaa Sajid
reproducibility@worldbank.org
September 29, 2026

This review verifies the reproducibility of the output datasets of the *"Global Financial Development Indicators and Indices Database"*. There is no accompanying manuscript; the outputs were identified from the README.

Contents in this review:

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

Main findings

Every output dataset in the package (core variables, supplementary variables and financial development indices) has been reproduced accurately.

- The code was successfully executed on a new computer after adjusting the root file path (`global path`) on line 10 of `01 Fin Data codemaster.do`. This master script runs all other scripts, from raw data to the output datasets and indices. All required Stata packages are included in the `code/ado/` folder.
- The output demonstrates consistent stability across multiple runs. Specifically, executing the code 2 times consecutively yielded identical results.
- The code takes approximately ∼2 minutes to run.
- We conducted our reproducibility analysis based on the output datasets described in the README.
- **Verification Process and Data Handling**:
  - The package relies on three types of data: publicly available data (IMF, World Bank, BIS), accessible data (WB Enterprise Surveys), and proprietary/subscription data (AXCO, Dealogic, Fitch, Bloomberg, PitchBook).
  - Open data is included in the public reproducibility package. Proprietary data was provided directly by the authors to conduct the verification and is not included in the public package (see README for sources and access).
  - *data_hash_report.csv* lists the SHA256 hashes of all files in the raw data folder, including the proprietary files used for verification, to support data integrity checks.

  - *comparison_report.csv* compares each of the 42 datasets produced by the reviewers' run with the authors' version. All datasets have the same number of observations, the same variables and the same data types. All common variable values match exactly or within 0.0006%, consistent with floating-point rounding.

- **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 output datasets) 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 output datasets and reproducibility status

Output datasets

- **Core variables (*core_variables.dta*)** Reproduced. The number of observations, variables, and data types match the authors' file. All variable values match within 0.0006%.
- **Supplementary variables (*sup_variables.dta*)** Reproduced. The number of observations, variables, and data types match the authors' file. All variable values match within 0.0006%.
- **Financial development indices (*idx_6_data_indices.xlsx*)** Reproduced. The number of observations, variables, and data types match the authors' file. All variable values match within 0.0006%.
- **Core variables (*core_variables.xlsx*)** Reproduced. The number of observations, variables, and data types match the authors' file. All variable values match within 0.0006%.
- **Supplementary variables (*sup_variables.xlsx*)** Reproduced. The number of observations, variables, and data types match the authors' file. All variable values match within 0.0006%.
- **Financial development indices (*idx_6_data_indices.xlsx*)** Reproduced. The number of observations, variables, and data types match the authors' file. All variable values match within 0.0006%.
- **Final FinData dataset (*FinData_Indices and Indicators.xlsx*)** Reproduced (manual changes). Final dataset with the 16 financial development indices, 34 core indicators, and 33 supplementary indicators for 218 economies, 1980–2024. The file combines three processed datasets: the 16 indices from `idx_6_data_indices.xlsx`, the 34 core indicators from `core_variables.xlsx`, and the 33 supplementary indicators from `sup_variables.xlsx`. Both indicator files are produced by `02 Fin Data merging.do`. The package has no code for this last step; the indicator datasets were merged with the indices, and the header and label rows were added by hand. Reviewers compared every cell of the final workbook, by economy and year, against the processed datasets:
  - 16 indices against `idx_6_data_indices.xlsx`: all values reproduced within 0.0001%.
  - 34 core indicators against `core_variables.xlsx`: all values reproduced within 0.00001%.

- 33 supplementary indicators against `sup_variables.xlsx`: all values reproduced within 0.00002%.

  These very small differences are floating-point precision effects in the GDP-based indicators. They do not affect reproducibility.

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 processors)
- Memory available: 128 GB
- Software version: Stata 18 MP