

Replication Package for "The Contribution of Human Capital to Current and Future Growth: an Extension of the World Bank's Long Term Growth Model (LTGM-HC)"

By Arthur Mendes (agalegomendes@worldbank.org and agalegomendes@gmail.com) and Steven Pennings (spennings@worldbank.org and steven.pennings@gmail.com)
October 2025

This file explains how to replicate the figures and tables in the main text and appendix of the paper. All results are generated using Stata 16.1 for Mac. To install Stata package X type findit X. Code takes about 30 minutes to run on Mac with Intel Core i9.

Directions:

* Run: **MP2025.do in STATA**, after adjusting the directory on line 22. This file has 7 parts

- Part 1 will call "Dofiles/Masterdataset.do" to generate the master dataset and save as Datasets/dataset_3d
- The raw data used to generate the master dataset is described below
- Part 2 will use the data from Part 1 to generate scenarios for each human capital component (Schooling, Quality, ASR, and NSR) and save as "Datasets/dataset_hc_assumptions.dta"
- Part 3 will use the data from Part 2, call Dofiles/LTGMHC.do to generate human capital growth scenarios for each country (saved as "Datasets/dataset_ghc.dta")
 - Scenario s1: Middle of the road baseline
 - Scenario s2: Status quo
 - Scenario s3: Moderate trend
 - Scenario s4: Optimistic trend
- Part 4 will use the data from Part 3 to generate GDP growth scenarios (saved as "Datasets/dataset_gdp.dta")
- Dofiles/LTGMv56data will generate the inputs dataset "Datasets/LTGM_data_tab560_153" needed to run the LTGM for each country (153)
- Dofiles/GDP_assumptions.do will set the assumptions to run the LTGM or LTGM-NR for each country
- Dofiles/LTGM runs the main LTGM using Datasets/LTGM_data_tab560_153.dta
- Dofiles/LTGMNR runs the LTGM-NR using Datasets/LTGM_data_tab560_153.dta and LTGMNR/0_LTGMNR_master.xlsx (Simulated data, see below a description of how this data is generated)
 - Scenario s1-s4: As above (Part 3)
 - Scenario s5: No human capital growth
- Part 5 calls do files to generate all results figures and tables from the main text. (Note: Table 1 is a summary of parameters and data sources, not a "result", and so is not replicated)
- FIGURES: MP2025.do calls do files to produce each figure "/Dofiles/Figures/FiguresXX.do" and saves as PNG in /Figures

- The only exception is Figure 1, which is generated by Excel file Figure1.xlsx. The data is linked to "Figure/Figure1_inputfromstata.xlsx" which is generated by Stata code MP2025.do
- TABLES: MP2025.do calls do files for each table "/Dofiles/Tables/TablesXX.do" and saves output in /Tables
- ManuscriptTables234.xlsx constructs tables 2, 3 and 4 as in the paper (using inputs Tables/Tables1,2,3.xlsx generated from Stata)
- Parts 6 generates human capital scenarios including tertiary education (Appendix A)
- Calls /Dofiles/LTGMHC_tertiary.do
- Parts 7 and 8 will generate all figures and tables from the appendix, including the Online Appendix Tables
- MP2025.do calls a do file for each appendix figure XX "/Dofiles/Figures/AppendixFiguresXX.do" and saves most figures as PNG files in /Figures.
- The exception is Appendix Figure B1 is generated by AppendixFigureB1.xlsx, mostly using calculations in the sheet. Any data (from cell Z29 onwards) is from \RawData\HCI\hci_data_september_2020.xlsx
- MP2025.do calls do files for each appendix table "/Dofiles/Tables/AppendixTablesXX.do" and saves output in /Tables
- (1) ManuscriptAppendixTables123.xlsx constructs Appendix tables 1, 2,3 as in the paper
- (2) LTGMHC_OnlineAppendixTables_October2025.xlsx constructs the Online Appendix Tables

****Directory list****

Root directory:

- Master do file MP2025.do

Other directories:

/RawData Stores raw data

/RawData/LTGM Stores raw data needed to run the LTGM (see description below)

/DoFiles Stores STATA do files called in MP2025.do

/Datasets Stores main datasets

/Figures Stores figures (main text and appendix)

/Tables Stores table (main text and appendix).

/Logs Directory to store the Stata log files

/LTGMNR Simulated data in the LTGMNR

* LTGM-NR downloaded from the LTGM website: <https://www.worldbank.org/LTGM>

→ Section: "Natural Resource Extension (LTGM-NR)"

→ Latest version: LTGM-NR v1.0 (Excel file)

* The LTGM-NR is set to ~default calibration for each natural resource rich country in our

sample and saved as ISO_LTGM_NR.xlsx (e.g., AGO_LTGM_NR.xlsx).

* This step is based on (and described in): Decerf, Benoit Marie A.; Friedman, Jed; Galego Mendes, Arthur; Pennings, Steven Michael; Yonzan, Nishant. Lives, Livelihoods, and Learning: A Global Perspective on the Well-Being Impacts of the COVID-19 Pandemic. Policy Research Working Paper No. 10728. Washington, DC: World Bank. <http://documents.worldbank.org/curated/en/099537503202413127>

* 0_LTGMNR_master.xlsx to collect the simulated output across the 41 countries. This is called in Part 4 to generate the GDP simulations with for natural resource rich countries.

*** Raw data sources *****

1. WCDE: <https://dataexplorer.wittgensteincentre.org/wcde-v2/> Reference to underlying report: <https://publications.jrc.ec.europa.eu/repository/handle/JRC111148>
2. HCI: This can be downloaded from https://datacatalogfiles.worldbank.org/ddh-published/0038030/DR0046209/hci_data_september_2020.xlsx and <https://www.worldbank.org/en/publication/human-capital>
3. WDI: <https://data.worldbank.org/indicator>
4. WHO: [https://www.who.int/data/gho/data/indicators/indicator-details/GHO/gho-jme-country-children-aged-5-years-stunted-\(-height-for-age--2-sd\)](https://www.who.int/data/gho/data/indicators/indicator-details/GHO/gho-jme-country-children-aged-5-years-stunted-(-height-for-age--2-sd)) - Data downloaded on October 7, 2025
5. UNWPP 2024 and 2022: <https://population.un.org/wpp/>
6. BL: <http://barrolee.com>
7. PWT 10.1: <https://www.rug.nl/ggdc/productivity/pwt/pwt-releases/pwt1001>
8. Global Data Set on Education Quality (1965-2015) - Altinok, Angrist, and Patrinos (2018): <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/706141516721172989/global-data-set-on-education-quality-1965-2015> -
9. RawData/LTGM --- Raw data to run the LTGM:
 - *1 WDI - http://databank.worldbank.org/data/download/WDI_csv.zip;
 - *2 PWT 9 - <https://www.rug.nl/ggdc/docs/pwt90.dta>
 - *3 PWT 10.01 - <https://www.rug.nl/ggdc/productivity/pwt/pwt-releases/pwt1001>
 - *4 Class file from the World Bank that includes income groups and regions
 - *5 UNWPP 2024 Population Data - <https://population.un.org/wpp/>
 - *6 IMF WEO <https://www.imf.org/en/Publications/WEO/weo-database/2024/October>

Approximated data download date:

1. WCDE: Sep 18, 2024

2. HCI: April 24, 2024
3. WDI: August 28, 2025
4. WHO: August 26, 2025
5. UNWPP 2024 and 2022: August 21, 2025
6. BL: July 31, 2024
7. PWT 10.1: June 17, 2024
8. Global Data Set on Education Quality (1965-2015): October 13, 2025
9. RawData/LTGM --- Raw data to run the LTGM:
 - *1 WDI: July, 2024
 - *2 PWT 9: June 24, 2016
 - *3 PWT 10.01: June 17, 2024
 - *4 Class file from the World Bank that includes income groups and regions: July, 2024
 - *5 UNWPP 2024 Population Data: August 21, 2025
 - *6 IMF WEO: April 2024

Disclaimer: The views expressed here are the authors' and do not necessarily reflect those of the World Bank, its executive directors, or the countries they represent.