

README for the Reproducibility Package for Public Institutions Global Observatory: Annual Review 2025

Overview

This reproducibility package is designed to recreate all analyses featured in the Public Institutions Global Observatory (PIGO) Annual Review 2025, a publication produced by the EGVPI. It contains all the necessary data transformation scripts, custom functions, and plotting scripts required to replicate the report's figures precisely.

The package is structured to align with the report's chapters:

- Chapter 1: Workflows in this section enable the reproduction of analyses examining global and regional patterns and trends in institutional capacity.
- Chapter 2: This chapter's contents allow for a more granular, cluster-level analysis of operational engagements and the evidence based on institutional capacity. Additionally, it includes scripts to reproduce the display of selected highlights from World Bank lending operations and knowledge products focused on institutions.

It is important to note that the PIGO Annual Review 2025 is intended for internal use by World Bank staff. Therefore, the report's structure and findings are specifically tailored to meet the needs of staff, ensuring it is fit-for-purpose and providing actionable insights.

Instructions for Replicators

Replicating the results involves the following steps:

1. Restore Package Dependencies:
 - a. Open the `pigographs.Rproj` file in RStudio.
 - b. Execute the command `renv::restore()` in the RStudio console. This will install the exact package versions specified in the `renv.lock` file.
2. Execute the Full Workflow: To perform both stages conveniently, open the Main R script and run it. This will execute the following tasks in the required order.
 - a. Data Transformation:
 - Confirm that you have received and placed the `input/` folder provided by the EGVPI team in the project's root directory.
 - Source each of the raw data preprocessing scripts located in the `data-raw/source/` subdirectory. Execute these scripts sequentially.
 - b. Data Analysis & Visualization:
 - Source all analysis and plotting scripts found within the `analysis/` subdirectory.
 - Carefully verify that each plot generated in the `figures/` folder matches the corresponding figure in the PIGO Annual Review exactly.

All generated figures and outputs will be saved in the `figures` folder. For detailed information on each exhibit, please refer to the List of Exhibits.

Data Availability Statement

The PIGO report provides a concise, high-level synthesis of evidence on public institutions around the world, designed to inform World Bank staff about the latest trends in institutional capacity, reforms, and analytics. The report's core findings rest on the Country-Level Institutional Assessment and Review (CLIAR) database, augmented by complementary World Bank indicators. As these datasets are restricted to internal use, the following section provides clear guidance on how staff can request and secure access to each file in order to reproduce the analysis.

Chapter 1: Global and Regional Trends in Institutional Capacity

Filename : cliar\db_variables.xlsx

- Availability: World Bank Internal, based on public data.
- Source: CLIAR Dashboard
- URL: <https://datanalytics-int.worldbank.org/CLIAR/>
- Access year: May 2025.
- Note: This is a codebook that enables us to understand the clusters composition of different indicators and sources curated by the EGVPI team. To access it, go to the URL, select the data tab, go to the data download section and select the Data Dictionary button.

Filename : cliar\dynamic_ctf_032525.dta

- Availability: World Bank Internal, based on public data.
- Source: CLIAR Dashboard
- URL: <https://datanalytics-int.worldbank.org/CLIAR/>
- Access year: March 2025.
- Note: To retrieve the file, open the URL, select the Data tab, navigate to Data Download, choose CTF Dynamic (All Indicators), and in Interactive Data Access & Custom Download pick the .dta format.

Filename : cliar\static_ctf_032525.dta

- Availability: World Bank Internal, based on public data.
- Source: CLIAR Dashboard
- URL: <https://datanalytics-int.worldbank.org/CLIAR/>
- Access year: March 2025.
- Note: To retrieve the file, open the URL, select the Data tab, navigate to Data Download, choose CTF Static (All Indicators), and in Interactive Data Access & Custom Download pick the .dta format.

Filename : correlation\static_ctf_012225.dta

- Availability: World Bank Internal, based on public data.
- Source: CLIAR Dashboard
- URL: <https://datanalytics-int.worldbank.org/CLIAR/>
- Access year: January 2025.
- Note: To retrieve the file, open the URL, select the Data tab, navigate to Data Download, choose CTF Static (All Indicators), and in Interactive Data Access & Custom Download pick the .dta format.

Filename : wb\client_countries.csv and correlation\client_countries.csv

- Availability: Public.
- Source: File elaborated by the team with public information from:
<https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups>
- URL: NA
- Access year: February 2025.
- Note: Data has been transformed with subject matter expertise.

Filename : wb\CLASS.xlsx and correlation\CLASS.xlsx

- Availability: World Bank Public.
- Source: World Bank Country and Lending Groups
- URL: <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups>
- Access year: February 2025.
- Note: Raw data. Data has not been modified.

Filename : correlation\country_risk_index.xlsx

- Availability: Restricted
- Source: Fitch Solutions and BMI (Business Monitor International)
- URL: internal access, contact the World bank library to get an account created before following the steps detailed in the note
- Access year: March 2025.
- Note: Log in to the BMI platform using your credentials. Navigate to the Country Risk Index page via the left-hand navigation menu. Use the filters to select: – Country Risk Index pillar – Year – Region. Hover over geography to view the Country Risk Index score. Access the Country Risk methodology document by clicking the "Methodology" button. Download the data for the entitled geographies as an Excel file by clicking the 'Download' button. View the Country Risk Index scores for the entitled geographies in the table below. For a detailed profile of a geography's Country Risk Index, click on the desired geography in the table. Contact: Hannah Kang: hannah.kang@fitchsolutions.com

Filename : correlation\P_Data_Extract_From_World_Development_Indicators_updated.xlsx

- Availability: World Bank Public.
- Source: WDI
- URL: <https://databank.worldbank.org/source/world-development-indicators>
- Access year: 2025.
- Note: Raw data. Data has not been modified.

Chapter 2: Operational Engagements and Evidence on Institutional Capacity

Filename : boost\BOOST_country_data.xlsx

- Availability: Restricted. Internal use only.
- Source: File shared by Massimo Mastruzzi on October 2024
- URL: <https://www.worldbank.org/en/programs/boost-portal/boost-data-lab>
- Access year: October 2024
- Note: To access the data, email mmastruzzi@worldbank.org and request the Cross-Country Analysis full dataset in .xlsx format.

Filename : boost\budget_priorities.xlsx

- Availability: Internal.
- Source: Internal
- URL: NA
- Access year: November 2024
- Note: Based on BOOST_country_data.xlsx, Lead economist Rita Ramalho used subject expertise to define this list

Filename : wb/gspcs.csv

Original file name: GSPS_Indicators_Dataset_11_10_23.xlsx.

- Availability: Public.
- Source: Global Survey of Public Servants
- URL: https://www.globalsurveyofpublicservants.org/sites/default/files/2023-11/GSPS_Indicators_Dataset_11_10_23.xlsx
- Access year: May 2025.
- Note: Raw data. Data has not been modified.

Filename : wb/dgss.csv.

Original file name: WBG_GovTech Dataset_Dec2020.xlsx

- Availability: Public.
- Source: GovTech Maturity Index.
- URL: https://datacatalogfiles.worldbank.org/ddh-published/0037889/DR0065450/WBG_GovTech%20Dataset_Dec2020.xlsx
- Access year: May 2025.
- Note: Raw data. Data has not been modified.

Overview table

Source	Dataset	Link	Year	Access
CLIAR Dashboard	cliar\db_variables.xlsx	https://w0lxdrconn01.worldbank.org/CLIAR/ Codebook elaborated by the EGVPI team	May 2025	Limited
CLIAR Dashboard	cliar\dynamic_ctf_032525.dta	https://w0lxdrconn01.worldbank.org/CLIAR/	March 2025	Limited
CLIAR Dashboard	cliar\static_ctf_032525.dta	https://w0lxdrconn01.worldbank.org/CLIAR/	March 2025	Limited
World Bank	wb\client_countries.csv	Internally shared	February 2025	Public
World Bank	wb\CLASS.xlsx	https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups	February 2025	Public
World Bank	P_Data_Extract_From_World_Development_Indicators_updated.xlsx	https://databank.worldbank.org/source/world-development-indicators	March 2025	Public
World Bank	boost\BOOST_country_data.xlsx	Internally shared	October 2024	Limited
World Bank	boost\budget_priorities.xlsx	Internally processed	November 2024	Limited
Global Survey of Public Servants	gsps/GSPS_Indicators_Data_set_11_10_23.xlsx	https://www.globalsurveyofpublicservants.org/sites/default/files/2023-11/GSPS_Indicators_Dataset_11_10_23.xlsx	May 2025	Public
GovTech Maturity Index	wb/dgss.csv	https://datacatalogfiles.worldbank.org/ddh-published/0037889/DR0065450/WBG_GovTech20DataSet_Dec2020.xlsx	May 2025	Public

Statement about Rights

I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

I certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package. Appropriate permissions are documented in the LICENSE.txt file.

Software Requirements

R: Version 4.3 or higher

Dependency Management: renv for managing package dependencies

Development Tools: devtools for package loading

Memory, Runtime, and Storage Requirements

Memory: No minimum required, but 16 GB is recommended for optimal performance.

Run Time: ~5 minutes on recommended hardware.

Folder Structure

The repository is organized as follows:

— data-raw/	# Raw data and initial processing scripts.
— source	# Scripts for transforming raw data
— input	# Raw data files (provided by the EGVPI Team)
— output	# Transformed data for ad-hoc analysis
— analysis/	# Scripts for analysis and plotting.
— R/	# Custom R functions.
— documentation/	# Reproducibility documentation.
— data/	# Processed data for analysis.
— figures/	# Final generated plots.
— man/	# R package manual pages.
— renv/	# R package dependency management.
— README. Rmd	# Overview to run workflow.
— pigographs.Rproj	# RStudio project file.
— .gitignore	# Files ignored by Git.
— .Rbuildignore	# Files ignored when building.
— .Rprofile	# Project-specific R startup settings.
— DESCRIPTION	# Project metadata.
— LICENSE	# Licensing information.
— LICENSE.md	# License in Markdown format.
— NAMESPACE	# Package namespace definition.
— main.R	# Main file to reproduce the whole workflow.
— spielplatz/	# Sandbox/experimentation area.

List of Exhibits

Exhibit name	Output filename	Script	Note
Figure 1. Country-Level Institutional Assessment and Review (CLIAR) Framework	NA	NA	Image. Source: CLIAR Methodology note
Figure 2. Quality of institutions and levels of economic development are positively correlated	GDPpercapitaPPPconstant2021international_main	correlation-income-risk.R	
Figure 3. Higher levels of institutional capacity are associated with lower levels of poverty	Povertyheadcountratioat215aday2017PPPofpopulation_main	correlation-income-risk.R	
Figure 4. Higher levels of institutional capacity are correlated with lower country risk	CountryRisk_vs_CTF_main	correlation-income-risk.R	
Figure 5. Regional variation in the policy arena (political and social)	05-poli-arena-regional-dumbbells.png	regional-cluster-dumbbells.R	
Figure 5. Regional variation in the policy arena (political and social)	05-social-arena-regional-dumbbells.png	regional-cluster-dumbbells.R	
Figure 6. Regional variation in institutions at the center of government	06-center-of-gov-digital-regional-dumbbells.png	regional-cluster-dumbbells.R	
Figure 6. Regional variation in institutions at the center of government	06-center-of-gov-hrm-regional-dumbbells.png	regional-cluster-dumbbells.R	
Figure 7. Regional Variation in Degree of Integrity and Transparency and Accountability	07-integrity-regional-dumbbells.png	regional-cluster-dumbbells.R	
Figure 7. Regional Variation in Degree of Integrity and Transparency and Accountability	07-transp-regional-dumbbells.png	regional-cluster-dumbbells.R	
Figure 8. Regional variation in specific thematic areas: Justice and Energy and Environment	08-enviroment-regional-dumbbells.png	regional-cluster-dumbbells.R	
Figure 8. Regional variation in specific thematic areas: Justice and Energy and Environment	08-justice-regional-dumbbells.png	regional-cluster-dumbbells.R	
Figure 9. Change in political institutions over time (2018 vs 2022)	09-political-cluster-change.png	clusters-over-time.R	
Figure 10. Change in social institutions over time (2018 vs 2022)	10-social-cluster-change.png	clusters-over-time.R	
Figure 11. Change in public human resource management institutions over time (2018-2022)	11-hrm-cluster-change.png	clusters-over-time.R	
Figure 12. Change in degree of integrity over time (2018-2022)	12-integrity-cluster-change.png	clusters-over-time.R	

Figure 13. Regional variation in the share of workers with tertiary education in the public and private sectors	NA	NA	Image. Source: Worldwide Bureaucracy Indicators ¹
Figure 14. Global datasets on public financial management	NA	NA	Hyperlinks references, not applicable
Figure 15. Regional variation in capital budget execution rate in service delivery (2019)	15-service-delivery-education-budget-execution.png	budget-execution-by-region.R	
Figure 15. Regional variation in capital budget execution rate in service delivery (2019)	15-service-delivery-health-budget-execution.png	budget-execution-by-region.R	
Figure 16. Global datasets on government technologies	NA	NA	Hyperlinks references, not applicable
Figure 17. Share of countries with operational management information systems, by region	17-coverage_information_systems.png	analyze_information_systems.R	
Figure 18. Global datasets on management practices	NA	NA	Hyperlinks references, not applicable
Figure 19. Institutional-level variation in institutional target-setting (Guatemala and Ghana)	19_mgmt_target.png	gsps-management.R	
Annex 1. Additional Correlations between Institutional Dimensions and Prosperity Indicators (a)	GDPpercapitaPPPconstant2021international_annex	correlation-income-risk.R	
Annex 1. Additional Correlations between Institutional Dimensions and Prosperity Indicators (b)	Povertyheadcounratioat215aday2017PPPofpopulation_annex	correlation-income-risk.R	
Annex 1. Additional Correlations between Institutional Dimensions and Prosperity Indicators (c)	CountryRisk_vs_CTF_annex	correlation-income-risk.R	
Annex 1. Additional Correlations between Institutional Dimensions and Prosperity Indicators (d)	Povertyheadcounratioat215aday2017PPPofpopulation_annex_nossa	correlation-income-risk.R	
Annex 2. Change in Sector-specific Institutions (a) - justice	annex-a-justice-cluster-change.png	clusters-over-time.R	
Annex 2. Change in Sector-specific Institutions (b) - environmental	annex-b-enviromental-cluster-change.png	clusters-over-time.R	
Annex 5. Regression of Country Risk Index on organizational and governance dimensions	lm_risk_institutions.doc	correlation-income-risk.R	

¹ There are three steps: (1) select relevant regions (e.g., EAP, LAC), (2) export the visualization as a PowerPoint (bottom right corner of the dashboard), selecting a specific sheet from the dashboard, in our case it is figure 3, Tertiary education, (3) copy and paste the image from the PowerPoint