

README

Overview

This document explains the steps necessary to replicate the data, figures, and tables used in Ambel and Belete (2025): “*Who benefits from public spending on healthcare? Longitudinal evidence from Ethiopia*” using Stata. The associated files will run all of the codes to generate the final data as well as the 8 figures and 13 tables in the paper. The replicator should expect the code to run for about 10 minutes.

Data Availability and Provenance Statements

☐ This paper does not involve analysis of external data (i.e., no data are used, or the only data are generated by the authors via simulation in their code).

Statement about Rights

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

☐ We certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package. Appropriate permission are documented in the LICENSE.txt file.

Summary of Availability

☐ All data **are** publicly available.

☒ Some data **cannot be made** publicly available.

☐ **No data can be made** publicly available.

☐ Confidential data used in this paper and not provided as part of the public replication package will be preserved for ____ years after publication, in accordance with journal policies.

Details on each Data Source

All raw input data are available in the folder: "\\ReproducibilityPackage\\02_rawdata\\"

1. Ethiopia Socioeconomic Panel Survey (ESPS) 2018/19 (CSA & WB 2020):

Publicly available at the World Bank’s micro-database:

<https://microdata.worldbank.org/index.php/catalog/3823>

These raw .dta data files should be put in one folder as “w4” within folder “02_rawdata”.

2. Ethiopia Socioeconomic Panel Survey (ESPS) 2021/22 (ESS & WB 2023):

Publicly available at the World Bank’s micro-database:

<https://microdata.worldbank.org/index.php/catalog/6161>

These raw .dta data files should be put in one folder as “w5” within folder “02_rawdata”.

3. Healthcare cost per user by service-facility type and region (some data cannot be made publicly available):

The cost per user by service-facility type is computed from three sources: (1) the total healthcare budget by region, obtained from the MOF Ethiopia budget data (not publicly available, accessed in April 2024), (2) the parameters allotted for the healthcare services and facilities used in the study, obtained from the Ethiopia

National Health Accounts report which is publicly available (See Reference: MoH 2022), and (3) the weighted number of healthcare users, compiled from the ESPS 2018/19 (CSA & WB 2020) and 2021/22 (ESS & WB 2023) that are publicly available.

Users must request access to the budget data and obtain permission from the Ministry of Finance. Please contact the IBEX/IFMIS Project Office by phone at +251-11-40-20-00, Ext. 1747, or by email at info-ifmis-ethio@mofed.gov.et

Once the MOF budget data is accessed, users should proceed as follows.

- Aggregate the recurrent health expenditure by region for the two years (2018/19 and 2021/22).
- For 2021/22, include expenditures for the SNNP, Sidama, and South West regions under the SNNP region. Also aggregate the federal (central) government's recurrent health expenditure for both years (to be assigned later to each region as per their share of public healthcare users).
- Save these aggregate data as a separate Excel file, BOOST_recurrentexp.xlsx, with 2019 and 2022 as two sheet names. The headers of the aggregated variables should be "recexp_region" and "recexp_fed".

Finally, the data in 2019 and 2022 sheets should look like the following:

region	recexp_region	recexp_fed
Addis Ababa		
Afar		
Amhara		
Benishangul-Gumuz		
Dire Dawa		
Gambella		
Harari		
Oromia		
SNNP		
Somalie		
Tigray		

Notes: "recexp_region" refers to "Recurrent health expenditure, regional government"; "recexp_fed" refers to "Recurrent health expenditure, federal government". recexp_fed will be the same/single value, for each year, across region rows.

Final data files

Running the following three do-files will provide the final data sets used to generate the tables and figures. All final data outputs are saved in folder "\\ReproducibilityPackage\03_finaldata\". However, they can be run along with all other do-files from the master do-file [00_Master.do].

0a_BIA_datamgmt_w4.do

OUTPUT: sample_indi_w4.dta: Data set of all sampled individuals for 2018/19

0b_BIA_datamgmt_w5.do

OUTPUT: sample_indi_w5.dta: Data set of all sampled individuals for 2021/22

0c_Merged_data.do"

OUTPUTS:

sample_indi_w4w5.dta: Panel data set of all sampled individuals for 2018/19 and 2021/22
sample_indi_w4w5_users: Panel data set of users of healthcare services for 2018/19 and 2021/22

Computational requirements

Software Requirements

Software and version last used: Stata MP v19

☒ The replication package contains one or more programs to install all dependencies and set up the necessary directory structure.

Before running the master do-file [00_Master.do], please install the following three Stata packages: winsor2, covconc, and glcurve. Optionally, ado files of these packages are available in the folder "ReproducibilityPackage\06_ados\".

Controlled Randomness

☐ Random seed is set at line _____ of program _____.

☒ No Pseudo random generator is used in the analysis described here.

Memory, Runtime, Storage Requirements

Approximate time needed to reproduce the analyses on a standard (CURRENT YEAR) desktop machine:

- ☒ <10 minutes
- ☐ 10-60 minutes
- ☐ 1-2 hours
- ☐ 2-8 hours
- ☐ 8-24 hours
- ☐ 1-3 days
- ☐ 3-14 days
- ☐ > 14 days

Approximate storage space needed:

- ☐ < 25 MB
- ☐ 25 MB - 250 MB
- ☒ 250 MB - 2 GB
- ☐ 2 GB - 25 GB
- ☐ 25 GB - 250 GB
- ☐ > 250 GB
- ☐ Not feasible to run on a desktop machine, as described below.

Details

The full code was last run on **Windows 11 Enterprise, version 10.0.26100 Build 26100, 11th Gen Intel(R) Core i5-1145G7 (2.60 GHz), 16 GB RAM**. Computation took about **10 minutes**.

Description of programs/code

Folder path: "\\ReproducibilityPackage\\"

00_Master.do: Contains all folder paths and packages necessary to run the individual do files.

01_dofiles: Contains a folder of all do-files; see details below.

02_rawdata: Contains folders of input data files; see details above.

03_finaldata: All final datasets used to generate tables and figures are available here.

04_tables: Contains all final tables used in the paper.

05_figures: Contains all final figures used in the paper.

06_ados: Contains an option for ado files for Stata packages used: winsor2, covconc, and glcurve.

07_logfile: A log file of all the commands and results are saved in this folder.

Instructions to Replicators

- Please update the global directory to the local folder path on line 11 of the master do-file, "\\ReproducibilityPackage\00_Master.do".
- Download the raw data files referenced above.
- Run "\\ReproducibilityPackage\00_Master.do" to run all steps in sequence.

List of Tables/Figures and Programs

The provided code reproduces:

☒ All numbers provided in text in the paper

☒ All tables and figures in the paper

☐ Selected tables and figures in the paper, as explained and justified below.

Tables

Tables do-files folder path: "\\ReproducibilityPackage\01_dofiles\\"

OUTPUT: All tables are saved in Excel file in folder

"\\ReproducibilityPackage\04_tables\EthBIA_Tables.xlsx".

Table #	Program	Line Number	Output file: Sheet in Excel file EthBIA_Tables.xlsx	Note
Table 1	1a_Table1.do		T1	
Table 2	1b_Table2.do		T2	
Table 3	1c_Table3.do		T3	
Table 4	1d_Table4.do		T4	
Table 5	1e_Table5.do		T5	
Table 6	1f_Table6.do		T6	
Table 7	1g_Table7.do		T7	
Table 8	1h_Table8.do		T8	
Table A1	2a_TableA1.do		TA1	

Table A2	2b_TableA2.do	TA2
Table A3	2c_TableA3.do	TA3
Table A4	2d_TableA4.do	TA4
Table A5	2e_TableA5.do	TA5

Figures

Figures do-files folder path: "\\ReproducibilityPackage\01_dofiles\\"

OUTPUT: All figures, except Figure 5, are saved in folder "\\ReproducibilityPackage\05_figures\\".

Figure #	Program	Line Number	Output file	Note
Figure 1	3a_Figure1.do		Figure1.png	
Figure 2	3b_Figure2.do		Figure2.png	
Figure 3	3c_Figure3.do		Figure3.png	
Figure 4	3d_Figure4.do		Figure4.png	
Figure 5	3e_Figure5.do	04_tables\EthBIA_Tables.xlsx, Sheet name: Figure5		Next: Plotted in Excel and saved as Figure5.xlsx in folder \\ReproducibilityPackage\\"
Figure 6	3f_Figure6.do		Figure6.png	
Figure A1	4a_FigureA1.do		FigureA1.png	
Figure A2	4b_FigureA2.do		FigureA2.png	

References

CSA, & WB. (2020). Ethiopia Socioeconomic Survey (ESS) Report - Wave 4, 2018/19. Central Statistics Agency (CSA) of Ethiopia and World Bank (WB). Addis Ababa.

<https://Microdata.Worldbank.Org/Index.Php/Catalog/3823>

ESS, & WB. (2023). Ethiopia Socioeconomic Panel Survey (ESPS) Report - Wave 5, 2021/22. Ethiopian Statistical Service (CSA) and World Bank (WB). Addis Ababa.

<https://Microdata.Worldbank.Org/Index.Php/Catalog/6161>

MoH. (2022). Ethiopia National Health Accounts Report 2019/20. Ministry of Health - Ethiopia.