

README for the Reproducibility package for Industrial Policy for Development: Approaches in the 21st Century

Contents

Overview.....	1
Data Availability Statement.....	1
Data Sources and Availability.....	1
Instructions for Replicators	8
List of Exhibits.....	10
Requirements	12
Folder Structure.....	13

Overview

This readme file includes all the replication material for the World Bank Policy Research Report *Industrial Policy for Development* by Ana Margarida Fernandes and Tristan Reed.

Data Availability Statement

The authors of the report certify they have legitimate access to and permission to use the data used in this manuscript. Most of the data are included in the reproducibility package. Some of the data not included can be requested from the author team while other data cannot be shared.

Data Sources and Availability

The datasets labelled as “Restricted” are not publicly available, and external users cannot request access to them. Details on the data and sources can be found in the report’s Data Appendix or the datasets section of the reproducibility package metadata.

Data Sources

Filename: “Data\raw\DTA 1.0 - Horizontal Content (v2).xlsx”

- **Name of the data:** World Bank Deep Trade Agreements Dataset
- **Source:** Mattoo, A., N. Rocha, and M. Ruta. 2020. *Handbook of Deep Trade Agreements*. Washington, DC: World Bank.
- **URL:** [https://datacatalogfiles.worldbank.org/ddh-published/0065624/DR0093615/DTA%201.0%20-%20Horizontal%20Content%20\(v2\).xlsx](https://datacatalogfiles.worldbank.org/ddh-published/0065624/DR0093615/DTA%201.0%20-%20Horizontal%20Content%20(v2).xlsx)
- **Accessed:** May 2024

- **Description of the dataset:** This dataset maps the coverage of 52 selected policy areas in 400 PTAs notified at WTO signed between 1958 and 2023. It also includes information about legal enforceability of each policy area. This dataset captures the extensive margin of the content of Deep trade agreements. The latest version of the data includes 400 PTAs signed between 1958 and 2023.
- **Relevant for:** Figure B1.1.1

Filename: "Data\raw\WBG_IncomeClassification_FY25.dta"

- **Name of the data:** World Bank Country and Lending Groups
- **Source:** World Bank Group Country and Lending Groups. *Washington, DC: World Bank.*
- **URL:** <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519>
- **Accessed:** September 2025
- **Relevant for:** Figure B1.1.1

Filename: "Data\raw\comparative_advantage.xlsx" (sheet: solar potential)

- **Source:** ESMAP. 2020. Global Photovoltaic Power Potential by Country. Washington, DC: World Bank.
- **URL:** <https://globalsolaratlas.info/global-pv-potential-study>
- **Accessed:** 2025
- **Relevant for:** Figure 4.1

Filename: "Data\raw\comparative_advantage.xlsx" (sheet: industrial policy)

- **Source:** Climate actions and policies measurement framework
- **URL:** https://sdmx.oecd.org/public/rest/data/OECD.ENV.EPI,DSD_CAPMF@DF_CAPMF,1.0/all?dimensionAtObservation=AllDimensions&format=csvfilewithlabels
- **Accessed:** 2025
- **Relevant for:** Figure 4.1

Filename: "Data\raw\wgi_raw.dta"

- **Source:** World Bank Worldwide Governance Indicators (database)
- **URL:** <http://www.govindicators.org/>
- **Accessed:** July 2025
- **Relevant for:** Figure 2.9; Table 1.2

Filename: "Data\raw\GTA NIPO - August 2025.xlsx" [Restricted data]

- **Name of the data:** New Industrial Policy Observatory database
- **Source:** Evenett, S., Jakubik, A., Martín, F. and Ruta, M. (2024) The return of industrial policy in data. *The World Economy*, 47(7), pp.2762-2788.
- **URL:** <https://globaltradealert.org/reports/new-industrial-policy-observatory-nipo>
- **Accessed:** August 2025
- **Relevant for:** Figure B2.1.1

Filename: "Data\raw\EBRD_Juhaszetal_gpt_27May25.dta" [Restricted data]

- **Source:** Evenett, S., Jakubik, A., Martín, F. and Ruta, M. (2024) The return of industrial policy in data. *The World Economy*, 47(7), pp.2762-2788.

- **Accessed:** May 2025
- **Relevant for:** Figure B2.1.1

Filename:

"Data\raw\JLOP_Industrial_Policy_Data_for_World_Bank_Fernandes_Reed_April8_2025.dta"
[Restricted data]

- **Source:** Juhász, R., Lane, N., Oehlsen, E. and Pérez, V. (2023a) *The who, what, when, and how of industrial policy: A text-based approach*. STEG Working Paper No. 050.
- **Accessed:** May 2025
- **Relevant for:** Figure B2.1.1

Filename: "Data\raw\GTED_FullDatabase_20240701.xlsx"

- **Name of the data:** Global Tax Expenditure Database
- **Source:** Redonda, A., von Haldenwang, C. and Aliu, F. (2025) *Global Tax Expenditures Database*. Tax Expenditures Lab.
- **URL:** <https://gted.taxexpenditures.org/>
- **Accessed:** March 2025
- **Relevant for:** Figure 2.4

Filename: "Data\raw\OECD Database on Export Restrictions for Staple Crops 2024.xlsx"

- **Name of the data:** OECD Database on Export Restrictions for Staple Crops 2024
- **Source:** OECD (2024a) *OECD Inventory of Export Restrictions on Staple Crops*.
- **URL:** <https://www.oecd.org/en/topics/sub-issues/agro-food-trade/export-restrictions-on-staple-crops.html>
- **Accessed:** June 2025
- **Relevant for:** Figure 2.3

Filename: "Data\raw\oecd-industrial-raw-materials-complete-data_2023.xlsx"

- **Name of the data:** OECD Inventory of Export Restrictions on Industrial Raw Materials 2024
- **Source:** *OECD Inventory of Export Restrictions on Industrial Raw Materials 2024: Monitoring the use of export restrictions amid market and policy tensions*. Paris: OECD Publishing.
- **URL:** [https://data-explorer.oecd.org/vis?df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD_QDD_TAD_EXP_RESTRIC%40DF_QDD_TAD_EXP_RESTRIC&df\[ag\]=OECD.TAD.ADM&df\[vs\]=2.0&dq=....&pd=2009%2C&to\[TIME_PERIOD\]=false](https://data-explorer.oecd.org/vis?df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_QDD_TAD_EXP_RESTRIC%40DF_QDD_TAD_EXP_RESTRIC&df[ag]=OECD.TAD.ADM&df[vs]=2.0&dq=....&pd=2009%2C&to[TIME_PERIOD]=false)
- **Accessed:** June 2025
- **Relevant for:** Figure 2.3

Filename: "Data\raw\world-imf2024-v1.dta"

- **Source:** IMF (2024b) World Revenue Longitudinal Database
- **URL:** <https://www.imf.org/en/Topics/fiscal-policies/world-revenue-longitudinal-database>
- **Accessed:** May 2025
- **Relevant for:** Figure 2.8; Table 1.2

Filename: "Data\raw\wdi_lcu.csv"

- **Name:** Average LCU/USD exchange rate, (indicator PA.NUS.FCRF)
- **Source:** World Bank (2025d) *World Development Indicators (database)*.
- **URL:** <https://databank.worldbank.org/source/world-development-indicators>
- **Accessed:** July 2025
- **Relevant for:** Figures 2.10

Filename: "Data\raw\wdi gdp.csv"

- **Name:** GDP per capita (constant USD), (indicator NY.GDP.PCAP.KD)
- **Source:** World Bank (2025d) *World Development Indicators (database)*.
- **URL:** <https://databank.worldbank.org/source/world-development-indicators>
- **Accessed:** July 2025
- **Relevant for:** Figures 2.5

Filename: "Data\raw\gdp.csv"

- **Name:** Total GDP (constant USD), (indicator NY.GDP.MKTP.KD)
- **Source:** World Bank (2025d) *World Development Indicators (database)*.
- **URL:** <https://databank.worldbank.org/source/world-development-indicators>
- **Accessed:** July 2025
- **Relevant for:** Figures 2.10; Table 1.2

Filename: "Data\raw\gdp lcu.csv"

- **Name:** Total GDP (current LCU), (indicator NY.GDP.MKTP.CN)
- **Source:** World Bank (2025d) *World Development Indicators (database)*.
- **URL:** <https://databank.worldbank.org/source/world-development-indicators>
- **Accessed:** July 2025
- **Relevant for:** Figures 2.10

Filename: "Data\raw\income group.xlsx"

- **Name:** Historical classification of countries by income
- **Source:** World Bank (2025d) *World Development Indicators (database)*.
- **URL:** <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups>
- **Accessed:** July 2025
- **Relevant for:** Figures 2.2–2.4, 2.6, 2.7, 6.3

Filename: "Data\processed\clean\aggregate tariffs.dta"
"Data\processed\clean\t_2023.dta"

- **Name:** World Integrated Trade Solution (WITS) TRAINS tariff data (from World Bank) and BACI trade value data (from CEPII).
- **Source:** Ileana Cristina Costantinescu (ineagu@worldbank.org) [for raw data and code – *please do not contact yet*]
- **Accessed:** September 2025
- **Relevant for:** Figures 2.2, 2.8, 6.3

Filename: "Data\processed\clean\Core data.dta" [Restricted data]

- **Name:** IMF Government Finance Statistics; World Bank BOOST
- **Source:** Eric Anthony Lacey (elacey@worldbank.org) [for raw data and code]
- **Accessed:** November 2025

- **Relevant for:** Figure 2.4

Filename: "Data\raw\gppd_data.xlsx"

- **Name:** Global Public Procurement Database
- **Source:** World Bank (2020), Global Public Procurement Database (online).
- **URL:** <https://www.globalpublicprocurementdata.org/gppd/>
- **Accessed:** December 2025
- **Relevant for:** Figure 2.6

Filename: "Data\raw\population.csv"

- **Name:** Population, total (indicator SP.POP.TOTL)
- **Source:** World Bank (2025d) *World Development Indicators (database)*.
- **URL:** <https://databank.worldbank.org/source/world-development-indicators>
- **Accessed:** July 2025
- **Relevant for:** Table 1.2

Filename: "Data\raw\debt.xls"

- **Name:** General government gross debt (% of GDP)
- **Source:** IMF (2025) *World Economic Outlook: Global Economy in Flux, Prospects Remain Dim*. October 2025. Washington DC: International Monetary Fund.
- **URL:** <https://www.imf.org/en/publications/weo/issues/2025/10/14/world-economic-outlook-october-2025#tools>
- **Accessed:** October 2025
- **Relevant for:** Table 1.2

Filename: "Data\raw\gqii2023_public_dataset_v2.xlsx"

- **Name:** Global Quality Infrastructure Index scores
- **Source:** Harmes-Liedtke, U., M. Muñoz, and E. Waltos. 2024. *Global Quality Infrastructure Index Report 2023*.
- **URL:** <https://gqii.org/gqii-2023/>
- **Accessed:** August 2025
- **Relevant for:** Figure 2.5

Filename: "Data\raw\OECD_RDTAX_DB2025_1.xlsx"

- **Name:** Government direct funding and tax support for business R&D
- **Source:** OECD. 2025b. *R&D Tax Incentives Database*.
- **URL:** <https://oe.cd/rdtax>
- **Accessed:** September 2025
- **Relevant for:** Figure 2.7

Filenames: "Data\raw\GTA+Database+2008-2010.xlsx"

"Data\raw\GTA+Database+2011-2015.xlsx"

"Data\raw\GTA+Database+2016-2020.xlsx"

"Data\raw\GTA+Database+2021-latest.xlsx"

- **Name:** Global Trade Alert database
- **Source:** SGEPT. 2024. *Global Trade Alert (Database)*. St. Gallen Endowment for Prosperity through Trade.
- **URL:** <https://globaltradealert.org/data-center>

- **Accessed:** December 2024 – March 2025
- **Relevant for:** Figures B2.1.1, 2.8–2.10

Filename: *"Data\raw\Survey on policy dialogue about industrial policy(1-42)_cleaned.xlsx"*

- **Name:** Industrial policy survey data
- **Source:** Author-conducted survey
- **URL:** n/a
- **Notes:** Survey conducted using Microsoft Forms and then cleaned by hand to create variable names, code sectors, and code ministries.
- **Accessed:** April 2025
- **Relevant for:** Figure 1.1

Filename: *"Data\processed\clean\ndps_industries.dta"*

- **Name:** National Development Plans data
- **Source:** Created by the authors
- **URL:** n/a
- **Notes:** Includes information on national development plans read and recorded from multiple government sources, along with keyword analysis.
- **Accessed:** January – March 2025
- **Relevant for:** Figures 2.1, B5.1.1, 6.1

Filename: *"Data\processed\clean\wb investment projects.dta"*

- **Name:** Prior Actions and Projects data
- **Source:** World Bank
- **URL:** <https://www.worldbank.org/en/what-we-do/products-and-services/financing-instruments/development-policy-financing#3>
- **Notes:** Created by authors through manual cleaning and keyword searches.
- **Accessed:** March 2025
- **Relevant for:** Figure B5.1.1

Filenames: *"Data\processed\clean\wb prior actions.dta"*
"Data\processed\clean\dpad_database_FY23.dta"

- **Name:** World Bank Prior Actions and Projects data
- **Source:** World Bank
- **URL:** n/a
- **Notes:** Created by authors through manual cleaning and keyword searches.
- **Accessed:** June 2025
- **Relevant for:** Figure B5.1.1, B5.1.2

Filenames: *"Data\processed\clean\wb_soe_keywords.dta"*
"Data\processed\clean\wb_subtax_keywords.dta"
"Data\processed\clean\wb_trade_keywords.dta"

- **Name:** Prior Actions and national development plans keywords data
- **Source:** World Bank and national development plans for multiple governments
- **URL:** n/a
- **Notes:** Created by authors through manual cleaning and keyword searches.
- **Accessed:** June 2025

- **Relevant for:** Figure B5.1.2

Filename: "Data\processed\clean\Income Group data.dta"

- **Name:** Income group classification
- **Source:** World Bank
- **URL:** <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups>
- **Notes:** Generated by cleaning the file for historical classification.
- **Accessed:** June 2025
- **Relevant for:** Figure 2.1, B5.1.1, 6.1

Filename: "Data\processed\clean\GDP data.dta"

- **Name:** GDP data
- **Source:** World Bank
- **URL:** <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups>
- **Notes:** Downloaded from World Bank Open Data using the Stata command *wbopendata* and cleaning the data using the indicators NY.GDP.PCAP.PP.CD and NY.GDP.MKTP.PP.CD.
- **Accessed:** June 2024
- **Relevant for:** Figure 2.1, B5.1.1

Filename: "Data\processed\clean\green_policies.dta"

- **Name:** Climate Policy Database
- **Source:** NewClimate Institute, Wageningen University and Research, and PBL Netherlands Environmental Assessment Agency. 2024. *Climate Policy Database*
- **URL:** <https://climatepolicydatabase.org/about-legal>
- **Notes:** Cleaned and aggregated data from the online database.
- **Accessed:** July 2025
- **Relevant for:** Figure 6.2

Filename: "Data\processed\clean\S&T Fig 3.xlsx"

- **Name:** Share of economic sectors' global GHG emissions covered by an ETS or carbon tax
- **Source:** World Bank. 2025b. *State and Trends of Carbon Pricing 2025*. Washington, DC: World Bank
- **URL:** <https://www.worldbank.org/en/publication/state-and-trends-of-carbon-pricing>
- **Notes:** Data extracted from Figure 3 of the report.
- **Accessed:** August 2025
- **Relevant for:** Figure 6.2

Filename: "Data\raw\agency_year.xlsx"

- **Name:** Export promotion agencies and years of establishment
- **Source:** Authors
- **URL:** n/a
- **Notes:** Data aggregated by authors through online search.

- **Accessed:** July 2025
- **Relevant for:** Figure 5.2

Filename: “Data\processed\clean\group.dta”

- **Name:** Country regional and income group classification
- **Source:** World Bank
- **URL:** <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups>
- **Notes:** Generated by cleaning the file for historical classification.
- **Accessed:** July 2025
- **Relevant for:** Figure 5.2

Filename: “Data\raw\GVC5_product_taxonomy.dta”

- **Name:** GVC taxonomy
- **Source:** Rosenow, S.K. and Mealy, P. 2024. Turning Risks into Reward: Diversifying the Global Value Chains of Decarbonization Technologies. *Policy Research Working Paper*, no. 10696. Washington, DC: World Bank.
- **URL:** <https://openknowledge.worldbank.org/entities/publication/75964bc2-4115-4178-9d80-fb7518ebd4b9>
- **Notes:** Updated data based on Rosenow and Mealy (2024).
- **Accessed:** October 2025
- **Relevant for:** Figure 6.3

Instructions for Replicators

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

1. Save the zipped replication package in a folder of your choice and unzip it. The unpacking will include three subfolders named “Code”, “Data” and “Output”, respectively. Note that the “Output” folder contains an empty subfolder (“figures”), where the replication code will save the graphs produced.
2. Users must first have access to all data files if they are not included in the reproducibility package. They should place them in the data folder.
3. Update the following files with your local directory path (i.e., the path for the folder where the unzipped replication package is saved):
 - a. The setup script *1-main.do* in the folder “Code”
 - b. “Code\Data analysis and outputs\fig 4_1.R” (for Figure 4.1)
 - c. Users can also add the local directory path in individual .do files for all Figures for which individual codes are available (i.e., all except Figs. 4.2 and 5.1)

- all_figures.do
- fig 4.1.R

List of Exhibits

The scripts listed above reproduce all figures as presented in the report except for:

- Figure B1.1.1, to be created in Microsoft Excel. The .xlsx file is created by running the Stata .do file “fig b1.1.1.do” or “all_figures.do”
- Figure 4.2, taken from Reed (2024)¹
- Figure 5.1, adapted from Criscuolo and Palmade (2008)²

Additionally, separate scripts are provided in the same folder to replicate each individual figure:

Exhibit name	Output filename	Script	Note
Figure 1.1	fig 1_1.png	Fig 1.1.do	- Found in Replicability\Output\figures - Also produced by script all_figures.do
Figure B.1.1.1	fig b1_1_1.xlsx	fig b1.1.1.do	- Found in Replicability\Output\figures - Also produced by script all_figures.do - The data in the MS Excel file can be used to replicate Figure B1.1.1 by selecting a combination graph with: line graph for “Cumulative Preferential Trade Agreements in Force” (scale on secondary axis, dash pattern, 1pt width, color orange accent 2); stacked column for “With legally enforceable Subsidies provisions” (automatic color); and stacked column for “Without legally enforceable Subsidies provisions” (color white, background 1, darker 25%)
Figure B2.1.1	fig b2_1_1.png	fig b2.1.1.do	- Found in Replicability\Output\figures - Also produced by script all_figures.do
Figure 2.1	fig 2_1.png	fig 2.1.do	- Found in Replicability\Output\figures - Also produced by script all_figures.do

¹ Reed, T. (2024) Export-Led Industrial Policy for Developing Countries: Is There a Way to Pick Winners? *Journal of Economic Perspectives*. Fall 2024.

² Criscuolo, A., and V. Palmade. 2008. *Reform Teams*. Public Policy for the Private Sector; Note No. 318. Washington, DC: World Bank.

Exhibit name	Output filename	Script	Note
Figure 2.2	fig 2_2a.png fig 2_2b.png	fig 2.2.do	- Found in Output\figures - Also produced by script all_figures.do
Figure 2.3	fig 2_3a.png fig 2_3b.png	fig 2.3.do	- Found in Output\figures - Also produced by script all_figures.do
Figure 2.4	fig 2_4a.png fig 2_4b.png	fig 2.4.do	- Found in Output\figures - Also produced by script all_figures.do
Figure 2.5	fig 2_5.png	fig 2.5.do	- Found in Output\figures - Also produced by script all_figures.do
Figure 2.6	fig 2_6.png	fig 2.6.do	- Found in Output\figures - Also produced by script all_figures.do
Figure 2.7	fig 2_7.png	fig 2.7.do	- Found in Output\figures - Also produced by script all_figures.do
Figure 2.8	fig 2_8a.png fig 2_8b.png	fig 2.8.do !	- Found in Output\figures - Also produced by script all_figures.do
Figure 2.9	fig 2_9.png	fig 2.9.do !	- Found in Output\figures - Also produced by script all_figures.do
Figure 2.10	fig 2_10a.png fig 2_10b.png	fig 2.10.do !	- Found in Output\figures - Also produced by script all_figures.do
Figure 4.1	fig 15a.png fig 15b.png	fig 4.1.R	- Found in Output\figures
Figure 5.2	fig 5_2.png	fig 5.2.do	- Found in Output\figures - Also produced by script all_figures.do
Figure B5.1.1	fig b5_1_1.png	fig b5.1.1.do	- Found in Output\figures - Also produced by script all_figures.do
Figure B5.1.2	fig b5_1_2.png	fig 5.1.2.do	- Found in Output\figures - Also produced by script all_figures.do
Figure 6.1	fig 6_1.png	fig 6.1.do	- Found in Output\figures - Also produced by script all_figures.do
Figure 6.2	fig 6_2a.png fig 6_2b.png	fig 6.2.do	- Found in Output\figures - Also produced by script all_figures.do

Exhibit name	Output filename	Script	Note
Figure 6.3	fig 6_3.png	fig 6.3.do	- Found in Output\figures - Also produced by script all_figures.do

Table 1.2 can be found in the folder “Output\tables”. It can be replicated by extracting the following data:

- Average GDP for 2021-2023 using Output\raw\gdp.csv;
- Average government effectiveness scores for 2022-2024 using Output\raw\wgi_raw.dta;
- Average population for 2021-2023 using Output\raw\population.csv;
- Average debt/GDP ratios for 2021-2023 using Output\raw\debt.xls
- Average tax revenue/GDP ratios for 2020-2022 using Data\raw\world-imf2024-v1.dta

All data can then be included in one sheet by using MS Excel’s xlookup function (sheet “all” in *table 1_2.xlsx*), only retaining countries for which all data are available. Median values should be calculated for GDP, GE scores, tax/GDP and debt/GDP ratios in order to compare individual country values and establish relative size long the three dimensions of interest.

Requirements

Software Requirements

All figures in the report – except for Figures B1.1.1, 4.1, 4.2 and 5.1 – have been obtained using Stata MP version 18.5, with the following packages installed:

- isocodes
- tolong
- sxpose

For Figure B1.1.1, the data was initially cleaned using Stata, with the resulting data exported as a Microsoft Excel file (to “Output\figures\fig b1_1_1.xlsx”) by the do file.

Figure 4.1 requires R version 15 with the following packages:

- dyp1r
- ggplot2
- haven
- readxl
- writexl

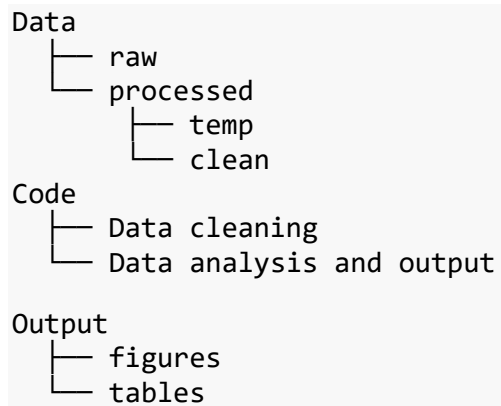
Table 1.2 is generated entirely using Microsoft Excel.

Memory and Runtime and Storage Requirements

All data (raw, temporary and clean) and outputs require approximately 75GB of storage. Certain scripts are computationally intensive (marked with a “!”) and may require longer processing times. As such, at least 16GB (ideally 32GB) of RAM or access to a cloud computing environment are recommended.

Folder Structure

The following schematic describes the folder structure for this replication package:



```
graph LR; Data --> raw; Data --> processed; processed --> temp; processed --> clean; Code --> "Data cleaning"; Code --> "Data analysis and output"; Output --> figures; Output --> tables;
```

The diagram illustrates the folder structure of the replication package. It is organized into three main categories: Data, Code, and Output. The Data category includes 'raw' and 'processed' folders, with 'processed' further subdivided into 'temp' and 'clean'. The Code category includes 'Data cleaning' and 'Data analysis and output'. The Output category includes 'figures' and 'tables'.