Readme File for Replication Package for 2022 Update of Worldwide Governance Indicators

September 23, 2022

Download contents of .zip archive into desired folder on your computer. The folder will contain:

1. **readme.docx** – i.e. this file.
2. **wgicode1.do** – this is the master .do file that needs to be executed to generate aggregate indictors.
3. **importmerge1.do, ucmcode1.do, nonrepcode1.do, rankcalc1.do** – these are all subsidiary .do files called by wgicode.do. Descriptions of these files are included at the top of each file
4. **countrycodes.xlsx** – this file contains the country codes and country names as used in the WGI
5. **sourcedata** – this folder contains 34 Excel files **XXX.xlsx** where XXX is a three-letter code corresponding to each of the 34 WGI data sources. The Excel files contain the data from each source, and also are posted separately on the WGI website. This folder also contains a copy of **countrycodes.xlsx.**
6. **sourcedofiles** – this folder contains 34 Stata .do files **XXX_merge.do.** These files are called by **importmerge1.do** and are used to retrieve the data from each Excel file into Stata format.
7. **reproducibilitycheck.pdf** – this file contains the summary findings of a code review of the replication package conducted by DIMEAnalytics at the World Bank.

For a complete technical description of the Unobserved Components Model used to construct the aggregate WGI indicators, please refer to Section 3 of Kaufmann, Daniel, Aart Kraay and Pablo Zoido-Lobaton (1999). "Aggregating Governance Indicators". World Bank Policy Research Working Paper No. 2195.

To replicate the WGI aggregation calculations:

1. At the beginning of wgicode1.do, set the global $filepath to match the folder to which you have downloaded the replication package on your computer.
2. Open Stata and run wgicode1.do.

This will generate the following output files that will be stored in the same folder to which you downloaded the replication package:

1. **wgidataset.dta** – this is a Stata dataset containing the six aggregate indicators, standard errors, number of sources, and percentile ranks identical to the one posted on the home page of www.govindicators.org. This dataset is in "long" format with countries and time periods stacked vertically.
2. **aggindsdownloadable.xlsx** – this is an Excel file containing an unformatted version of the Excel file posted on the home page of www.govindicators.org, containing all of the aggregate indicators. This dataset is in "wide" format with one tab for each indicator and separate columns for each variable and time period, and separate tabs for each of the six indicators.
3. **sourcedata.dta** -- this is a Stata dataset containing all the individual indicators used to calculate the aggregate indicators. This dataset is in "wide" format with countries in rows and separate columns for each variable and period.

4. **resultsdata.dta** – this is a Stata dataset containing all the individual indicators, all the aggregate indicators, and all the estimated parameters of the unobserved components model used to generate the aggregates. This dataset is in "wide" format with countries in rows and separate columns for each variable and period.
5. **params.xlsx** – this is an Excel file containing the estimated parameters of the unobserved components model for each data source. This dataset is in "long" format with individual indicators and periods stacked vertically and three columns for the three parameters for each source.
6. **individualindstableau.xlsx** and **aggindstableau.xlsx** are two Excel files containing all of the individual indicators and all of the aggregate indicators. These contain no new information relative to the previous files, but are used as inputs to the Tableau dashboards on the interactive portion of www.govindicators.org.

For questions please contact Aart Kraay (akraay@worldbank.org).