

Public Institutions Flagship Report — Reproducibility Package

This repository contains all the analysis produced for the Public Institutions (PI) Flagship Report: *Enabling Private Sector Development with a Focus on Procurement and Regulatory Institutions*. It is structured to adhere to the [World Bank Reproducible Research Repository guidelines](#).

Contents

1. [Overview](#)
 2. [Data Availability](#)
 3. [Instructions for Replicators](#)
 4. [List of Exhibits](#)
 5. [Requirements](#)
 6. [Code Description](#)
 7. [Folder Structure](#)
-

Overview

This reproducibility package contains all code and documentation necessary to replicate the analysis in the Public Institutions Flagship Report. The package includes data-cleaning scripts, analysis scripts, and the code required to produce all tables and figures presented in the report. The analysis is organized by chapter, each with its own data/ and source/ folders.

Replicators should begin by reviewing the [Requirements](#) section to ensure all necessary software and dependencies are installed, then follow the step-by-step instructions in [Instructions for Replicators](#).

Folder Structure

```
pi-flagship-report/  
├── chapter-1/  
│   ├── data/  
│   │   ├── input/  
│   │   └── output/  
│   ├── figs/  
│   └── source/  
├── chapter-2/  
│   ├── data/  
│   │   ├── input/  
│   │   └── output/  
└──
```

```
├── figs/
├── source/
├── chapter-3/
│   ├── data/
│   │   ├── input/
│   │   └── output/
│   ├── figs/
│   ├── source/
│   └── stata-excel
├── manuscript/
├── LICENSE
└── README.md
```

Data Availability

- ☐ All data are publicly available.
- ☒ Some data cannot be made publicly available.
- ☐ No data can be made publicly available.

Data Sources

Chapter 1

- **Filename:** gsps.csv
 - **Source:** World Bank Global Survey of Public Servants (GSPS)
 - **URL:** https://www.globalsurveyofpublicservants.org/sites/default/files/2023-11/GSPS_Indicators_Dataset_11_10_23.xlsx
 - **Access year:** 2024
 - **Description:** Survey data on public service providers across multiple countries, covering topics including recruitment, management, independence, and transparency.
 - **License:** Creative Commons Attribution 4.0 International (CC BY 4.0)
-

- **Filename:** WBG_GovTech_Dataset_Oct2022.xlsx
 - **Source:** World Bank GovTech Maturity Index (GTMI)
 - **URL:** <https://www.worldbank.org/en/programs/govtech/gtmi>
 - **Access year:** 2022
 - **Description:** Cross-country data on the maturity of government technology systems including FMIS, TSA, HRMIS, e-procurement, PIMS, and customs systems.
 - **License:** Creative Commons Attribution 4.0 International (CC BY 4.0)
-

- **Filename:** regional_classification.dta
 - **Source:** World Bank country and lending groups classification
 - **URL:** <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519>
 - **Access year:** 2024
 - **Description:** World Bank country classification by region, income group, and lending category.
 - **License:** Creative Commons Attribution 4.0 International (CC BY 4.0)
-

- **Filename:** boost/boost_arm_2006_2022.xlsx
 - **Source:** World Bank BOOST Public Expenditure Database — Armenia
 - **URL:** <https://www.worldbank.org/en/programs/boost-portal>
 - **Access year:** 2024
 - **Description:** Detailed public expenditure data for Armenia (2006–2022), disaggregated by ministry and economic classification.
 - **License:** World Bank internal data; not publicly redistributable
-

- **Filename:** boost/boost_pgy_2006_2022.xlsx
 - **Source:** World Bank BOOST Public Expenditure Database — Paraguay
 - **URL:** <https://www.worldbank.org/en/programs/boost-portal>
 - **Access year:** 2024
 - **Description:** Detailed public expenditure data for Paraguay (2006–2022), disaggregated by ministry and economic classification.
 - **License:** World Bank internal data; not publicly redistributable
-

- **Filename:** boost/boost_uga_2004_2022.xlsx
 - **Source:** World Bank BOOST Public Expenditure Database — Uganda
 - **URL:** <https://www.worldbank.org/en/programs/boost-portal>
 - **Access year:** 2024
 - **Description:** Detailed public expenditure data for Uganda (2004–2022), disaggregated by ministry and economic classification.
 - **License:** World Bank internal data; not publicly redistributable
-

Chapter 2

- **Filename:** maps.xlsx
- **Source:** Methodology for Assessing Procurement Systems (MAPS)

- **URL:** <https://www.mapsinitiative.org/>
 - **Access year:** 2024
 - **Description:** Country-level assessments of public procurement systems across pillars and indicators, including personnel, independence, transparency, accountability, and information systems dimensions.
 - **License:** Available upon request from the MAPS Secretariat
-

- **Filename:** `cwg.xlsx`
 - **Source:** World Bank Business Ready (B-READY) — Contracting with Government
 - **URL:** <https://www.worldbank.org/en/businessready>
 - **Access year:** 2024
 - **Description:** Country scores on the ease of contracting with the government, including procedure, time, legal, and e-procurement sub-indices (191 economies).
 - **License:** Creative Commons Attribution 4.0 International (CC BY 4.0)
-

- **Filename:** `gtmi.xlsx`
 - **Source:** World Bank GovTech Maturity Index (GTMI)
 - **URL:** <https://www.worldbank.org/en/programs/govtech/gtmi>
 - **Access year:** 2022
 - **Description:** Same as WBG_GovTech_Dataset_Oct2022.xlsx (Chapter 1), used here for panel and cross-sectional analysis of e-procurement systems.
 - **License:** Creative Commons Attribution 4.0 International (CC BY 4.0)
-

- **Filename:** `regions.dta`
 - **Source:** World Bank country and lending groups classification
 - **URL:** <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519>
 - **Access year:** 2024
 - **Description:** Same as `regional_classification.dta` (Chapter 1).
 - **License:** Creative Commons Attribution 4.0 International (CC BY 4.0)
-

- **Filename:** `wdi.csv`
- **Source:** World Bank World Development Indicators (WDI)
- **URL:** <https://databank.worldbank.org/source/world-development-indicators>
- **Access year:** 2024
- **Description:** Macroeconomic covariates including GDP per capita (2015 USD), population, and other country-level indicators used to construct control variables.

- **License:** Creative Commons Attribution 4.0 International (CC BY 4.0)
-

- **Filename:** bfa.xlsx
 - **Source:** World Bank BOOST Public Expenditure Database — Burkina Faso
 - **URL:** <https://www.worldbank.org/en/programs/boost-portal>
 - **Access year:** 2024
 - **Description:** Detailed public expenditure data for Burkina Faso (2017–2020), disaggregated by ministry and economic classification, used to analyze procurement office spending relative to investment spending.
 - **License:** World Bank internal data; not publicly redistributable
-

Chapter 3

- **Filename:** itu_appointment.csv, itu_workforce.csv, itu_broadband.csv, itu_coverage.csv, itu_decision_making.csv, itu_dispute_resolution.csv, itu_institutional_structure.csv, itu_investment.csv, itu_mobile_dropped_call_ratio.csv, itu_mobile_subscription.csv, itu_revenue.csv
 - **Source:** International Telecommunication Union (ITU) — Regulatory Tracker
 - **URL:** <https://regtracker.itu.int/>
 - **Access year:** 2024
 - **Description:** Country-level data on telecom regulatory institutions, covering appointment authority, workforce composition, dispute resolution mechanisms, institutional structure, investment levels, broadband traffic, mobile subscriptions, coverage, and revenue.
 - **License:** ITU terms of use; available for research purposes
-

- **Filename:** oecd_gsr.xlsx
 - **Source:** OECD — Indicators of Regulatory Policy and Governance (iREG) / Government at a Glance
 - **URL:** <https://www.oecd.org/en/topics/regulatory-governance.html>
 - **Access year:** 2024
 - **Description:** Processed country-level scores on regulatory governance dimensions (independence, accountability, scope) across energy, communications, air, rail, and water sectors.
 - **License:** OECD terms of use
-

- **Filename:** oecd_gsr_original.xlsx
 - **Source:** OECD — Indicators of Regulatory Policy and Governance (iREG), raw database
 - **URL:** <https://www.oecd.org/en/topics/regulatory-governance.html>
 - **Access year:** 2024
 - **Description:** Full question-level regulatory governance data (2023 edition) used to construct sector-level scores for strategic objectives, financial independence, and accountability.
 - **License:** OECD terms of use
-

- **Filename:** bready_business_entry.csv, bready_business_insolvency.csv, bready_business_location.csv, bready_dispute_resolution.csv, bready_financial_services.csv, bready_international_trade.csv, bready_labor.csv, bready_market_competition.csv, bready_taxation.csv, bready_utility_services.csv
 - **Source:** World Bank Business Ready (B-READY) 2024
 - **URL:** <https://www.worldbank.org/en/businessready>
 - **Access year:** 2024
 - **Description:** Pillar I (de jure) and Pillar II (de facto) scores by topic for 184 economies, used to construct regulatory implementation gap measures.
 - **License:** Creative Commons Attribution 4.0 International (CC BY 4.0)
-

- **Filename:** banking_regulator.csv
 - **Source:** World Bank Bank Regulation and Supervision Survey (BRSS) — 2019 edition
 - **URL:** <https://datacatalog.worldbank.org/int/search/dataset/0038632>
 - **Access year:** 2024
 - **Description:** Country-level data on bank regulatory authority characteristics including workforce size, share of specialized staff, and staff tenure.
 - **License:** Creative Commons Attribution 4.0 International (CC BY 4.0)
-

- **Filename:** gdp_pc_ppp.csv
- **Source:** World Bank World Development Indicators — GDP per capita, PPP (constant 2017 international \$)
- **URL:** <https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.KD>
- **Access year:** 2024
- **Description:** GDP per capita in PPP terms (1960–2022) used as a covariate to order countries in regulatory governance charts.
- **License:** Creative Commons Attribution 4.0 International (CC BY 4.0)

-
- **Filename:** regulatory_governance.csv (in world-bank/girg/)
 - **Source:** World Bank Global Indicators of Regulatory Governance (GIRG)
 - **URL:** <https://rulemaking.worldbank.org/>
 - **Access year:** 2024
 - **Description:** Country-level indicators on regulatory governance practices including consultation, transparency, and impact assessment (2018 edition).
 - **License:** Creative Commons Attribution 4.0 International (CC BY 4.0)
-

- **Filename:** afdb_regulatory_gov.csv
 - **Source:** African Development Bank (AfDB) — Regulatory Governance Dataset
 - **URL:** https://africa-energy-portal.org/sites/default/files/2023-02/ERI%202022_AFDB%20EN.pdf
 - **Access year:** 2024
 - **Description:** Country-level transparency scores for regulatory institutions in African countries, used to produce the Africa regulatory transparency map.
 - **License:** Available upon request from AfDB
-

- **Filename:** WB_countries_Admin0_10m (shapefile set)
 - **Source:** World Bank — Official Boundaries Shapefiles
 - **URL:** <https://datacatalog.worldbank.org/int/search/dataset/0038272>
 - **Access year:** 2024
 - **Description:** World Bank official country boundaries at 1:10m scale, used for mapping regulatory governance scores across African countries.
 - **License:** Creative Commons Attribution 4.0 International (CC BY 4.0)
-

- **Source (downloaded at runtime):** World Bank country classification
 - **URL:** <https://ddh-openapi.worldbank.org/resources/DR0095333/download/>
 - **Access year:** 2024
 - **Description:** World Bank country classification by income group and region, downloaded programmatically in chapter_3_prepdata.R.
 - **License:** Creative Commons Attribution 4.0 International (CC BY 4.0)
-

- **Filename:** Updated_PMR_Sector_Indicator_2023-24_and_2018.xlsx (sheet PMR_Sector_Network_2023-24_clea)
 - **Source:** OECD Product Market Regulation (PMR) Sector Indicators, 2023–24 wave
-

- **URL:** https://www.oecd.org/content/dam/oecd/en/topics/policy-sub-issues/product-market-regulation/PMR_Sector_Indicator_2023-24_and_2018_March2026.xlsx
 - **Access year:** 2025 (August)
 - **Description:** PMR sector sub-indicators used to measure regulatory stringency in e-communications. Variable on the x-axis: e_communications_mobile_ecomm. Column names were manually adjusted to comply with Stata naming rules before import. Albania (ALB) is dropped as an outlier.
 - **License:** OECD terms of use
 - **Note:** The OECD has refreshed this file since the access date; replication against the current version may yield slightly different values.
-

- **Filename:** ITU_ICTPriceBaskets_2008-2024.xlsx (sheet economies_2008-2024)
 - **Source:** ITU ICT Price Baskets, 2008–2024
 - **URL:** <https://www.itu.int/en/ITU-D/Statistics/Pages/ICTprices/default.aspx>
 - **Access year:** 2025 (June 12)
 - **Description:** Mobile-broadband basket prices expressed as a percent of GNI per capita (i271mb_ts_gni), restricted to the year 2023. Used on the y-axis of Figure 3.3.
 - **License:** ITU terms of use
 - **Note:** ITU has refreshed the data since the access date; replication against the current version may yield different results.
-

- **Filename:** CLASS_2025_07_02.xlsx
 - **Source:** World Bank country income classification, July 2025 release
 - **URL:** <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519>
 - **Access year:** 2025
 - **Description:** Used to assign income group labels (HIC, UMC, LMC) to countries in the PMR dataset via a VLOOKUP step applied to the Excel input before Stata import. Must be applied before running 1_Figures.do.
 - **License:** Creative Commons Attribution 4.0 International (CC BY 4.0)
-

- **Filename:** GSR_Data.xlsx (sheet Database_clean)
- **Source:** OECD 2023 Indicators on the Governance of Sector Regulators (GSR)
- **URL:** [https://www.oecd.org/content/dam/oecd/en/publications/support-materials/2025/03/the-2023-indicators-on-the-governance-of-sector-regulators_eb33f9bb/2023%20Indicators%20on%20the%20Governance%20of%20Sector%20Regulators%20\(Febuary%202025\).xlsx](https://www.oecd.org/content/dam/oecd/en/publications/support-materials/2025/03/the-2023-indicators-on-the-governance-of-sector-regulators_eb33f9bb/2023%20Indicators%20on%20the%20Governance%20of%20Sector%20Regulators%20(Febuary%202025).xlsx)
- **Access year:** 2025 (September)

- **Description:** Question-level governance data for sector regulators across telecoms, rail, air, and energy sectors. Used in Figure 3.6 panels (a–d) to show financing mechanisms by income tercile. Brazil’s second regulator entries (variables suffixed BRA2) are dropped to avoid double-counting. The original OECD file was lightly cleaned (sheet structure flattened) before import.
 - **License:** OECD terms of use
-

- **Filename:** World Development Indicators (pulled live via wbopendata)
 - **Source:** World Bank World Development Indicators
 - **URL:** <https://databank.worldbank.org/source/world-development-indicators>
 - **Access year:** 2025
 - **Description:** Retrieved live in Stata using the wbopendata package. Indicators used: NY.GNP.PCAP.CD (GNI per capita, current USD, for income tercile construction in Figure 3.6) and NY.GDP.PCAP.PP.KD (GDP per capita, PPP, retained for reference).
 - **License:** Creative Commons Attribution 4.0 International (CC BY 4.0)
-

- **Filename:** Excel compilation workbook and PDF norms in Datain/Ministerial_Turn_Over/
 - **Source:** Ministry of Justice of Peru (legal norms appointing ministers); Ministry of Transportation and Communications of Peru; OSIPTEL; ProInversión; Plataforma Digital Única del Estado Peruano
 - **URLs:** <https://spijweb.minjus.gob.pe/> · <https://www.gob.pe/institucion/mtc/normas-legales/> · <https://www.osiptel.gob.pe/> · <https://www.investinperu.pe/> · <https://www.gob.pe/>
 - **Access year:** 2025
 - **Description:** Individual legal norms recording ministerial appointments and key regulatory milestones for Peru’s telecommunications sector. Used to produce Figure 3.12 (ministerial turnover and regulatory milestones) entirely in Excel. The compilation workbook contains hyperlinks to each source norm; the corresponding PDF files are stored in Datain/Ministerial_Turn_Over/MCT_appointments/ and Datain/Ministerial_Turn_Over/Milestones/.
 - **License:** Public domain (Peruvian government legal records)
-

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 file.
-

Instructions for Replicators

New users should follow these steps to run the package successfully:

1. Ensure all required software and dependencies are installed as listed in the [Requirements](#) section.
2. For data files not included in the package (BOOST, GSPS, MAPS, AfDB), request access via the links in the [Data Sources](#) section and place them in the appropriate data/input/ subfolder or use the OneDrive folder shared by the team.
3. Open an R session with the working directory set to the root of this repository (simply opening replication.Rproj).
4. Run `renv::restore()` to install all R package requirements.
5. Run each chapter's scripts in order. If you have access to the OneDrive folder, you may choose to skip the prepdata step:

```
# Chapter 1
source("chapter-1/source/chapter_1_prepdata.R")
source("chapter-1/source/chapter_1_figures.R")

# Chapter 2
source("chapter-2/source/chapter_2_prepdata.R")
source("chapter-2/source/chapter_2_figures.R")

# Chapter 3
source("chapter-3/source/chapter_3_prepdata.R")
source("chapter-3/source/chapter_3_figures.R")
```

5. For the Stata/Excel figures in Chapter 3 (Figures 3.3, 3.6, and 3.12), use the folder chapter-3/stata-excel/:
 - Run Do/1_Figures.do in Stata. This script auto-installs the wbopendata package if needed and generates Do/2_label_vars.do at runtime — do not edit that file manually.
 - Figure 3.12 is produced in Excel using the workbook in Datain/Ministerial_Turn_Over/.
-

List of Exhibits

The provided code reproduces:

- ☐ All numbers provided in text in the paper

- ☒ All tables and figures in the paper
- ☐ Selected tables and figures in the paper, as explained and justified below

Exhibit	Output filename	Script
Figure 1.2	chapter-1/figs/fig_1_2.png	chapter-1/source/chapter_1_figures.R
Figure 1.3	chapter-1/figs/fig_1_3.png	chapter-1/source/chapter_1_figures.R
Figure 1.4	chapter-1/figs/fig_1_4.png	chapter-1/source/chapter_1_figures.R
Figure 1.5	chapter-1/figs/fig_1_5.png	chapter-1/source/chapter_1_figures.R
Figure 1.6	chapter-1/figs/fig_1_6.png	chapter-1/source/chapter_1_figures.R
Figure 1.7	chapter-1/figs/fig_1_7.png	chapter-1/source/chapter_1_figures.R
Figure 2.2	chapter-2/figs/fig2_2.png	chapter-2/source/chapter_2_figures.R
Figure 2.3	chapter-2/figs/fig2_3.png	chapter-2/source/chapter_2_figures.R
Figure 2.4	chapter-2/figs/fig2_4.png	chapter-2/source/chapter_2_figures.R
Figure 2.5	chapter-2/figs/fig2_5.png	chapter-2/source/chapter_2_figures.R
Figure 3.2	chapter-3/figs/fig_3_2.png	chapter-3/source/chapter_3_figures.R
Figure 3.3	chapter-3/stata-excel/Output/Charts/Figure_3.3_comms_price.png	chapter-3/stata-excel/Do/1_Figure s.do (Stata)
Figure 3.4	chapter-3/figs/fig_3_4.png	chapter-3/source/chapter_3_figures.R
Figure 3.5	chapter-3/figs/fig_3_5.png	chapter-3/source/chapter_3_figures.R

Exhibit	Output filename	Script
Figure 3.6a	chapter-3/stata-excel/Output/Charts/Fig_3_6_Comms_Sources_graph_Q2b_a_22.png	chapter-3/stata-excel/Do/1_Figures.do (Stata)
Figure 3.6b	chapter-3/stata-excel/Output/Charts/Fig_3_6_Rail_Sources_graph_Q3ab_a_22.png	chapter-3/stata-excel/Do/1_Figures.do (Stata)
Figure 3.6c	chapter-3/stata-excel/Output/Charts/Fig_3_6_Air_Sources_graph_Q3bb_a_22.png	chapter-3/stata-excel/Do/1_Figures.do (Stata)
Figure 3.6d	chapter-3/stata-excel/Output/Charts/Fig_3_6_Energy_Sources_graph_Q1b_a_22.png	chapter-3/stata-excel/Do/1_Figures.do (Stata)
Figure 3.7	chapter-3/figs/fig_3_7.png	chapter-3/source/chapter_3_figures.R
Figure 3.8	chapter-3/figs/fig_3_8.png	chapter-3/source/chapter_3_figures.R
Figure 3.9	chapter-3/figs/fig_3_9.png	chapter-3/source/chapter_3_figures.R
Figure 3.10	chapter-3/figs/fig_3_10.png	chapter-3/source/chapter_3_figures.R
Figure 3.11a	chapter-3/figs/fig_3_11a.png	chapter-3/source/chapter_3_figures.R
Figure 3.11b	chapter-3/figs/fig_3_11b.png	chapter-3/source/chapter_3_figures.R
Figure 3.12	chapter-3/stata-excel/Datain/Ministerial_Turn_Over/	Excel workbook (manual)

Requirements

Computational Requirements

The analysis was developed and tested on the following system:

- **Operating System:** Windows 11
- **Processor:** Intel Core i7
- **RAM:** 16 GB

Software Requirements

- **R** *(version 4.5.2)
- **Stata** (*version 17 or later*)
 - `wbopendata` — auto-installed by `1_Figures.do` if not present (requires internet connection at runtime)
- **Microsoft Excel** (*any recent version*) — required to reproduce Figure 3.12 manually from the workbook in `chapter-3/stata-excel/Datain/Ministerial_Turn_Over/`

Memory, Runtime, and Storage Requirements

- **Approximate runtime:** < 10 minutes for all chapters combined
 - **Storage required:** ~500 MB (including all input data)
-

Code Description

Each chapter follows the same two-script structure:

- **chapter_X_prepdata.R** — Reads raw input data, cleans and recodes variables, and writes processed datasets to `data/output/`.
- **chapter_X_figures.R** — Reads processed data from `data/output/` and produces all figures saved to `figs/`.

Chapter 3 also uses: - **chapter-3/source/funs.R** — Helper functions (`plot_scores`, `summarize_scores`) used by `chapter_3_figures.R`.

Three figures in Chapter 3 are produced outside of R, using materials in `chapter-3/stata-excel/`: - **Do/1_Figures.do** — Master Stata script. Produces Figures 3.3 and 3.6 (panels a–d). Reads PMR, ITU, and GSR Excel inputs from `Datain/`, pulls WDI data live via `wbopendata`, and writes PNG output to `Output/Charts/`. Also auto-generates `Do/2_label_vars.do` at runtime — do not edit that file manually. - **Figure 3.12** — Produced manually in Excel from the workbook in `Datain/Ministerial_Turn_Over/`. Not generated by any script.