

Reproducibility Package for Drivers and Trends for the Equality of Opportunity for Sexual and Gender Minorities: A Panel Approach

Paola Ballon & Omar Alburqueque
Global Department for Social Development, World Bank

July 9, 2026

Overview

This package reproduces the core results of the paper *Drivers and Trends for the Equality of Opportunity for Sexual and Gender Minorities: A Panel Approach*. It generates the EQOSOGI panel dataset, computes period-specific statistics, performs time-series cluster analysis, and assembles external covariates for econometric analysis. The workflow spans Stata and R and uses a directory layout with raw inputs and intermediate files.

1. Content

This reproducibility package contains the following folders:

- **inputs:** Contains the raw data files used by the package. These include the EQOSOGI panel input, the legal-system classification input, Penn World Table, the WDI extract, and the Quality of Government (QoG) Basic Dataset (Time-Series). The QoG file is not redistributed with the package; users must download it and place `qog_bas_ts_jan25.dta` in the `inputs/` folder before running the package.
- **scripts:** Contains the Stata and R scripts called by the master file.
- **intermediate files:** Contains files generated by the scripts and used by later steps of the workflow. Users should not edit these files manually.
- **outputs:** Contains the final EQOSOGI panel dataset and all reproduced tables and figures after running the package.

2. Software & Dependencies

2.1. Core Environment

- **Stata:** 16 MP or newer. The following packages should be installed by all users before running this package: `outreg2`, `ivreg2`, `ranktest` and `xtoverid`.
- **R:** Version 4.4.1. or newer. The following libraries should be installed by all users before running this package: `readxl`, `writexl`, `boot`, `dplyr`, `tidyr`, `stringr`, `tidyverse`, `cluster`, `e1071`, `clusterCrit`, `xts`, `lubridate`, `kofdata`, `ggplot2` and `purrr`.

2.2. Notes

- Internet access is required to successfully run the package.

3. Data Sources and Access

Most raw data files used in the analysis are included in this reproducibility package. The Quality of Government (QoG) Basic Dataset (Time-Series) is the only exception and must be downloaded by the user from the provider and placed in the `inputs/` folder. The dataset is freely accessible and does not require prior registration. Each dataset is also publicly available from its original provider; for transparency we report the corresponding URLs and the date on which each file was accessed. We redistribute the bundled raw data to ensure reproducibility, as some providers periodically update their datasets and later releases may not reproduce the exact results reported in the paper. Users who wish to work with updated versions can download them directly from the sources listed in Table 1 and replace the copies in the `inputs` folder. The KOF series are obtained programmatically using the script `1c_KOF_data.R`, which downloads the KOF Globalisation Database series and stores the resulting file in `intermediate files/`. Table 1 summarizes the raw inputs, their sources, and access dates, and indicates whether each file is included with the package or must be downloaded by the user.

4. How to Run

All empirical results are replicated through a single Stata master file, `0_master_file.do`, which coordinates both the Stata and R components of the package. Before running it, the user must open `0_master_file.do` and set two local paths:

1. `global eqosogi_userpath` must point to the root folder where the reproducibility package has been copied on the local machine.
2. `global Rscript_path` must point to the `Rscript.exe` file in the local R installation (for example, `C:\Program Files\R\R-4.4.1\bin\x64\Rscript.exe`).

Once these two globals are correctly set, the user can save and run `0_master_file.do` in Stata. The master file sequentially calls all numbered scripts in the recommended execution order and manages the interaction between Stata and R. All tables and figures reproduced by the package are written to the `outputs` subfolder.

5. Reproduced Outputs

This section summarizes the outputs reproduced by the master file.

- **Data:** Final panel dataset, `EQOSOGI_panel_data.dta`.
- **Tables:** Table 3, Table 4, Table 5, Table 6, Table 7, Table A2, Table A2.1, Table A2.2, Table A2.3, Table A3, Table A4, and Table A5.
- **Figures:** Figure 1 and Figure 2.

Table 1: Raw data files, sources, and access dates

Filename	Data	Availability	Source	Access date
EQOSOGI_Panel_Data.xlsx	EQOSOGI panel legal dataset (authors' compilation)	Included with the package	Provided by authors	November 2025
Legal_Systems.xlsx	Legal-system classification input used to construct the colonial condition–legal system variable	Included with the package	Provided by authors; based on a review of QoG and JuriGlobe sources	May 2026
pwt1001.dta	Penn World Table v10.01	Included with the package	Penn World Table website: Link to source	June 2025
qog_bas_ts_jan25.dta	Quality of Government Institute Basic Dataset (Time-Series)	User must download and place in the inputs/ folder	Quality of Governance website: Link to source	December 2025
WDI_data.xlsx	WDI (only GDP per capita, constant 2015 US\$ series)	Included with the package	World Development Indicators: Link to source	June 2025

The corresponding files are written to the `outputs/` folder. Table A2.1, Table A2.2, and Table A2.3 are exported as separate sheets in the workbook `Tables_A2-1_to_A2-3.xlsx`.

6. Access, Rights & Disclaimer

This work is distributed under the Creative Commons Attribution 4.0 International license (CC BY 4.0). Data remain subject to the licenses of their respective owners (PWT, QoG, WDI, KOF). Any findings, interpretations, and conclusions expressed are those of the authors and do not necessarily reflect the views of the World Bank, its Executive Directors, or the governments they represent.

7. Citation

Ballon, P. & Alburquerque, O. (2025). *Reproducibility Package for Drivers and Trends for the Equality of Opportunity for Sexual and Gender Minorities: A Panel Approach*. World Bank.

8. Contact

For inquiries, please contact Paola Ballon – pballon@worldbank.org and Omar Alburquerque – oalburquequechav@worldbank.org