

README for Reproducibility Package

Paper: “Measuring Agency Through Psychological Constructs in Lower-Income Settings”

Authors: Clare Clingain, Aletheia Donald, and Maria Hernandez-de-Benito

Date: March 20, 2026

1. Overview

This reproducibility package contains all code, data, and output files necessary to replicate the tables and figures in “Measuring Agency Through Psychological Constructs in Lower-Income Settings.” The paper introduces and validates four psychological agency scales – Goal-Setting Capacity (GSC), Agricultural Self-Efficacy (AGSE), Generalized Livelihoods Self-Efficacy (GLSE), and Short-form Locus of Control (S-LOC) – tested across nine surveys in six Sub-Saharan African countries (Kenya, Malawi, Tanzania, Uganda, Benin, and Cote d’Ivoire). The package reproduces all main text tables and figures (Tables 1, 2, 4; Figures 1-3), as well as all appendix tables and figures (Tables A2-A3, B2-B26, C1-C37; Figures A1-A4).

To replicate all results, open code/Setup.do, set the root path to the location of this package on your machine, and run the file. This master script installs required Stata packages and sequentially executes all analysis do-files.

2. Data Availability Statement

Data Availability

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

The data used in this paper come from nine different surveys conducted across six countries in Sub-Saharan Africa. The data were collected under heterogeneous data agreements and cannot all be made publicly available. The datasets included in this package are de-identified analytical extracts. The data are embargoed until journal publication.

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.

Data Sources

Data Source 1: Kenya IDRC – CAP-Youth Empowerment Institute (CAPYEI) Baseline Survey

Field	Description
Filename(s)	Kenya_IDRC_descsample.dta
Source	World Bank Africa Gender Innovation Lab (GIL) and CAP-Youth Empowerment Institute (CAPYEI)
URL	Not publicly available. Data were collected as part of a baseline survey for a gender-sensitive technical and vocational education training program in Kenya.
Access date	April-May 2021 (data collection); accessed for this study in 2024
Citation	World Bank Africa Gender Innovation Lab (GIL) and CAP-Youth Empowerment Institute (CAPYEI). 2021. "Kenya IDRC – CAP-Youth Empowerment Institute (CAPYEI) Baseline Survey" [dataset]. Accessed 2024.
License	Restricted access. Data collected under project-specific data agreements. Not available for public redistribution. Data embargoed until journal publication. Contact the World Bank Africa Gender Innovation Lab for access requests.

Description: Baseline survey of 1,664 young women aged 15-48 across 18 counties in Kenya, enrolled in a gender-sensitive vocational training program. Contains Goal-Setting Capacity (GSC) and Generalized Livelihoods Self-Efficacy (GLSE) scale items.

Data Source 2: Kenya KYEOP – Kenya Youth Employment Opportunities Project Endline Survey

Field	Description
Filename(s)	kyeop_gse.dta
Source	Government of Kenya and the World Bank
URL	Not publicly available. Data were collected as part of an impact evaluation of the Kenya Youth Employment Opportunities Project (KYEOP).
Access date	September-December 2023 (data collection);

Field	Description
	accessed for this study in 2024
Citation	Government of Kenya and the World Bank. 2023. "Kenya KYEOP – Kenya Youth Employment Opportunities Project Endline Survey" [dataset]. Accessed 2024.
License	Restricted access. Data collected under project-specific data agreements. Not available for public redistribution. Data embargoed until journal publication. Contact the World Bank for access requests.

Description: Endline survey of 9,308 young individuals (5,588 women and 3,720 men) aged 18-37 across 15 counties in Kenya. Contains Generalized Livelihoods Self-Efficacy (GLSE) scale items.

Data Source 3: Tanzania LSMS – Tanzania Methodological Survey Experiment on Household Consumption Measurement

Field	Description
Filename(s)	Tanzania_scale_tool17.dta, Tanzania_scale_tool20.dta, Tanzania_scale_tool22.dta
Source	World Bank Living Standards Measurement Study (LSMS) Program
URL	Not publicly available. Data were collected as part of a nationally representative household survey. See Dang et al. (2023) for survey documentation.
Access date	April-July 2022 (data collection); accessed for this study in 2024
Citation	Dang, H.-A. H., T. Kilic, V. Hlasny, K. Abanokova, and C. Carletto (2023). "Using Survey-to-Survey Imputation to Fill Poverty Data Gaps at a Low Cost: Evidence from a Randomized Survey Experiment," Policy Research Working Paper WPS 10738, World Bank Group, Washington, D.C.
License	Restricted access. Data collected by the World Bank LSMS program under specific data use agreements. Data embargoed until journal publication. Contact the World Bank LSMS team for access requests.

Description: Nationally representative household survey spanning 143 enumeration areas across Mainland Tanzania and Zanzibar. Sample comprises 1,025 female and 392 male respondents. Contains Goal-Setting Capacity (GSC), Agricultural Self-Efficacy (AGSE), and Generalized Livelihoods Self-Efficacy (GLSE) scale items.

Data Source 4: Malawi LSMS – Malawi Time Use Measurement Study

Field	Description
Filename(s)	Malawi_scale_tool_agse.dta, Malawi_scale_tool_gsc.dta, Malawi_scale_tool_loc.dta
Source	World Bank Living Standards Measurement Study (LSMS) Program
URL	Not publicly available. Data were collected as part of a methodological experiment on time use measurement. See Kilic et al. (2024) for survey documentation.
Access date	July 2022-March 2023 (data collection); accessed for this study in 2024
Citation	Kilic, T., T. L. G. Daum, H. Buchwald, G. Seymour, P. M. Mvula, A. C. Munthali, M. Kachinjika, and G. B. Koolwal (2024). “Recording the Time Divide: A Comparative Study of Smartphone- and Recall-Based Approaches to Time Use Measurement,” Policy Research Working Paper 10695, World Bank, Washington, D.C.
License	Restricted access. Data collected by the World Bank LSMS program under specific data use agreements. Data embargoed until journal publication. Contact the World Bank LSMS team for access requests.

Description: Nationally representative survey across 72 enumeration areas in nine districts of Southern and Central Malawi. Sample includes 718 female and 717 male respondents. Contains Goal-Setting Capacity (GSC), Agricultural Self-Efficacy (AGSE), and Short-form Locus of Control (S-LOC) scale items.

Data Source 5: Malawi IFPRI – Women’s Empowerment in Agriculture Index (pro-WEAI)

Field	Description
Filename(s)	malawi_ifpri_loc.dta
Source	International Food Policy Research Institute (IFPRI)
URL	Not publicly available. Data were collected as part of

Field	Description
	the pro-WEAI project in Malawi.
Access date	Accessed for this study in 2024
Citation	International Food Policy Research Institute (IFPRI). "Malawi IFPRI – Women's Empowerment in Agriculture Index (pro-WEAI)" [dataset]. Accessed 2024.
License	Restricted access. Data collected under IFPRI data use agreements. Data embargoed until journal publication. Contact IFPRI for access requests.

Description: Survey of 2,675 women and 1,664 men in Malawi. Contains Short-form Locus of Control (S-LOC) scale items.

Data Source 6: Uganda IRC – Financial Integration in Displacement Survey

Field	Description
Filename(s)	Uganda_IRC_descsample.dta
Source	International Rescue Committee (IRC)
URL	Not publicly available. Data were collected as part of an endline survey of refugee and host community farmers. See Clingain and Rincon (2022) for project documentation.
Access date	Spring 2022 (data collection); accessed for this study in 2024
Citation	Clingain, C. and S. Rincon (2022). "Financial integration in displacement: Pilot findings from Uganda," Tech. rep., International Rescue Committee, https://airbel.rescue.org/projects/financial-integration-in-displacement/ .
License	Restricted access. Data collected under IRC project-specific data agreements. Data embargoed until journal publication. Contact the International Rescue Committee for access requests.

Description: Endline survey of refugee and host community farmers in Yumbe, Uganda. Sample comprises 148 female and 115 male respondents; 38% of women and 66% of men were refugees of South Sudanese origin. Contains Agricultural Self-Efficacy (AGSE) scale items.

Data Source 7: Uganda Kampala – Female Market Vendors Baseline Survey

Field	Description
Filename(s)	kampala_loc.dta
Source	World Bank Africa Gender Innovation Lab (GIL)
URL	Not publicly available. Data were collected as part of a baseline survey for a childcare RCT in Kampala, Uganda. See https://www.socialscienceregistry.org/trials/11467 for trial registration.
Access date	Accessed for this study in 2024
Citation	World Bank Africa Gender Innovation Lab (GIL). "Uganda Kampala – Female Market Vendors Baseline Survey" [dataset]. Accessed 2024.
License	Restricted access. Data collected under project-specific data agreements. Data embargoed until journal publication. Contact the World Bank Africa Gender Innovation Lab for access requests.

Description: Baseline survey of 956 female market vendors in six markets in the Greater Kampala Metropolitan Area, Uganda. Eligibility required working in one of the selected markets and having at least one young child. Contains Short-form Locus of Control (S-LOC) scale items.

Data Source 8: Cote d'Ivoire OLAM – Factory Workers Survey

Field	Description
Filename(s)	civ_olam_loc.dta
Source	World Bank Africa Gender Innovation Lab (GIL)
URL	Not publicly available. Data were collected as part of a survey of factory workers at a cashew processing plant in Cote d'Ivoire. See Donald and Grosset-Touba (2025) for study documentation.
Access date	Accessed for this study in 2024
Citation	Donald, A. and F. Grosset-Touba (2025). "Complementarities in Labor Supply: How Joint Commuting Shapes Work Decisions," Working Paper.
License	Restricted access. Data collected under project-specific data agreements. Data embargoed until journal publication. Contact the World Bank Africa Gender Innovation Lab for access requests.

Description: Survey of factory workers at a cashew processing plant in the outskirts of Abidjan, Cote d'Ivoire. Sample includes 637 female and 120 male respondents. Contains Short-form Locus of Control (S-LOC) scale items.

Data Source 9: Benin SWEDD – Sahel Women's Empowerment and Demographic Dividend Survey

Field	Description
Filename(s)	swedd_ado_loc.dta, swedd_ado1_loc.dta, swedd_ado2_loc.dta, swedd_mothers_loc.dta
Source	World Bank Sahel Women's Empowerment and Demographic Dividend (SWEDD) Project
URL	Not publicly available. Data were collected as part of the SWEDD project baseline survey in Benin.
Access date	Accessed for this study in 2024
Citation	World Bank Sahel Women's Empowerment and Demographic Dividend (SWEDD) Project. "Benin SWEDD – Sahel Women's Empowerment and Demographic Dividend Survey" [dataset]. Accessed 2024.
License	Restricted access. Data collected under SWEDD project data agreements. Data embargoed until journal publication. Contact the World Bank for access requests.

Description: Baseline survey in Benin administered to adolescent girls and their mothers as part of the SWEDD project. Contains Short-form Locus of Control (S-LOC) scale items.

Additional Derived Data

Field	Description
Filename(s)	reg_masterfile.dta
Source	Created by the authors
Description	Merged analytical dataset combining harmonized variables from all nine surveys above. This file is used for the main regression analyses in the paper.

3. Instructions for Replicators

Steps to Reproduce

1. **Set the root path:** Open code/Setup.do and change the global root path in line 21 to the location of this replication package on your machine.
2. **Set variable abbreviation:** In Stata, run `set varabbrev` on before executing the code (or ensure your Stata configuration allows variable abbreviation).
3. **Run the master file:** Execute code/Setup.do. This file will:
 - Install all required Stata packages (if not already installed)
 - Run code/Main_tables_figures.do (Tables 1, 2, 4; Figures 1-3)
 - Run code/CFA_EFA.do (Tables B9-B22)
 - Run code/Appendix.do (Tables A2-A3, B2-B8, B23-B26, C1-C37; Figures A1-A4)
4. All outputs will be saved in the output/ folder.

4. List of Exhibits

All tables and figures in the paper are reproduced by this package.

Exhibit	Output Filename	Script	Note
Table 1	Table_1.tex	Main_tables_figures.do	Scale Content
Table 2	Table_2.tex	Main_tables_figures.do	Summary Statistics
Table 4	Table_4.tex	Main_tables_figures.do	Agency and Socioeconomic Outcomes
Figure 1	Figure_1.png	Main_tables_figures.do	Gender Differences
Figure 2	Figure_2.png	Main_tables_figures.do	Agency and Mental Health
Figure 3	Figure_3.png	Main_tables_figures.do	Agency and Empower

Exhibit	Output Filename	Script	Note
Table A2	Table_A2.tex	Appendix.do	ment Summary Statistics – Female Responde nts
Table A3	Table_A3.tex	Appendix.do	Summary Statistics – Male Responde nts
Figures A1a-A1f	Figure_A1a.png – Figure_A1f.png	Appendix.do	GSC Scale Distributio n
Figures A2a-A2f	Figure_A2a.png – Figure_A2f.png	Appendix.do	AGSE Scale Distributio n
Figures A3a-A3f	Figure_A3a.png – Figure_A3f.png	Appendix.do	GLSE Scale Distributio n
Figures A4a-A4l	Figure_A4a.png – Figure_A4l.png	Appendix.do	S-LOC Scale Distributio n
Table B2	Table_B2.tex	Appendix.do	Cronbach’ s Alpha – GSC
Table B3	Table_B3.tex	Appendix.do	Cronbach’ s Alpha – AGSE
Table B4	Table_B4.tex	Appendix.do	Cronbach’ s Alpha – GLSE
Table B5	Table_B5.tex	Appendix.do	Cronbach’ s Alpha – S-LOC
Table B6	Table_B6.tex	Appendix.do	Inter-item Correlation

Exhibit	Output Filename	Script	Note
Table B7	Table_B7.tex	Appendix.do	– GSC Inter-item Correlation
Table B8	Table_B8.tex	Appendix.do	– AGSE Inter-item Correlation
Tables B9-B10	Table_B9.tex – Table_B10.tex	CFA_EFA.do	– S-LOC CFA/EFA – GSC
Tables B11-B12	Table_B11.tex – Table_B12.tex	CFA_EFA.do	CFA/EFA – AGSE
Tables B13-B14	Table_B13.tex – Table_B14.tex	CFA_EFA.do	CFA/EFA – GSC (additional)
Tables B15-B16	Table_B15.tex – Table_B16.tex	CFA_EFA.do	CFA/EFA – GLSE
Tables B17-B18	Table_B17.tex – Table_B18.tex	CFA_EFA.do	CFA/EFA – GLSE (additional)
Tables B19-B20	Table_B19.tex – Table_B20.tex	CFA_EFA.do	CFA/EFA – S-LOC
Tables B21-B22	Table_B21.tex – Table_B22.tex	CFA_EFA.do	CFA/EFA – S-LOC (additional)
Table B23	Table_B23.tex	Appendix.do	Robustnes s – Alternative Aggregatio n (GSC)
Table B24	Table_B24.tex	Appendix.do	Robustnes s – Alternative Aggregatio n (GLSE)
Table B25	Table_B25.tex	Appendix.do	Robustnes s – Alternative

Exhibit	Output Filename	Script	Note
Table B26	Table_B26.tex	Appendix.do	Aggregation (AGSE) Robustness – Alternative Aggregation (S-LOC)
Tables C1-C21	Table_C1.tex – Table_C21.tex	Appendix.do	Regression Results – Well-being and Socioeconomic Outcomes
Tables C22-C25	Table_C22.tex – Table_C25.tex	Appendix.do	Robustness – Socioeconomic Outcomes (Alternative Aggregation)
Tables C26-C37	Table_C26.tex – Table_C37.tex	Appendix.do	Additional Regression Results

5. Requirements

Computational Requirements

- **Operating System:** Windows, macOS, or Linux
- **Processor:** Standard modern processor sufficient
- **Estimated Runtime:** Approximately 15-30 minutes on a standard desktop computer

Software Requirements

- **Stata** (version 16 or higher recommended)

Stata Packages (installed automatically by Setup.do)

- **reghdfe** – Linear regressions with multiple fixed effects

- `ftools` – Fast Stata tools (required by `reghdfe`)
- `gtools` – Faster Stata commands
- `estout` – Estimation output formatting
- `texsave` – Save Stata datasets as LaTeX tables
- `coefplot` – Coefficient plots

Memory and Storage Requirements

- **RAM:** 4 GB minimum
- **Storage:** Approximately 50 MB for the full package

6. Code Description

Script	Purpose
<code>Setup.do</code>	Master file. Sets the root path, installs required Stata packages, and sequentially executes all other do-files.
<code>Main_tables_figures.do</code>	Produces the main text exhibits: Table 1 (Scale Content), Table 2 (Summary Statistics), Table 4 (Agency and Socioeconomic Outcomes), and Figures 1-3.
<code>CFA_EFA.do</code>	Produces Confirmatory Factor Analysis (CFA) and Exploratory Factor Analysis (EFA) structural validity tables (Tables B9-B22).
<code>Appendix.do</code>	Produces all remaining appendix exhibits: summary statistics by gender (Tables A2-A3), reliability and validity tables (Tables B2-B8, B23-B26), scale distribution figures (Figures A1-A4), and all regression result tables (Tables C1-C37).

7. Folder Structure

26WB replication package/

```
|
|-- README.md                <- This file
|-- code/
|   |-- Setup.do             <- Master do-file (run this)
|   |-- Main_tables_figures.do <- Main text tables and figures
|   |-- CFA_EFA.do           <- CFA/EFA validity tables
|   |-- Appendix.do          <- Appendix tables and figures
```

```
| -- data/
|   |-- reg_masterfile.dta      <- Merged analytical dataset
|   |-- Kenya_IDRC_descsample.dta
|   |-- kyeop_gse.dta
|   |-- Tanzania_scale_tool17.dta
|   |-- Tanzania_scale_tool20.dta
|   |-- Tanzania_scale_tool22.dta
|   |-- Malawi_scale_tool_agse.dta
|   |-- Malawi_scale_tool_gsc.dta
|   |-- Malawi_scale_tool_loc.dta
|   |-- malawi_ifpri_loc.dta
|   |-- Uganda_IRC_descsample.dta
|   |-- kampala_loc.dta
|   |-- civ_olam_loc.dta
|   |-- swedd_ado_loc.dta
|   |-- swedd_ado1_loc.dta
|   |-- swedd_ado2_loc.dta
|   |-- swedd_mothers_loc.dta
|
| -- output/
|   |-- Table_1.tex ... Table_C37.tex  <- All tables in LaTeX format
|   |-- Figure_1.png ... Figure_A41.png <- All figures in PNG format
```