Replication Data for:
"Preparing for an Aging Africa:
Data-Driven Priorities for Economic Research and Policy"

by Madeline Duhon, Edward Miguel, Amos Njuguna,
Daniela Pinto Veizaga, and Michael Walker

Replication materials overview

This folder contains raw data, intermediate data, and code to replicate the tables, figures, and analysis included in the paper and the online appendix. All data cleaning and some of the analysis is done using Stata. Most of the figures produced are using R.

Features

- All raw data, aside from the country-level survey data, is included in the replication folder.
- Instructions on where to obtain the raw country-level survey data are included in this document.
- Code for:
    - Cleaning and analysis.
    - Generating the paper and online appendix figures.

Directory structure

- **"data"** is the subfolder for all cleaned data files.
    - Contains all cleaned data necessary to replicate the analysis, including the cleaned country-level survey data (`allcountries.dta`).
- **"rawdata"** is the repository for raw data.
    - Contains all raw data, aside from raw country-level survey data.
- **"code"** subfolder contains all cleaning and analysis code.
    - `master.do` runs the cleaning and the analysis code (for all analyses done in Stata).
    - `code/cleaning` contains code for cleaning all raw data, aside from the country-level survey data (as Stata .do files).
    - `code/cleaning/countries` contains code for cleaning country-level survey data (as Stata .do files); this code can be run only after downloading the raw survey data files from original sources as described below.

- `code/analysis` contains code for analysis (for all analyses done in Stata; as Stata .do files)
- `master.R` runs the figures code (for all figures done in R)
- `renv.lock` the lockfile for R package versions.
- `code/figures` contains code for replicating the figures
  - `./constants.R` contains custom constants used in the R scripts.
  - `./funcs.R` contains custom functions for data and figure processing.
  - `./clean_files.R` script to clean data for figures.
  - `./main_graph.R` script to generate the main figures.
  - `./appendix.R` script to generate the appendix figures.

- **"output"** is the repository for all tables and figures.
- ``README.pdf'' is this README file.

Cleaning and analysis

- Software: Stata (version 18).
- `master.do`, contains a description of the purpose, inputs, and outputs of all do files used for cleaning and analysis.
- Set working directory in `master.do` (row 13) to run do files for cleaning and analysis.
- As necessary, comment out or run commands at the start of `master.do` (rows 19-23) to install user-generated commands.
- Note that code calling on do files used to clean survey data and produce `allcountries.dta` is commented out, since these cannot be run without the raw country-level survey data.
- User-generated commands and packages: `shp2dta`, `egenmore`, `_gwtmean`, `spmap`, `texsave`, `unique`

Generating the paper and online appendix figures

- Prerequisites
  - Software:
    - R version 4.2.0 or higher installed.
    - Operating System: Compatible with Windows, macOS, and Linux.
  - Required R Packages:
    - This project uses the `renv` package to manage dependencies. The `renv.lock` file ensures that the correct package versions are installed for consistent results.

- Setup Instructions
  - Download the repository to your local machine.
  - Install `renv` and restore environment.
    - Install the renv package in R (if not already installed):

      ```
      install.packages("renv")
      ```

    - Navigate to the directory where the `renv.lock` is located and restore the project environment:

      ```
      renv::restore()
      ```

      This will install all the required packages at the versions specified in `renv.lock`.
- Running the code
  - Set up the working directory.
    - The `main_script.R` automatically sets the working directory to the folder containing the script using the `rstudioapi` package. If you encounter an issue, manually set the working directory in R:

      ```
      setwd("path/to/your-repo-name")
      ```

  - Run the master script. This will load the data, preprocess the datasets for figures, and generate the main and appendix figures. Figures will be saved to the `output/`.

    ```
    source("master.R")
    ```

- The paper and apppedix figures will be saved in:

  ```
  output/*
  ```

- Troubleshooting
  - `rstudioapi` not available: If you are not using RStudio, the script will prompt you to set the working directory manually. Update the `setwd()` function in `master.R`.
  - Missing data files: Ensure all required data files are in the `data/` directory.
  - `renv::restore()` issues: If `renv::restore()` fails, try manually installing missing packages, eg.,

```
install.packages(c("ggplot2", "dplyr", "patchwork"))
```

- Custom scripts not found: Ensure the `figures/` directory contains all the required scripts (`constants.R, funcs.R`, etc.). If any are missing, check the dataverse repository for updates.

Where to obtain data

- Raw country survey data can be obtained as listed in Table 1.
- All other raw data used in the analysis (and included in the replication folder) or discussed in the text can be obtained as listed in Table 2.

Table 1: Country-level survey data

Database and Survey	Years	Date Accessed	Link or DOI
World Bank Microdata Library: Living Standards Measurement Study			
Nigeria General Household Survey	2018-2019	March 2023	https://doi.org/10.48529/8gbe-z155
	2015-2016	March 2023	https://doi.org/10.48529/7xmj-q133
	2012-2013	March 2023	https://doi.org/10.48529/kxpy-aa72
	2010-2011	March 2023	https://doi.org/10.48529/y9e2-b753
Ethiopia Socioeconomic Survey	2018-2019	March 2023	https://doi.org/10.48529/k739-c548
	2015-2016	March 2023	https://doi.org/10.48529/ampf-7988
	2013-2014	March 2023	https://doi.org/10.48529/mccp-v123
Malawi Integrated Household Survey	2019-2020	March 2023	https://doi.org/10.48529/yqn3-zv74
	2016-2017	March 2023	https://doi.org/10.48529/g2p9-9r19
	2010-2011	March 2023	https://doi.org/10.48529/w1jq-qh85
	2004-2005	March 2023	https://doi.org/10.48529/2ked-2t88
Tanzania National Panel Survey	2020-2021	March 2023	https://microdata.worldbank.org/index.php/catalog/5639
	2014-2015	March 2023	https://doi.org/10.48529/y3qj-d018
	2012-2013	March 2023	
	2010-2011	March 2023	
	2008-2009	March 2023	https://doi.org/10.48529/hz8s-3489
Uganda National Panel Survey	2019-2020	March 2023	
	2018-2019	March 2023	https://doi.org/10.48529/ttvg-m774
	2015-2016	March 2023	
	2013-2014	March 2023	https://doi.org/10.48529/c1c4-h654
	2011-2012	March 2023	https://doi.org/10.48529/5cpp-r373
	2010-2011	March 2023	
	2005-2009	March 2023	
Niger Enquête Harmonisée sur le Conditions de Vie des Ménages	2018-2019	August 2023	
/ National Survey on Household Living Conditions and Agriculture	2014	August 2023	https://doi.org/10.48529/3xnb-sd96
	2011	August 2023	https://doi.org/10.48529/bp16-s524
Harvard Dataverse			
Ghana Socioeconomic Panel Survey	2009-2018	March 2023	https://doi.org/10.7910/DVN/E5QP0F
DataFirst			
South Africa National Income Dynamics Study	2017	March 2023	https://doi.org/10.25828/fw3h-v708
	2014-2015	March 2023	https://doi.org/10.25828/f4ws-8a78
	2012	March 2023	https://doi.org/10.25828/7pgq-q106
	2010-2011	March 2023	https://doi.org/10.25828/j1h1-5m16
	2008	March 2023	https://doi.org/10.25828/e7w9-m033

Table 2: Other data used in analysis or discussed in text

Database and Variables	Date accessed	Link
ArcGIS Hub		
Africa country boundaries	March 2023	https://hub.arcgis.com/datasets/07610d73964e4d39ab62c4245d548625_0/explore?location=35.401272%2C34.619903%2C22.77
Health and Retirement Survey		
In text: Patterns of employment, marriage, and living alone among older individuals in the US	June 2024	https://hrsdata.isr.umich.edu/data-products/2020-hrs-core
Institute for Health Metrics and Evaluation: Global Burden of Disease Results		
In text: Statistics related to years lived with disability, leading causes of years lived with disability, deaths due to communicable diseases, and rates of depression	July 2024	https://vizhub.healthdata.org/gbd-results/
International Labor Organization ILOSTAT Explorer		
In text: Employment among older individuals	September 2023	https://ilostat.ilo.org/data/data-explorer/
International Labor Organization Social Protection Data Dashboard		
Pension coverage	May 2023	https://www.social-protection.org/gimi/WSPDB.action?id=19
Social protection expenditure on older persons (as % of GDP)	May 2023	https://www.social-protection.org/gimi/WSPDB.action?id=19
Sustainable Development Goals Indicators Database		
In text: Employment in the informal sector in SSA	September 2023	https://unstats.un.org/sdgs/dataportal/database
United Nations World Population Prospects 2022		
Population projections	March 2023	https://population.un.org/wpp/Download/Standard/Population/
Regional classifications	March 2023	https://population.un.org/wpp/Download/Documentation/Documentation/
World Bank		
Income groups	March 2023	https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups
World Bank Development Indicators		
Population	March 2023	https://data.worldbank.org/indicator/SP.POP.TOTL
Population, female	May 2023	https://data.worldbank.org/indicator/SP.POP.TOTL.FE.IN
Population, male	May 2023	https://data.worldbank.org/indicator/SP.POP.TOTL.MA.IN
GDP per capita, PPP	June 2023	https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD
Domestic government health expenditure (as % of GDP)	March 2023	https://data.worldbank.org/indicator/SH.XPD.GHED.GD.ZS
Private health expenditure (as % of total health expenditure)	March 2023	https://data.worldbank.org/indicator/SH.XPD.OOPC.CH.ZS
Physicians per 1000 (year)	March 2023	https://data.worldbank.org/indicator/SH.MED.PHYS.ZS
Hospital beds (per 1000)	March 2023	https://data.worldbank.org/indicator/SH.MED.BEDS.ZS
World Health Organization Global Health Observatory		
Life expectancy at birth	May 2023	https://www.who.int/data/gho/data/indicators/indicator-details/GHO/life-expectancy-at-birth-(years)
Life expectancy at age 60	May 2023	https://www.who.int/data/gho/data/indicators/indicator-details/GHO/life-expectancy-at-age-60-(years)
SDG UHC Service Coverage Index	March 2023	https://www.who.int/data/gho/data/themes/topics/service-coverage