

Guide on how to use the Reproducibility Package for the paper “*CLARE: a Causal machine Learning Approach to Resilience Estimation*”

1. Replication

This reproducibility package contains the codes and data needed to allow interested parties to replicate the results presented in the working paper.

The analysis has been carried out using the software R (version 4.2.0) and Stata (version 18).

All the files needed for this replication are contained in the folder “**Replication package for WBWP–October 2025**”.

The folder “**Codes**” contains the *.R files and the *.do files used to estimate the results presented in the paper (**Tables 3-4-5-6-7-A.2-B.1** and **Figures B.1-B.2**).

The folder “**Data**” contains the datasets file used for the analysis.

All the outputs, intermediate and final ones, are stored in the “**Outputs**” folder. Note that some outputs are also intermediate inputs that lead to the other outputs.

To proceed with the replication, please **change your working directory** in the scripts and **install the required R libraries and Stata commands**. The summary table in the following page reports the input datasets, the corresponding codes, and the outputs.

2. Data Availability and Provenance Statement

The data employed for this analysis come from two main data sources:

- **LSMS-ISA Collection.** We use multi-topic longitudinal household data for four Sub-Saharan African countries (Malawi, Nigeria, Tanzania, Uganda) from the **World Bank’s Living Standard Measurement Survey – Integrated Surveys on Agriculture (LSMS-ISA)**. These data are available [here](#), but permission and formal request are needed to download and use the data. Please refer to the working paper for the list of survey waves employed for each countries, and other details about the construction of the household dataset.
- **EDDI** (Evaporative Drought Demand Index) data collected by NOAA (the US National Oceanic and Atmospheric Administration). This data are publicly available and can be downloaded from [here](#). For more information on the EDDI index, refer to Hobbins et al.

(2016). For this analysis, EDDI values were spatially joined to confidential household coordinates, and the merged dataset is not publicly shareable.

For the full details regarding the construction of the starting datasets, including all the specific variables used in the analysis, please refer to the paper.

Note: Because the LSMS-ISA microdata was merged to restricted data, the public reproducibility package begins from intermediate data, not raw household microdata.

All intermediate datasets used in the analysis, and needed to replicate the results in the paper, are included in the reproducibility package.

The Stata do-files that transform the raw LSMS-ISA data into these intermediate datasets are provided in the folder **“Do files – Variable Construction”** to ensure full transparency.

This allows users who have authorized access to LSMS-ISA and the restricted data to regenerate the intermediate data exactly. For users without access to the restricted raw data, the provided intermediate datasets allow full replication of all results in the paper.

Name of datasets used raw

COUNTRY	INPUT DATA USED
Nigeria	NGA_2010-2019_NUPD_v01_M_Stata NGA_2010_GHSP-W1_v03_M_STATA NGA_2012_GHSP-W2_v02_M_STATA NGA_2015_GHSP-W3_v02_M_Stata NGA_2018_GHSP-W4_v03_M_Stata12
Malawi	MWI_2010-2019_IHPS_v06_M_Stata MWI_2010_IHS-III_v01_M_STATA8 MWI_2010-2013_IHPS_v01_M_Stata MWI_2016_IHS-IV_v04_M_STATA14 MWI_2019_IHS-V_v06_M_Stata
Tanzania	TZA_2008_NPS1_v02_M_STATA_English_labels TZA_2010_NPS-R2_v03_M_STATA8 TZA_2012_NPS-R3_v01_M_STATA8_English_labels TZA_2014_NPS-R4_v03_M_v03_A_EXT_STATA11 TZA_2019_NPD-SDD_v06_M_STATA12
Uganda	UGA_2005_2009_UNPS_v02_M_STATA8 UGA_2010_UNPS_v03_M_STATA12 UGA_2011_UNPS_v02_M_Stata UGA_2013_UNPS_v02_M_STATA8 UGA_2015_UNPS_v02_M_STATA8 UGA_2018_UNPS_v02_M_STATA12 UGA_2019_UNPS_v03_M_STATA14
Please note: Datasets referenced in the do-files but not listed in this table cannot be shared, as they contain sensitive information. These datasets include the real confidential household coordinates used for merging with weather information. Specifically, these can be found in the following do files: nig_panel.do (lines: 943, 957, 958, 967, 968, 969, 970); mlw_w1_long.do (line 420); mlw_w2_long.do (line 407); mlw_w3_long.do (line 388); mlw_w4_long.do (line 390); panel_malawi.do (lines: 97, 98, 99, 100, 240, 241);	

tanzania_panel_new.do (lines: 361, 379, 395, 396, 403, 412, 413, 414); uga_panel.do (lines: 160, 176, 197, 198, 199, 206, 207, 208).

3. References

Hobbins, M. T., Wood, A., McEvoy, D. J., Huntington, J. L., Morton, C., Anderson, M., & Hain, C. (2016). The evaporative demand drought index. Part I: Linking drought evolution to variations in evaporative demand. *Journal of Hydrometeorology*, 17(6), 1745-1761.

Datasets	Codes	Outputs
- dataset_descriptives.csv	STEP 0 – Descriptive_Statistics.do	- Table A.2 in Appendix A exported from the excel file descriptive_statistics.xlsx;
- dataset_start.csv - dataset_all.csv	STEP 1 - Estimation_Forecasting.R	- Figures B.1 and B.2 in Appendix B; - Table B.1 in Appendix B; - dataset_allexp_forecasting_main.csv - results_forecasting_main.csv
- dataset_allexp_forecasting_main.csv - results_forecasting_main.csv	STEP 2 - Aggregation and Evaluation_Forecasting.do	- weights_main.dta - Table 3 (cf code lines 62-63-64-113) - Table 4 (code line 84) - Table 5 (code line 104)
- dataset_noMalawi.csv - dataset_noNigeria.csv - dataset_noTanzania.csv - dataset_noUganda.csv - dataset_all.csv	STEP 1 - Estimation_Out-Of-Country_Malawi.R STEP 1 - Estimation_Out-Of-Country_Nigeria.R STEP 1 - Estimation_Out-Of-Country_Tanzania.R STEP 1 - Estimation_Out-Of-Country_Uganda.R NB: For this step, the sequence in which you execute the four R scripts is not important, as long as each script is run independently.	- results_OOC_Malawi.csv" - results_OOC_Nigeria.csv" - results_OOC_Tanzania.csv" - results_OOC_Uganda.csv" - dataset_Malawi.csv - dataset_Nigeria.csv - dataset_Tanzania.csv - dataset_Uganda.csv
- results_OOC_Malawi.csv" - results_OOC_Nigeria.csv" - results_OOC_Tanzania.csv" - results_OOC_Uganda.csv" - dataset_Malawi.csv - dataset_Nigeria.csv - dataset_Tanzania.csv - dataset_Uganda.csv	STEP 2 - Aggregation and Evaluation_Out-Of-Country.do	- Malawi.dta - Nigeria.dta - Tanzania.dta - Uganda.dta - weights_Malawi.dta - weights_Nigeria.dta - weights_Tanzania.dta - weights_Uganda.dta - Table 6 (cf code lines 157-158-159-187) - Table 7 (code line 178)

Summary table for reproducing the results presented in the paper

