

# Are the Poorest Catching Up? README

**Authors: James Cust, Paul Collier and Alexis Rivera-Ballesteros**

**May 2, 2024**

## Contents

1. Overview
2. Computational requirements
3. Data Availability Statement
4. Data sources description
5. Description of files
6. License

## 1. Overview

The code of this reproducibility package cleans the data from raw files obtained from different sources and creates the figures for the World Bank's Policy Research Working Paper 10622 "Are the Poorest Catching Up?" using Stata.

The `0_Master` do file in the Code runs all the code and produces all the figures in the paper. To run all the do files, please change the project path with yours at line 19 of the `0_Master` do file.

## 2. Computational requirements

The code was last run with Stata 15 SE version on a Macbook. It requires to install two Stata packages which can be installed by running the file `1_Package_Setup` in the Code folder:

- `wbopendata`
- `mipolate`

All files are in Stata or Excel format.

## 3. Data Availability Statement

All data are obtained from public sources except the data required to produce figures 13, 14, 16 and 17, where we only report aggregated data.

## 4. Data sources description

This paper uses publicly available data from 9 sources downloaded to the `Data/base` and `Data/clean` folders and are described below:

- `wdi.dta` contains macroeconomic indicators obtained from the [World Bank's World Development Indicators 2023 update](#) via the [DataBank API for Stata](#). The indicators used are:
  - `NY.GDP.MKTP.KD` – GDP (constant 2015 US\$)
  - `NY.GDP.PCAP.KD` – GDP per capita (constant 2015 US\$)

- NY.GDP.MKTP.PP.KD – GDP, PPP (constant 2017 international \$)
  - SP.POP.TOTL – Population, total
  - NY.GDP.TOTL.RT.ZS – Total natural resources rents (% of GDP)
  - SP.POP.GROW – Population growth (annual %)
  - NY.GNP.MKTP.KD – GNI (constant 2015 US\$)
  - NE.GDI.FTOT.KD – Gross fixed capital formation (constant 2015 US\$)
- **pwt100.dta** contains GDP and other macroeconomic indicators from [Penn World Table 10.0](#).
  - **pip.csv** contains surveyed income data from [World Bank's Poverty and Inequality Platform](#). The data is downloaded from the [Poverty and Inequality Indicators dashboard](#), selecting 2011 as the base year from the top dropdown menu, and hitting the "Download all country data" from the download options at the bottom of the left sidebar.
  - **lm\_wpid\_web.dta** contains the inputs to replicate the analysis from [Lakner and Milanovic \(2013\)](#).
  - **CWON2021 Country Tool – Full Dataset.xlsx** contains wealth accounting data from the [Changing Wealth of Nations 2021](#) report.
  - **CMO-Historical-Data-Monthly.xlsx** contains data on commodity prices from [World Bank Commodity Markets data](#)
  - **combined\_data.dta** contains the inputs to replicate the analysis from [Patel, Sandefur, and Subramanian \(2021\)](#). The replication code and data are obtained from [this link](#) that the authors provided in their [Working Paper version](#).
  - **si.pov.dday.csv** contains the [Poverty headcount ratio at \\$1.90 a day \(discontinued\)](#) series that are no longer part of the WDI indicators, due to a [change in the methodology of poverty estimation](#).
  - **bb\_countries.dta** contains the list of countries that are classified as part of the Bottom Billion group in [Collier \(2009\)](#)
  - **oecd.dta** contains the list of OECD countries as of December 2023 listed in the [OECD official webpage](#).
  - **f\_povhc\_bygroup**, **h\_povhc\_bygroup**, and **povhc\_byregion** are datasets with historical and projected poverty headcount estimates constructed from country-year estimates from multiple sources that include publicly available data from World Bank's PovCalnet data but also from other sources that are not public including [World Bank's Global Poverty and Inequality Data](#) team and linear interpolation for gap filling.

## 5. Description of files

- Code folder: this folder contains all the Stata do files needed to clean the raw data and produce the figures in the manuscript. The `0_Master.do` file runs all the codes. The `1_Package_Setup.do` file installs the `wbopendata` Stata package to call WDI data from the World Bank's API. The `2_Data_Construction.do` file cleans the raw data and performs the necessary adjustments for the data. The `3_Tables_Figures.do` data produces the 59 charts in the manuscript, including those inserted in Annex A and B.
- Data folder: the folder contains two subfolders – base and clean.
  - Base: this folder contains all the raw data described in section 4 in Stata and Excel format.
  - Clean: this folder contains data that has been processed and it is ready to use to produce the figures. All files are in Stata format.

- Output folder: this folder contains 59 charts in PDF format for the 24 figures in the manuscript and the 6 figures in the Annexes. Each file is labeled with its corresponding figure number (and figure panel) for easy identification in the manuscript.

## 6. License

This README is under a CC-BY license.

