



Reproducible Research Repository

Worldwide Governance Indicators (WGI) 2026

First Submission: FR_WLD_2026_749

Maria Reyes Retana

reproducibility@worldbank.org

This review verifies the reproducibility of the output datasets produced by the reproducibility package “*Worldwide Governance Indicators (WGI) 2026*”.

Contents in this review:

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

Main findings

Every output dataset produced by this reproducibility package has been reproduced accurately.

- The code was successfully executed on a new computer after:
 1. Setting the folderpath global in 0. WGI master.do to the local machine path.
 2. Running the code.
- Reviewers verified the reproduced output datasets against the reference copies included in the authors’ resubmitted package. All three output datasets (wgidataset_with_sourcedata, wgiparameterestimates, all_inputs_long, in both .dta and .xlsx formats) matched on row count, variable coverage, and every common variable’s values and data types. File-level checksums differ from the authors’ shipped copies only because Stata embeds an internal save-timestamp in each .dta/.xlsx file header, which changes the file hash without changing any data value; this is documented in *comparison_report.csv*.
- The code takes approximately ~10 minutes to run end-to-end.
- We conducted our reproducibility analysis based on the reproducibility package as resubmitted by the authors on September 18, 2026.
- **Verification Process and Data Handling:**
 - The reproducibility package relies on two types of data: open data (30 of the 35 governance data sources) and limited-access/licensed commercial data (5 sources: EIU Operational Risk Service, WEF Executive Opinion Survey, Gallup World Poll, PRS Group ICRG, and S&P Global Country Risk Scores).
 - Open data is included in the public reproducibility package.

- The five licensed commercial sources are not included in raw form; only the standardized, rescaled values derived from them (`averaged_commercial_data.dta`) are included, per the data providers' redistribution terms documented in the README's Data Availability Statement.
- **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 raw input to final 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

- **wgidataset_with_sourcedata** (.dta / .xlsx) — Final WGI governance indicators (six dimensions) with source-level input data. **Reproduced.**
- **wgiparameterestimates** (.dta / .xlsx) — Unobserved Components Model parameter estimates (alpha, beta, sigma, weight) by source and dimension. **Reproduced.**
- **all_inputs_long** (.dta / .xlsx) — All raw indicator inputs to the WGI, long format, original scales. **Reproduced.**
- **non_commercial** (.dta) **Reproduced.**

Reproduction Environment

Output datasets were reproduced on a computer with the following specifications:

- OS: Windows 11 Enterprise
- Processor: Intel(R) Xeon(R) Platinum 8562Y+
- Memory available: 128 GB
- Software version: Stata 19.0 MP