Reproducibility Package for A Global Profile of Work and Poverty: Two Decades of Evidence

This reproducibility package contains the Stata code, input data structure, intermediate files, and generated outputs used for two products: Chapter 4 of the Poverty, Prosperity, and Planet Report 2026 (3PR) and the background paper *A Global Profile of Work and Poverty: Two Decades of Evidence*. The workflow is controlled by 01.code/dofile/master.do, whose comments provide the crosswalk between report and background-paper exhibits. The master prepares percentile and country-specific datasets, runs simulation inputs, generates figures, and exports annex tables.

The country survey files in 03.intermediate/2003, 03.intermediate/2013, and 03.intermediate/2024 are the authoritative country-file folders for this package. The current master workflow processes 2003 and 2024 in the percentile and country-specific preparation loops; 2013 is retained as an authoritative country-file year and can be used by do-files that explicitly include it or by future extensions to the master workflow.

This working-paper package is linked to the separate 3PR Chapter 4 exhibit package in ../07.reproducibility_3pr. The `DataFig*.dta` and `DataTab*.dta` files used by the Chapter 4 package are produced here in 04.output/DataFig, then copied to the Chapter 4 package's `1-data/processed` folder, where the report-specific do-files render the final Chapter 4 figures and tables.

Contents

1. Overview
2. Data Availability
3. Instructions for Replicators
4. List of Exhibits
5. Requirements
6. Code Description
7. Folder Structure
8. Notes and Caveats

Overview

The package is organized as a Stata replication archive. The master do-file sets directory globals, points to the input and intermediate folders, and then runs the data preparation, simulation, figure, and table scripts in sequence.

Main inputs include country-level survey microdata, PIP poverty inputs, population and growth data, WDI population-age files, country classification files, sector classification inputs, and ILOSTAT employment-rate extracts (ICLS 13th concept) used for GMD-vs-ILO comparison checks. The main outputs are PNG figures in 04.output/Figures and table workbooks in 04.output/Tables.

The exhibit-level data outputs in 04.output/DataFig are the handoff from this working-paper workflow to the 3PR Chapter 4 exhibit package. Re-run this package to regenerate those `DataFig*.dta` and `DataTab*.dta` files before updating ../07.reproducibility_3pr/1-data/processed.

At a high level, the workflow does the following:

1. Prepares percentile country files and combines them into global datasets.
2. Prepares country-specific group datasets and simulation inputs.
3. Generates figures shared by 3PR Chapter 4 and the background paper, including their annexes.
4. Produces annex tables and supporting output datasets.

Data Availability

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

The package uses prepared survey, distribution, population, price, PPP, and classification files included in this archive. Some upstream source microdata are restricted by licensing or access permissions. The exhibit-level source-data files in 04.output/DataFig can be made

publicly available; the access restrictions apply to upstream source microdata and restricted country-level intermediate files. Replicators should confirm data rights before redistributing or rebuilding outside authorized environments.

Restricted survey-derived inputs have been archived on the World Bank's Development Data Hub (DDH) to support reproducibility. The relevant DDH record is Lineup Data for A Global Profile of Work and Poverty: Two Decades of Evidence., dataset 0067178. Access is confidential/requested and remains subject to the terms attached to each resource.

Data archived in DDH

The DDH record contains the following six upload resources. The table gives the archive name, the package location it restores, and its role in the workflow.

DDH archive	Contents / provenance	Extract or place in	Use in the reproducibility workflow
`DDHdofile.zip`	Project Stata do-files for the reproducibility workflow.	01.code/dofile	Runs the preparation, combination, analysis, figure, simulation, and table commands.
`02.input.zip`	Fixed-vintage prepared input data, including macro-demographic, population, classification, price/PPP, and sector inputs.	02.input	Supplies the support data used by the do-files. These files are used as provided; the package does not recreate all upstream source systems.
`2003.zip`	Restricted country-survey files for the 2003 reference-year lineup.	03.intermediate/2003	Input to the 2003 country-lineup preparation and downstream percentile and country-specific datasets.
`2013.zip`	Restricted country-survey files for the 2013 reference-year lineup.	03.intermediate/2013	Input to the 2013 country-lineup preparation and downstream percentile and country-specific datasets.
`2024.zip`	Restricted country-survey files for the 2024 reference-year lineup.	03.intermediate/2024	Input to the 2024 country-lineup preparation and downstream percentile and country-specific datasets.
`03.intermediate.zip`	Additional prepared intermediate resources, including survey-list and PIP-related inputs used by the workflow.	03.intermediate, preserving the archive's folder structure	Provides supporting lineup metadata, poverty inputs, and other intermediate resources used by preparation, table, and analysis scripts.

`02.input.zip`: input data

The `02.input.zip` archive contains the fixed-vintage support inputs used by the do-files. These files are provided as prepared inputs; the package does not recreate all upstream APIs, catalogs, or institutional systems.

File	Contents and provenance	Use in this package
growth_gdp_pce_pop.dta	Country-year population, GDP, private consumption, per-capita values, and growth measures, prepared from PIP/API macro-demographic inputs (1977--2026; updated April 2026).	Population weighting and macro-demographic merges throughout the figure, simulation, and table workflows.

File	Contents and provenance	Use in this package
Survey_price_framework.dta	Survey-level PIP price and welfare metadata, including datatype, comparability, CPI/PPP replication fields, coverage, and domain settings (1963--2025). Obtained from the World Bank `Support_2005_CPI_V15_M` GMDRAW resource.	Adds survey datatype and comparability fields to the survey-list tables.
wdi_1564_pop.dta	Female, male, and total population ages 15--64 from WDI indicators `SP.POP.1564.FE.IN`, `SP.POP.1564.MA.IN`, and `SP.POP.1564.TO` (1960--2025; downloaded March 2026).	Sex-specific and total working-age population weights.
WDI_code_pop_un1564.dta	Population ages 15--64 as a share of total population from WDI indicator `SP.POP.1564.TO.ZS` (1960--2024).	Working-age population shares used in percentile, growth, figure, and table calculations.
code_pop_un_cohorts_sh.dta	Country-year population by sex and five-year age cohort from ages 15--64, with cohort shares (1950--2100; prepared from UN population data, March 2026).	Age-specific robustness analysis, including the 25--64 variants.
CLASS.dta	Historical World Bank country classifications for fiscal years 1989--2026, including region, income group, IDA/lending, Sub-Saharan Africa subregion, and FCV status (FY26 classification updated September 2025).	Region, income-group, and related classifications used by figures and simulations.
country_list_name.dta	Country names and labelled World Bank/PIP regional, income, FCV, IDA, and world-group classifications for 218 economies.	Country names in listing-table outputs.
country_list2.dta	Numeric-coded country lookup for 218 economies, including region, income group, FCV, IDA, Africa split, and PIP regional classifications.	Income-group classification in listing-table outputs.
CHN_IND_list.dta	Six-record China/India survey crosswalk with country code, lineup year, survey year, survey name, datatype, and comparability.	Appended to the survey metadata used to construct Tables A1 and A2.
List_5sectors.xlsx	Survey-to-file crosswalk for the five-sector analysis; the `Data` sheet maps country, year, survey, and analysis filename, while `Datanew` contains extended workflow fields.	The five-sector simulation imports the `Data` sheet to identify eligible survey files.
ILO/clist_with_wb_code.csv	Country-code crosswalk used to align ILOSTAT records to package country codes.	Mapping input for ILOSTAT processing in the GMD-vs-ILO comparison workflow.
ILO/EMP_5WAP_SEX_AGE_RT_A-20260831T1556.dta	ILOSTAT employment-to-population rate extract for 15+ (WAP) dimensions.	Raw ILOSTAT input read by the ILO preparation script.
ILO/EMP_DWAP_SEX_AGE_RT_A-20260831T1555.dta	ILOSTAT detailed working-age employment-rate extract by age/sex dimensions.	Raw ILOSTAT input read by the ILO preparation script; filtered to ICLS13th-compatible series in code.

ILOSTAT files were downloaded from the ILO employment topic portal and processed in this package using the ICLS 13th concept: https://ilostat.ilo.org/topics/employment/.

`03.intermediate.zip`: supporting intermediate data

The `03.intermediate.zip` archive contains the prepared survey-list and PIP inputs used alongside the country-survey lineups.

Files or folder	Contents and provenance	Use in this package
03.intermediate/LNlist	Prepared survey metadata and listing files for the 2003, 2013, and 2024 lineups.	Identifies the survey files used in each reference year and supports listing and table outputs.
03.intermediate/PIPinput	Prepared PIP poverty inputs by lineup year and poverty line.	Merged by country code, reference year, and poverty line to bring headcount, population, and mean welfare values into the analysis.

The country-survey archives contain restricted survey-derived data. LIS/LISSY-derived files in 03.intermediate/LIS are produced by the authors by running project code in the LISSY system and retrieving the resulting extracts. The source LIS microdata are not redistributed in this package. The LISSY code can be provided upon request and must be run under the user's own authorized LIS account.

License: Confidential/restricted. Do not redistribute DDH-hosted survey microdata, LIS/LISSY extracts, or other restricted resources.

Access instructions: Request the required resources from DDH dataset 0067178. Extract each archive into the package location shown above, preserving its internal folder structure and filenames. Users must obtain their own permissions for the underlying survey and LIS data and comply with all applicable data-use and confidentiality agreements.

Data URL:
https://datacatalog.worldbank.org/int/search/dataset/0067178/lineup_data_for_a_global_profile_of_work_and_poverty_two_decades_of_evidence

Data Access Request

For restricted inputs, replicators may request access using the guidance below.

1. DDH confidential/requested files: Request the six archives listed in the DDH table above: `DDHdofile.zip`, `02.input.zip`, `2003.zip`, `2013.zip`, `2024.zip`, and `03.intermediate.zip`.
2. LIS/LISSY-derived files: Request the author-run LISSY code and related LIS-derived outputs in 03.intermediate/LIS from the corresponding authors.
3. Include in request: intended replication purpose, institutional affiliation, and confirmation that all source-data licensing and confidentiality terms will be respected.

Suggested request language:

"I am requesting access to the restricted replication inputs for the `A Global Profile of Work and Poverty: Two Decades of Evidence' reproducibility package, including DDH-hosted country-level files (2003, 2013, 2024) and, if needed, the author-run LISSY code used to generate LIS-derived inputs. This request is for non-commercial replication/research use under applicable data-use and confidentiality terms."

Data Sources

The package uses prepared intermediate and output files from the locations below. The file-level provenance for the DDH-provided `02.input.zip` archive is documented in the DDH section above.

Data category	Location in package	Source/provenance	Description and cleaning in this package
Harmonized country survey microdata	03.intermediate/2003, 03.intermediate/2013, 03.intermediate/2024	Harmonized country survey files assembled for this analysis; access depends on the original survey, LIS, or institutional permissions.	Authoritative country-level analysis files by reference year. File names follow `COUNTRY_SURVEYYEAR_SURVEY_LNYYYY_IND.dta`. The Stata prep scripts treat these files as analysis-ready inputs, then create anonymous household/person IDs, drop non-analysis metadata, standardize missing/not-asked labels, derive age groups, urban status, agriculture status, gender and education categories, fill selected variables from `_year` alternatives where

Data category	Location in package	Source/provenance	Description and cleaning in this package
			available, remove negative welfare values, apply PPP floors, and assign percentile groups.
ILOSTAT employment inputs (ICLS13th)	02.input/ILO, 03.intermediate/ILO, 03.intermediate/ILO_ICLS13th_3years.dta	ILOSTAT employment-rate downloads from the ILO employment portal (ICLS 13th concept), mapped to package country codes.	`05-ILO check 15-64.do` imports the raw ILO extracts, keeps ICLS13th-compatible observations, aligns country codes, and saves the harmonized three-year panel used by the GMD-vs-ILO comparison scripts.

The main cleaning and transformation performed by this replication package begins in 03a-Percentile country lineup.do and the country-specific variants. These scripts do not reconstruct the upstream input files; they prepare analysis datasets from the provided inputs and write derived `_100`, `_100cs`, combined global, simulation, figure, and table outputs.

The ten detailed `02.input` files are listed under `02.input.zip` in the DDH section. A `global CPI` macro in the do-files points to `Final_CPI_PPP_to_be_used.dta`, which is not included, but the macro is not dereferenced by the active workflow. The three `GlobalDist1000bins_*` files and `GMD_variables_count.dta` previously listed here are also not present or referenced by the current master workflow.

Data Prep For Intermediate Files (Restricted Inputs Required)

Users cannot rebuild the country-level intermediate files unless they have access to the restricted lineup microdata in DDH. The public/support files in 02.input alone are not sufficient.

- Place authorized country files in 03.intermediate/2003, 03.intermediate/2013, and 03.intermediate/2024.
- Retain LIS-derived files only under the access and redistribution conditions of LIS/LISSY.
- Ensure the prepared PIP poverty files are available in 03.intermediate/PIPinput.
- Set `global upath2` in master.do to the package root before running Stata.

The active master creates the percentile and country-specific intermediate files in this order:

```
// Create country-level 03.intermediate/YYYY_100/*_100.dta files
do "${upath2}\01.code\dofile\03a-Percentile country lineup.do"

// Combine them into 03.intermediate/Global_100_YYYY.dta
do "${upath2}\01.code\dofile\03c-Combine country percentile.do"

// Create country-level 03.intermediate/YYYY_100cs/*_100cs.dta files
do "${upath2}\01.code\dofile\03d-Country specific groups.do"

// Combine them into 03.intermediate/Global_100_YYYY_cs.dta
do "${upath2}\01.code\dofile\03c-Combine country percentile - country specific.do"
```

The two preparation scripts read the restricted country files and prepared PIP inputs, derive harmonized analysis variables, and write one transformed file per source survey. The two combine scripts append those survey files by lineup year.

The authors of the manuscript have legitimate access to and permission to use the data used in this manuscript.

Statement about Rights

- [x] The authors have legitimate access to and permission to use the data used in this analysis.
- [ ] The authors have documented permission to redistribute/publish all restricted source microdata contained in this replication package.

Because some source survey microdata may be restricted, replicators should confirm their own access rights before running the package or redistributing any source data.

Instructions for Replicators

1. Install Stata and the required user-written package listed in Requirements.
2. Confirm that all folders are created (see folder structure) and the country survey files are available in the authoritative country-year folders: 03.intermediate/2003, 03.intermediate/2013, and 03.intermediate/2024.
3. Confirm that the support input files are available in 02.input and that the PIP and survey-list folders are available under 03.intermediate.
4. Open 01.code/dofile/master.do and update the `global upath2` path to the root of this reproducibility package on your machine.
5. Run 01.code/dofile/master.do from Stata.
6. Review generated figures in 04.output/Figures and generated tables in 04.output/Tables.

The current master workflow uses 2021 PPP values and includes poverty-line values defined in the do-files as `300 420 830`.

List of Exhibits

The active commands in master.do produce exhibits for both 3PR Chapter 4 and the background paper. Figure outputs are written to 04.output/Figures, and table outputs are written to 04.output/Tables. The crosswalk below uses the numbering recorded in the master do-file. A dash means that the master does not identify a counterpart in that product.

- ☑ Tables and figures listed below
- ☑ Supporting generated datasets used by listed tables and figures

Detailed figure and table inventory (main reference)

This is the authoritative, row-per-exhibit inventory for the package: it covers every figure and table, rebuilt from the compiled PDF (05.paper/Jobs_and_poverty_paper.pdf) -- the PDF text was extracted and every `Figure N:`/`Figure X.N:`/`Table N:` caption was matched to its source in 05.paper/main.tex and to its producing do-file. The **Report Figure** column gives the corresponding 3PR Chapter 4 report exhibit number as previously recorded in the Report-to-paper crosswalk; those values were not independently re-verified against a 3PR Chapter 4 PDF, so treat them as carried over rather than confirmed. `--` means no 3PR counterpart was recorded. Split below into Main Figures (background-paper Figures 1-15), Main Tables (output workbooks), and Annex (appendix Figures A-G).

Main Figures

Exhibit	Report Figure	Name or panel	Output filename	Producing do-file	Notes
Fig1a	Figure 4.1	a) By quintile	`RFig4_1_growth_quintile.png`	04g-Percentile growth figures mean mu emp sh1564.do	Growth decomposition.
Fig1b	Figure 4.1	b) By percentile	`RFig4_1_growth_p100.png`	04g-Percentile growth figures mean mu emp sh1564.do	Growth decomposition.
Fig2a	Figure 4.5a	a) Employment rate of the poor, 2003 vs. 2024 (ages 15-64)	`RFig1_1_emp_scatter_poor830_b.png`	03f-Scatter country specific groups.do	Country-specific scatter.

Exhibit	Report Figure	Name or panel	Output filename	Producing do-file	Notes
Fig2b	Figure 4.5b	b) Annualized change in employment rate (ages 15-64)	RFig1_1_annemp_scatter_poor830_b.png	03f-Scatter country specific groups.do	Annotated country-specific scatter.
Fig3a	--	a) Employment rate of the poor, 2003 vs. 2024 (ages 25-64)	RFig1_1_emp_scatter_poor830_2564.png	03f-Scatter country specific groups 2564.do	Ages-25-64 companion to Figure 2.
Fig3b	--	b) Annualized change in employment rate (ages 25-64)	RFig1_1_annemp_scatter_poor830_2564.png	03f-Scatter country specific groups 2564.do	Ages-25-64 companion to Figure 2.
Fig4a	Figure 4.2a	a) By poverty status	RFig1_2_lstatus_poorcat_2024.png	03f-Bar chart by poor status.do	2024 bar chart by poverty status.
Fig4b	Figure 4.2b	b) Along the global welfare distribution	RFig1_1_lstatus_dep_2024.png	03d-Percentile figures (within p).do	Dependency-ratio panel.
Fig4c	Figure 4.3 (annex, previously)	c) By poverty status and income group	RFig1_4_lstatus_incgrp_lines.png	03f-4 Labor status by poor and inc group-specific lines.do	Distinct from the single-line _300 file (Figure G.2).
Fig5a	Figure 4B.1.1a (annex, previously)	a) Wage workers, 2024	RFig4_1_hours_ww_scatter_poor300.png	03d-Scatter Country specific groups - hours.do	Hours, $3.00 line.
Fig5b	Figure 4B.1.1b (annex, previously)	b) Self-employed, 2024	RFig4_1_hours_se_scatter_poor300.png	03d-Scatter Country specific groups - hours.do	Hours, $3.00 line.
Fig6a	Figure 4D.2c (annex, previously)	a) Wage vs. self-employed	RFig1_1_mu_wage_self_scatter_all2024.png	03f-Scatter Country specific groups - mu.do	Welfare per worker by type of work.
Fig6b	Figure 4D.1c (annex, previously)	b) By education	RFig1_1_mu_edusec01_scatter_all2024.png	03f-Scatter Country specific groups - mu.do	Welfare per worker by education.

Exhibit	Report Figure	Name or panel	Output filename	Producing do-file	Notes
Fig6c	Figure 4D.3c (annex, previously)	c) By industry	RFig1_1_mu_scatter_all2024.png	03f-Scatter Country specific groups - mu.do	Welfare per worker by sector.
Fig7a	Figure 4D.3a (annex, previously)	a) Sector along the welfare distribution, 2024	RFig1_2_industry_2024.png	03d-Percentile figures (within p).do	Composition by sector.
Fig7b	Figure 4D.3b (annex, previously)	b) Agriculture share, poor vs. non-poor, 2024	RFig1_1_agrsha_scatter_poor830_01.png	03f-Scatter country specific groups - industry.do	Agriculture-share scatter.
Fig8a	Figure 4D.2a (annex, previously)	a) Employment type along the welfare distribution, 2024	RFig1_3_empstat_2024.png	03d-Percentile figures (within p).do	Composition by employment type.
Fig8b	Figure 4D.2b (annex, previously)	b) Self-employment share, poor vs. non-poor, 2024	RFig1_1_selfsha_scatter_poor830_01.png	03f-Scatter country specific groups - emp status.do	Self-employment-share scatter.
Fig9	--	Employment status and sector jointly, 2024	RFig1_3_secemp_2024.png	03d-Percentile figures (within p).do	Employment by sector and status.
Fig10a	Figure 4.7a	a) $3.00 line	RFig1_2_agseedu_poor300_2024.png	03f-3 Bar chart by poor status 300.do	Education/agriculture bar chart.
Fig10b	Figure 4.7b	b) $8.30 line	RFig1_2_agseedu_poor830_2024.png	03f-3 Bar chart by poor status 800.do	Education/agriculture bar chart.
Fig11	Figure 4.4	Gender working (U-shape), 2003--2024	RFig3_1_gender_working_p_3years.png	03d-Percentile figures (within p) male female 3 years.do	Single-panel figure.
Fig12a	Figure 4D.4a (annex, previously)	a) Wage employment	RFig3_2_gender_ww_p.png	03d-Percentile figures (within p) male female.do	Employment type and sector of employment by gender.

Exhibit	Report Figure	Name or panel	Output filename	Producing do-file	Notes
Fig12b	Figure 4D.4b (annex, previously)	b) Self-employment and other arrangements	RFig3_2_gender_se_p.png	03d-Percentile figures (within p) male female.do	Employment type and sector of employment by gender.
Fig12c	Figure 4D.5a (annex, previously)	c) Agriculture	RFig3_1_gender_agr_within_p.png	03d-Percentile figures (within p) male female.do	Employment type and sector of employment by gender.
Fig12d	Figure 4D.5b (annex, previously)	d) Non-agriculture	RFig3_1_gender_nagr_within_p.png	03d-Percentile figures (within p) male female.do	Employment type and sector of employment by gender.
Fig13a	Figure 4.6 (panel not specified)	a) $3.00 line	RFig2_1_simemp75_global3_300.png	04f-Scatter country sim potential workers 75.do	Household-level employment simulation.
Fig13b	Figure 4.6 (panel not specified)	b) $8.30 line	RFig2_1_simemp75_global3_830.png	04f-Scatter country sim potential workers 75.do	Household-level employment simulation.
Fig13c	Figure 4.6 (panel not specified)	c) $3.00 line; female potential workers	RFig2_1_simemp75f_global3_300.png	04f-Scatter country sim female potential workers 75.do	Gender-specific simulation.
Fig13d	Figure 4.6 (panel not specified)	d) $8.30 line; female potential workers	RFig2_1_simemp75f_global3_830.png	04f-Scatter country sim female potential workers 75.do	Gender-specific simulation.
Fig13e	Figure 4.6 (panel not specified)	e) $3.00 line; male potential workers	RFig2_1_simemp75m_global3_300.png	04f-Scatter country sim male potential workers 75.do	Gender-specific simulation.
Fig13f	Figure 4.6 (panel not specified)	f) $8.30 line; male potential workers	RFig2_1_simemp75m_global3_830.png	04f-Scatter country sim male potential workers 75.do	Gender-specific simulation.
Fig14a	Figure 4.8a	a) Raising education	RFig2_1_simsecedu_global3_300.png	04d-figures educational changes and pov distribution.do	$3.00-line earnings switch.
Fig14b	Figure 4.8b	b) Self-employment to wage	RFig2_1_simseww_global3_300.png	04d-figures emp status gains and pov distribution.do	$3.00-line earnings switch.

Exhibit	Report Figure	Name or panel	Output filename	Producing do-file	Notes
Fig14c	Figure 4.8c	c) Agriculture to non-agriculture	RFig2_1_simsec_global3_300.png	04d-figures sector gains and pov distribution.do	$3.00-line earnings switch.
Fig15a	Figure 4.10a	a) $3.00 line	RFig4_1_4sims_300_GIC.png	04d-figures Compare 3 simulations.do	Growth-incidence curves, four counterfactuals.
Fig15b	Figure 4.10b	b) $8.30 line	RFig4_1_4sims_830_GIC.png	04d-figures Compare 3 simulations.do	Growth-incidence curves, four counterfactuals.

Main Tables

Exhibit	Report Figure	Name or panel	Output filename	Producing do-file	Notes
Table1	Table A1	Survey and country listing	Tables/TableA1.xlsx	04h-Table - listing.do	Survey and country listing table.
Table2	Table A2	Survey and country listing	Tables/TableA2.xlsx	04h-Table - listing.do	Survey and country listing table.
Table3	Table A3	Annex workbook	Tables/TableA3.xlsx	04h - table A3.do	Annex Table A3 workbook output.
TableA.1/A.2	--	GMD vs. ILO comparison	Tables/TableX1.xlsx	03f-GMD ILO compare.do	Appendix Table A.1 and Appendix Table A.2 are sourced from this TableX1.
TableH.1	--	Supporting combined workbook	Tables/Tables_combined.xlsx	04h-Table - listing.do	Combined supporting workbook generated by the table listing script.

Annex

Exhibit	Report Figure	Name or panel	Output filename	Producing do-file	Notes
FigA.1a	--	a) 2003-2013, poor employment	RFig1_1_emp_scatter_poor830_0313.png	03f-Scatter country specific groups 2003 2013 2024.do	Sub-period country-level changes.
FigA.1b	--	b) 2013-2024, poor employment	RFig1_1_emp_scatter_poor830_1324.png	03f-Scatter country specific groups 2003 2013 2024.do	Sub-period country-level changes.
FigA.1c	--	c) 2003-2013, annualized changes	RFig1_1_annemp_scatter_poor830_0313.png	03f-Scatter country specific groups 2003 2013 2024.do	Sub-period country-level changes.

Exhibit	Report Figure	Name or panel	Output filename	Producing do-file	Notes
FigA.1d	--	d) 2013-2024, annualized changes	RFig1_1_annemp_scatter_poor830_1324.png	03f-Scatter country specific groups 2003 2013 2024.do	Sub-period country-level changes.
FigC.1a	Figure 4C.1a	a) $3.00 line	RFig2_1_simemp75_scatter_poor300.png	04f-Scatter country sim potential workers 75.do	Counterfactual vs. observed headcount.
FigC.1b	Figure 4C.1b	b) $8.30 line	RFig2_1_simemp75_scatter_poor830.png	04f-Scatter country sim potential workers 75.do	Counterfactual vs. observed headcount.
FigC.2a	Figure 4C.2a	a) $3.00 line	RFig2_1_simemp75_scatter_pgap300.png	04f-Scatter country sim potential workers 75.do	Counterfactual vs. observed poverty gap.
FigC.2b	Figure 4C.2b	b) $8.30 line	RFig2_1_simemp75_scatter_pgap830.png	04f-Scatter country sim potential workers 75.do	Counterfactual vs. observed poverty gap.
FigC.3a	Figure 4C.3a	a) $3.00 line	RFig2_1_simemp75_scatter_ppgap300.png	04f-Scatter country sim potential workers 75.do	Counterfactual vs. observed percent poverty gap.
FigC.3b	Figure 4C.3b	b) $8.30 line	RFig2_1_simemp75_scatter_ppgap830.png	04f-Scatter country sim potential workers 75.do	Counterfactual vs. observed percent poverty gap.
FigC.4	Figure 4C.4	Global distribution of welfare per capita, 2024	RFig4_2_kdensity75.png	04f-Scatter country sim potential workers 75.do	Kernel density figure.
FigD.1a	Figure 4D.1a	a) Education levels along the welfare distribution, 2024	RFig1_1_educat4_2024.png	03d-Percentile figures (within p) educat4.do	Education annex figure.
FigD.1b	Figure 4D.1b	b) Secondary education or more, poor vs. non-poor, 2024	RFig1_1_edusec_scatter_poor830_01.png	03f-Scatter country specific groups - educat4.do	Education annex figure.

Exhibit	Report Figure	Name or panel	Output filename	Producing do-file	Notes
FigD.1c	Figure 4D.1c	c) Welfare per worker by education, 2024 (log scales)	RFig1_1_mu_edusec01_scatter_all2024.png	03f-Scatter Country specific groups - mu.do	Same file as Figure 6b.
FigD.2a	Figure 4D.2a	a) Employment type along the welfare distribution, 2024	RFig1_3_empstat_2024.png	03d-Percentile figures (within p).do	Same file as Figure 8a.
FigD.2b	Figure 4D.2b	b) Self-employment share, poor vs. non-poor, 2024	RFig1_1_selfsha_scatter_poor830_01.png	03f-Scatter country specific groups - emp status.do	Same file as Figure 8b.
FigD.2c	Figure 4D.2c	c) Welfare per worker by type of employment, 2024 (log scales)	RFig1_1_mu_wage_self_scatter_all2024.png	03f-Scatter Country specific groups - mu.do	Same file as Figure 6a.
FigD.3a	Figure 4D.3a	a) Sector along the welfare distribution, 2024	RFig1_2_industry_2024.png	03d-Percentile figures (within p).do	Same file as Figure 7a.
FigD.3b	Figure 4D.3b	b) Agriculture share, poor vs. non-poor, 2024	RFig1_1_agrsha_scatter_poor830_01.png	03f-Scatter country specific groups - industry.do	Same file as Figure 7b.
FigD.3c	Figure 4D.3c	c) Welfare per worker by sector, 2024 (log scales)	RFig1_1_mu_scatter_all2024.png	03f-Scatter Country specific groups - mu.do	Same file as Figure 6c.
FigD.4	--	Employment by three sectors	RFig3_1_gender_within_p.png	03d-Percentile figures (within p) male female.do	Employment by three sectors
FigE.1	--	Construction of the conditional premium and the participation penalty	fig_annex_schematic.png	*Not Stata-generated*	Authors' manual illustration, per caption.

Exhibit	Report Figure	Name or panel	Output filename	Producing do-file	Notes
FigE.2	Figure 4D.7	Conditional participation penalty and predicted earnings gain from leaving agriculture, 2024	RFig2_1_secgains_scatter_all830.png	04d-figures sector gains and pov distribution.do	Country-level scatter.
FigF.1a	Figure 4E.1a	a) Multiplier	RFig4_1_maxworker75_multiplier.png	04f-Scatter country sim potential workers 75.do	Employment-GIC decomposition factor.
FigF.1b	Figure 4E.1b	b) Treated share	RFig4_1_maxworker75_treated.png	04f-Scatter country sim potential workers 75.do	Employment-GIC decomposition factor.
FigF.2a	Figure 4.9a (previously recorded as main)	a) Raising education	RFig2_1_simsecedu_global3_830.png	04d-figures educational changes and pov distribution.do	$8.30-line earnings switch (NOT a main-text figure, despite being previously recorded as Fig9a).
FigF.2b	Figure 4.9b (previously recorded as main)	b) Self-employment to wage	RFig2_1_simseww_global3_830.png	04d-figures emp status gains and pov distribution.do	$8.30-line earnings switch (NOT a main-text figure, despite being previously recorded as Fig9b).
FigF.2c	Figure 4.9c (previously recorded as main)	c) Agriculture to non-agriculture	RFig2_1_simsec_global3_830.png	04d-figures sector gains and pov distribution.do	$8.30-line earnings switch (NOT a main-text figure, despite being previously recorded as Fig9c).
FigG.1a	--	a) By region, 2024	RFig1_1_lstatus_2024_regions.png	03d-Percentile figures (within p) - inc reg groups.do	Employment by regions
FigG.1b	--	b) By country	RFig1_1_lstatus_2024_incgroup.png	03d-Percentile figures (within	Employment by income

Exhibit	Report Figure	Name or panel	Output filename	Producing do-file	Notes
		income group, 2024		p) - inc reg groups.do	groups
FigG.2	Figure 4.3 (previously recorded as main)	Employment composition of poor/non-poor by income group ($3.00 line)	RFig1_4_lstatus_incgrp_300.png	03f-4 Labor status by poor and inc group.do	Single-line variant; NOT the same file as Figure 4c (which uses _lines).
FigG.3	--	Global trends in employment ratio and sectoral shares, 2003--2024	RFig2_1_overtime_within_p.png	03e-Percentile figures (within p) over time.do	Employment over years.
FigG.4	--	Labor force status including missing observations, 2024	RFig1_1_lstatus_2024_annex.png	03d-Percentile figures (within p).do	Employment category with missing.
FigG.5a	--	a) Employment rate of the poor, 2003 vs. 2024	RFig1_1_emp_scatter_poor300_b.png	03f-Scatter country specific groups.do	$3.00-line country-level changes.
FigG.5b	--	b) Annualized change in employment rate	RFig1_1_annemp_scatter_poor300_b.png	03f-Scatter country specific groups.do	$3.00-line country-level changes.
FigG.6a	Figure 4F.2a (previously)	a) Ages 15-64	RFig1_1_annemp_scatter_poor830_inconly.png	03f-Scatter country specific groups.do	Income-only concept, by income group.
FigG.6b	Figure 4F.2b (previously)	b) Ages 25-64	RFig1_1_annemp_scatter_poor830_inconly2564.png	03f-Scatter country specific groups 2564.do	Income-only concept, by income group.
FigG.7a	--	a) Ages 15-24	RFig2_1_overtime_work1524_p.png	03e-Percentile figures (within p) over time P1524.do	Employment rate 15-24
FigG.7b	--	b) Ages 25-64	RFig2_1_overtime_work2564_p.png	03e-Percentile figures (within p) over time P2564.do	Employment rate 15-24

Exhibit	Report Figure	Name or panel	Output filename	Producing do-file	Notes
FigG.8a	--	a) Wage workers, 2024	`RFig4_1_hours_ww_scatter_poor830.png`	03d-Scatter Country specific groups - hours.do	Hours, $8.30 line.
FigG.8b	--	b) Self-employed, 2024	`RFig4_1_hours_se_scatter_poor830.png`	03d-Scatter Country specific groups - hours.do	Hours, $8.30 line.
FigG.9a	--	a) All workers, by gender	`RFig4_1_hours_all_scatter_m_f.png`	03d-Scatter Country specific groups - hours.do	Weekly hours by gender.
FigG.9b	--	b) Female workers, poor vs. non-poor	`RFig4_1_hours_female_scatter_poor300.png`	03d-Scatter Country specific groups - hours.do	Weekly hours, women poor vs. non-poor.

Requirements

Software

- Stata 17 or later recommended. The workflow uses Stata data management, `collect`, graph export, and Excel export commands.

Stata User-Written and External Dependencies

- Stata package `grstyle` is used by 01.code/dofile/00-color setting.do to set the graph style and color palette.
- External Stata command `groupfunction` is used by the data-preparation, figure, simulation, and table scripts, including 03a-Percentile country lineup.do, 03d-Country specific groups.do, 04f-sim potential workers on h pg ppg 75%.do, and 04h - table A3.do.
- The ado folder is included in this working-paper folder (_ebin.ado).
- The prepared input and intermediate files described in Data Availability must be present before running the full workflow.

If needed, install the Stata dependencies before running the package:

```
ssc install grstyle, replace
ssc install groupfunction, replace
ssc install grc1leg2, replace
```

Runtime

- Full Stata pipeline: 180 minutes. Runtime depends heavily on the machine, storage speed, and whether restricted country-level intermediate files are already prepared.
- Exhibit-only reruns from existing intermediate outputs are expected to be faster than rebuilding percentile, country-specific, and simulation datasets from the country files.

Storage

- The current size of this working-paper reproducibility package is approximately 39.9 GB across inputs, intermediate files, outputs, and code.

- Recommended disk space: at least 50 GB for this package, with additional space if logs, temporary files, backups, or alternative output versions are retained.

Computational Requirements

The code uses Windows-style paths and a hardcoded root directory in 01.code/dofile/master.do. Replicators on other operating systems may need to adjust path separators and directory globals before running the workflow.

The package reads and writes Stata `.dta` files, PNG graphics, and Excel workbooks. Ensure enough local storage is available for the source country files, generated `_100` and `_100cs` files, combined intermediate datasets, simulation datasets, figures, and table outputs.

Code Description

The master do-file controls the workflow in the following order.

Step	Do-file(s)	Purpose
Setup	master.do, 00-color setting.do	Set root path, input paths, PPP base year, graph style, and global color macros.
Percentile preparation	03a-Percentile country lineup.do, 03c-Combine country percentile.do	Create and combine country percentile files, including generated `_100` files.
Country-specific groups	03d-Country specific groups.do, 03c-Combine country percentile - country specific.do	Create and combine country-specific group files, including generated `_100cs` files.
Hours preparation	03d-Country specific groups - hours.do	Create country-specific hours datasets used by hours figures.
Simulation preparation	04f-sim potential workers on h pg ppg 75%.do, female variant, male variant, 04f-sim education changes on h pg ppg.do, 04d-emp status gains and pov distribution.do, 04d-sector gains and pov distribution.do, 04f-sim 5sectors on h pg ppg pooled reg.do	Prepare simulation inputs for total, female, and male potential workers and for education, employment-status, sector, and five-sector scenarios.
Main figures	Figure do-files listed in Detailed figure and table inventory	Generate paper figures, write PNG files to 04.output/Figures, and save figure source-data outputs to 04.output/DataFig.
Annex figures	Annex figure do-files listed in Detailed figure and table inventory	Generate appendix/annex figures, write PNG files to 04.output/Figures, and save supporting figure source-data outputs to 04.output/DataFig.
GMD vs ILO comparison (optional): ILO panel	05-ILO check 15-64.do	Imports ILOSTAT extracts from 02.input/ILO, keeps ICLS13th-compatible observations, and harmonizes country codes, writing 03.intermediate/ILO/ILO_ICLS13th_3years.dta. These commands are currently commented out in master.do.
GMD vs ILO comparison (optional): GMD panels	03f-Emp ratio - nat - 3 lnyears.do, 03f-Emp ratio - nat - 3 lnyears poor830.do	Build GMD employment-rate panels across 2003, 2013, and 2024 for all population and for the poor-830 subgroup, writing 03.intermediate/GMD_empratio_3years.dta and 03.intermediate/GMD_empratio_3years_poor830.dta. These commands are currently commented out in master.do.
GMD vs ILO comparison (optional): comparison	03f-GMD ILO compare.do	Merges the GMD and ILO panels above and produces diagnostic comparisons of employment levels and changes, writing `Tables/TableX1.xlsx`. These commands are currently commented out in master.do.
Tables	04h-Table - listing.do, 04h - table A3.do	Export Table A1, Table A2, Table A3, and a combined supporting workbook to `04.output/Tables`.

Folder Structure

```
01.code/
  dofile/
    master.do
    00-color setting.do
    03a-Percentile country lineup.do
    03c-Combine country percentile.do
    03c-Combine country percentile - country specific.do
    03d-*.do
    03f-*.do
    04d-*.do
    04f-*.do
    04g-*.do
    04h-*.do

02.input/
  CHN_IND_list.dta
  CLASS.dta
  code_pop_un_cohorts_sh.dta
  country_list_name.dta
  country_list2.dta
  growth_gdp_pce_pop.dta
  ILO/
    clist_with_wb_code.csv
    EMP_5WAP_SEX_AGE_RT_A-20260831T1556.dta
    EMP_DWAP_SEX_AGE_RT_A-20260831T1555.dta
  List_5sectors.xlsx
  Survey_price_framework.dta
  wdi_1564_pop.dta
  WDI_code_pop_un1564.dta

03.intermediate/
  2003/          Authoritative country survey files for 2003 reference year
  2013/          Authoritative country survey files for 2013 reference year
  2024/          Authoritative country survey files for 2024 reference year
  2003_100/      Generated percentile files
  2003_100cs/    Generated country-specific percentile files
  2024_100/      Generated percentile files
  2024_100cs/    Generated country-specific percentile files
  ILO/
    ILO_ICLS13th_3years.dta
  LIS/
  LNlist/
  PIPinput/
  Global_100_*.dta
  GMD_empratio_3years*.dta
  ILO_ICLS13th_3years.dta
  results_*switch*.dta
  Sim_maxworker75_2024*.dta

04.output/
  DataFig/
    DataFig*.dta     Figure source-data outputs produced here and copied to
../07.reproducibility_3pr/1-data/processed
    DataTab*.dta     Table source-data outputs produced here and copied to
../07.reproducibility_3pr/1-data/processed
  Figures/
    RFig*.png
  Tables/
    TableA1.xlsx
    TableA2.xlsx
    TableA3.xlsx
    Tables_combined.xlsx
```

```
    CountrySpecific_*.dta
    Percentile_*.dta
```

Notes and Caveats

- 01.code/dofile/master.do contains a hardcoded `global upath2`; this must be updated before replication on another machine.
- The exhibit-number crosswalk follows comments in `master.do`. The report and paper manuscripts are not stored together in this repository, so rows marked "not specified" should be updated in the master when final numbering is assigned.
- The current master contains two figure-description lines beginning `Fig7a` and `Fig7b` without Stata comment markers. Prefix these lines with `*` before an unattended run; otherwise Stata will interpret them as commands.
- Country survey microdata may be restricted and should not be redistributed unless the relevant permissions allow it.
- The folders 03.intermediate/2003, 03.intermediate/2013, and 03.intermediate/2024 are the authoritative country-file folders. Several data-preparation scripts currently set `global lnyearlist 2003 2024`, while the hours scatter script includes `2003 2013 2024`.
- The code uses 2021 PPP values and poverty-line values defined in the do-files.
- 04.output/Figures contains additional generated PNG variants beyond the crosswalk above; these are supporting or alternative views produced by the figure scripts. The five-sector figure command is currently disabled in the master even though its preparation script remains active.
- The GMD-vs-ILO comparison block in master.do is present but currently commented out. To reproduce ILO comparison checks, uncomment the four ILO/GMD comparison calls after confirming 02.input/ILO is populated.
- Runtime and memory requirements were not recorded in the code. Replicators should expect runtime to depend on Stata version, storage speed, and the number and size of country survey files available locally.
- The 3PR Chapter 4 package in ../07.reproducibility_3pr depends on this package's 04.output/DataFig outputs. If Chapter 4 source data need to be updated, regenerate `DataFig*.dta` and `DataTab*.dta` here first, then copy them into the Chapter 4 package's `1-data/processed` folder.