Never Too Young to Dream Big:
Measuring youth aspirations in the time of the COVID-19 pandemic

Valentina Costa[1]   Ivette Contreras[2]   Amparo Palacios-Lopez[3]

[1]Corresponding Author. The World Bank. Email: vcosta@worldbank.org.

[2] The World Bank. Email: icontreras@worldbank.org.

[3] The World Bank. Email: apalacioslopez@worldbank.org.

Reproducibility Package README

Overview

The code within this replication package executes the analysis detailed in the paper "Never Too Young to Dream Big: Measuring youth aspirations in the time of the COVID-19 pandemic" using Stata. This package includes all codes required to go from the anonymized data to the replication of the tables and figures included in the analysis. The master file manages the execution of all codes, anticipated to run for approximately 10 minutes. Our primary objective in this reproducibility package is to ensure both accessibility and credibility in replicating our research methodology and results.

Data Availability and Provenance Statements

All datasets used are publicly available but cannot be redistributed and are not included in the reproducibility package. Replicators can download the datasets in the links provided below.

Information about youth's characteristics and aspirations, community characteristics and remittances are leveraged from High Frequency Phone Surveys:

1. World Bank. Ethiopia - High Frequency Phone Survey 2020-2023. Ref: ETH_2020-2023_HFPS_v13_M. Dataset downloaded from https://microdata.worldbank.org/index.php/catalog/3716 on January 2, 2024

2. Malawi National Statistical Office (NSO) (Government of Malawi). Malawi - High-Frequency Phone Survey 2020-2024 (HFPS-COVID-19 2020-2024). Ref: MWI_2020-2024_HFPS_v18_M. Downloaded from https://microdata.worldbank.org/index.php/catalog/3766 on January 2, 2024

3. National Bureau of Statistics. Nigeria COVID-19 National Longitudinal Phone Survey (COVID-19 NLPS) 2020-2021, Phase 1. Dataset downloaded from https://microdata.worldbank.org/index.php/catalog/3712 on January 2, 2024

Family characteristics are leveraged from the following LSMS-ISA surveys:

1. Central Statistical Agency of Ethiopia. Ethiopia Socioeconomic Survey (ESS4) 2018-2019. Public Use Dataset. Ref: ETH_2018_ESS_v03. Downloaded from https://microdata.worldbank.org/index.php/catalog/3823 on January 2, 2024

2. National Statistical Office. Integrated Household Panel Survey 2010-2013-2016-2019 (Long-Term Panel, 102 EAs). Public Use Dataset. Ref: MWI_2010-2019_IHPS_v06_M. Downloaded from https://microdata.worldbank.org/index.php/catalog/3819 on January 2, 2024

3. National Statistical Office. Malawi Integrated Household Panel Survey 2010-2013 (Short-Term Panel, 204 EAs), Ref. MWI_2010-2013_IHPS_v01_M. Dataset downloaded from https://microdata.worldbank.org/index.php/catalog/2248 on January 2, 2024

4. Nigeria National Bureau of Statistics. General Household Survey, Panel (GHS-Panel) 2018-2019. Dataset downloaded from https://microdata.worldbank.org/index.php/catalog/3557 on January 2, 2024

Statement about Rights

☒ I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

Summary of Availability

All data **are** publicly available.

Getting Started

Outline and Folder Structure

The reproducibility package consists of three folders:

- *Data anonymized:* this folder contains all the raw data files, the intermediate files, and the final data set used for the analysis. None of the data are provided in the reproducibility package due to data redistribution restrictions.

- *Do files:* this folder contains the related codes. The Master File labeled as "**master-dofile.do**" orchestrates the sequential execution of all code files according to the specified order.

- *Output:* this folder contains all the tables and figures reported in the paper in raw form. Output files are called appropriate names (e.g., Table1.xls, Figure1.png) and should be easy to correlate with the manuscript.

How to Reproduce

The initial action within this reproducibility package involves accessing the master file named master-dofile.do and execute its contents. Before proceeding, ensure Stata is installed on your system. This file establishes directory paths and executes the necessary Stata code to generate analytical outputs. Here's a concise overview of its contents:

PART 1: Harmonize settings.

PART 2: Prepare folder paths and define programs. This part runs **0_Globals.do.**

PART 3: Run do-files.

PART 3.1: This part runs **1_vrbs_Ethiopia.do** do-file that merges Ethiopia - High Frequency Phone Survey 2020-2023 with Ethiopia ESS4 2018-19 and constructs variables and indicators related to family characteristics, personal characteristics of the youth, community characteristics, youths' aspirations, remittances, and consumption aggregates. This do-file generates the dataset **ethiopia.dta** with the new variables and indicators useful for the analysis.

This process is replicated for Malawi and Nigeria. For Malawi, this part runs **2_vrbs_Malawi.do** do-file that merges Malawi - High-Frequency Phone Survey 2020-2024 with Malawi - Fifth Integrated Household Survey 2019-2020 and Malawi - Integrated Household Panel Survey 2010-2013 and constructs variables and indicators that are saved in the new dataset **malawi.dta**.

For Nigeria, this part runs **3_vrbs_Nigeria.do** do-file that merges COVID-19 NLPS 2020-2021 with GHS-Panel 2018-2019 and constructs variables and indicators that are saved in the new dataset **nigeria.dta**.

PART 3.2: This part runs **4_Merging datasets.do** do-file that creates a pooled dataset **pooled.dta** by merging the three country datasets ethiopia.dta, malawi.dta and nigeria.dta.

PART 3.3: This part runs **5_Figures.do** do-file that will generate figures 1, 2, 3, 4, 5, 6, 7 and 8 used in the main text of the paper.

PART 3.4: This part runs **6_Tables.do** do-file that will generate tables 1, 2, 3, 4 and 5 used in the main text of the paper.

PART 3.5: This part runs **7_Annex_tables.do** will generate tables A2 and A3 in annex.

Kindly designate the user root-path in line 2 within the Master do file (**master-dofile.do**). You have the option to execute the entire do file or select specific sections. It's crucial, however, to run the do file in sequence. If opting to run only one section, ensure that preceding sections have been executed beforehand.

List of tables and programs
Selected tables in the paper, as explained and justified below.

Table/Figure	Program	Line Number	Output file
Figure 1	Do files/ 5_Figures.do	45	Figures.xlsx, sheet 'Figure 1'
Figure 2	Do files/ 5_Figures.do	81	Figures.xlsx, sheet 'Figure 2'
Figure 3	Do files/ 5_Figures.do	116	Figures.xlsx, sheet 'Figure 3'
Figure 4	Do files/ 5_Figures.do	151	Figures.xlsx, sheet 'Figure 4'
Figure 5	Do files/ 5_Figures.do	187	Figures.xlsx, sheet 'Figure 5'
Figure 6	Do files/ 5_Figures.do	223	Figures.xlsx, sheet 'Figure 6'
Figure 7	Do files/ 5_Figures.do	278	Figures.xlsx, sheet 'Figure 7'
Figure 8	Do files/ 5_Figures.do	333	Figures.xlsx, sheet 'Figure 8'
Table 1	Do files/ 6_Tables.do	22	Table1.xlsx
Table 1	Do files/ 6_Tables.do	34	Table1.xlsx
Table 2	Do files/ 6_Tables.do	48	Table2.xlsx
Table 2	Do files/ 6_Tables.do	58	Table2.xlsx
Table 3	Do files/ 6_Tables.do	76	Table3.xlsx
Table 3	Do files/ 6_Tables.do	82	Table3.xlsx
Table 3	Do files/ 6_Tables.do	87	Table3.xlsx
Table 3	Do files/ 6_Tables.do	92	Table3.xlsx
Table 3	Do files/ 6_Tables.do	97	Table3.xlsx

Table 4	Do files/ 6_Tables.do	111	Table4.xlsx
Table 4	Do files/ 6_Tables.do	117	Table4.xlsx
Table 4	Do files/ 6_Tables.do	122	Table4.xlsx
Table 4	Do files/ 6_Tables.do	127	Table4.xlsx
Table 4	Do files/ 6_Tables.do	132	Table4.xlsx
Table 5	Do files/ 6_Tables.do	146	Table5.xlsx
Table 5	Do files/ 6_Tables.do	152	Table5.xlsx
Table 5	Do files/ 6_Tables.do	157	Table5.xlsx
Table 5	Do files/ 6_Tables.do	162	Table5.xlsx
Table 5	Do files/ 6_Tables.do	167	Table5.xlsx
Table A2	Do files/ 7_Annex_tables.do	17	TableA2.xlsx
Table A2	Do files/ 7_Annex_tables.do	316	TableA2.xlsx
Table A3	Do files/ 7_Annex_tables.do	634	TableA3.xlsx
Table A3	Do files/ 7_Annex_tables.do	640	TableA3.xlsx
Table A3	Do files/ 7_Annex_tables.do	645	TableA3.xlsx
Table A3	Do files/ 7_Annex_tables.do	650	TableA3.xlsx
Table A3	Do files/ 7_Annex_tables.do	655	TableA3.xlsx

Computational requirements

Stata 17 should be used to run the codes.

References

Lars Vilhuber, Connolly, M., Koren, M., Llull, J., & Morrow, P. (2022). A template README for social science replication packages (v1.1). Zenodo. https://doi.org/10.5281/zenodo.7293838

Acknowledgment