Replication package

Climate Change and Polarization: Some Inconvenient Evidence from Sub-Saharan Africa

F. Clementi, M. Fabiani, V. Molini, F. Schettino

Contact: Vasco Molini (vmolini@worldbank.org)

Overview

The package reproduces all tables and figures in the paper and in Appendices A–C from the analysis files described under Data availability. To run it, set `global root` in `main.do` (the only line to change) and run `main.do`. The steps run in a fixed order:

1. `01_weight_recalibration.do` builds the complete-case estimation samples with recalibrated weights (`data/intermediate/`), for temperature and for rainfall.
2. `02_descriptive_statistics.do` produces the descriptive statistics.
3. `03_reldist_analysis.do` runs the relative distribution analysis.
4. `04_shapley_decomposition.do` estimates the symmetrized Shapley decomposition in four specifications: temperature/eb (baseline), rainfall/eb, temperature/ipw and rainfall/ipw.
5. `main.do` copies the graphs of steps 3 and 4 to `outputs/figures/`, one file per figure of the paper (`Figure_1.gph` … `Figure_C4.gph`).

Steps 2–4 use the samples written by step 1.

Path	Content
`main.do`	master do-file; runs the whole workflow
`README.pdf`	this file
`code/`	do-files 01–04; `ado/reldist.ado` (patched)
`data/raw/`	`africa_raw_temperature.dta`, `africa_raw_rainfall.dta` (restricted; not included in the public package, see Data availability)
`data/intermediate/`	written by step 1
`outputs/`	one subfolder per step; `04_shapley/<climate>_<balance>/`; `figures/` (one file per figure of the paper)

Data availability

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.

☐ I certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package.

The household survey microdata (Source 1) are restricted and cannot be redistributed under the World Bank datalibweb terms of use. The analysis files built from them (`data/raw/africa_raw_temperature.dta` and `data/raw/africa_raw_rainfall.dta`) are therefore **not included in the public replication package**. They have been provided confidentially to the journal's data editor for the sole purpose of verifying reproducibility. The climate data (Source 2) are public.

To allow researchers who have access to the original microdata to rebuild the two analysis files, the procedure we followed to construct them is described in detail below, under "Construction of the pre-processed analysis files". Following it step by step should lead to the same two files used by this package.

Summary of data sources

#	Source	Access	Files used by the code
1	Harmonized household consumption survey microdata, World Bank (datalibweb)	Restricted (World Bank staff)	`data/raw/africa_raw_temperature.dta`, `data/raw/africa_raw_rainfall.dta` (survey variables); not included in the public package
2	Climate Change Knowledge Portal, observed climate data (CRU TS), World Bank	Public, free (CC BY 4.0)	merged into the same two files (temperature, precipitation)

Pre-processed analysis files

The two files in `data/raw/` are not the original survey files. They are derived datasets prepared by the authors outside this package: the harmonized surveys of 27 countries (40 surveys, listed in Table D2) were processed (welfare aggregate in 2017 PPP, person-level weights, covariates) and pooled into a single household-level file, to which the CCKP climate variables were merged by admin1 region and survey year. The two files differ only in the climate variable. Steps 1–4 of `main.do` start from these files; no step downloads, processes or merges the original sources.

Not all of these surveys enter the estimation sample: the complete-case selection performed by step 1 retains 35 surveys from 24 countries (Table D1). The construction of the two files is documented step by step under "Construction of the pre-processed analysis files", after the description of the two sources from which they are derived.

Source 1. Harmonized household consumption surveys (datalibweb)

- **Source and producer:** household budget and Living Standards Measurement Study surveys for 27 sub-Saharan African countries (2002–2021), collected by the respective national statistical offices and harmonized by the World Bank (Sub-Saharan Africa Team for Statistical Development, AFR-TSD), distributed through datalibweb, the World Bank Stata interface to its microdata catalogs.
- **URL / catalogue:** https://github.com/worldbank/datalibweb (access tool and documentation). Original, non-harmonized surveys: https://microdata.worldbank.org and https://www.worldbank.org/en/programs/lsms.
- **Access type:** restricted. datalibweb is available only on World Bank computers to users with a valid World Bank labour contract; each collection requires a separate permission from the regional/collection focal point.
- **Licence / terms of use:** datalibweb terms of use (https://github.com/worldbank/datalibweb, section "Terms of use"): use limited to World Bank business, no redistribution, aggregated reporting only, respondent privacy to be protected. The code of the datalibweb package is under the MIT licence; this does not apply to the data.
- **Access date:** June 2026.
- **Full citation:** AFR-TSD/World Bank (2026). SSAPOV harmonized household survey data, 27 sub-Saharan African countries, 2002–2021 (surveys listed in Table D2). Accessed June 2026 via datalibweb Stata package.
- **Files:** `data/raw/africa_raw_temperature.dta`, `data/raw/africa_raw_rainfall.dta`. Derived, pooled dataset (see "Pre-processed analysis files"): household-level variables for the 27 countries stacked in a single file (consumption per capita in 2017 PPP, weights, demographics, location, education, labour, assets). Not included in the public package; provided confidentially to the data editor.
- **Procedure to obtain access:**
  - World Bank staff can rebuild the source data by running `datalibweb` (or `dlw`) in Stata on a World Bank computer, after requesting access to the collection from the AFR focal point via the datalibweb website.

- Other researchers should contact the corresponding author, Vasco Molini (vmolini@worldbank.org), who will advise on the conditions under which access can be arranged in compliance with the World Bank's terms of use. General questions on datalibweb: datalibweb@worldbank.org (subject line "datalibweb helpdesk: …").
- The original national surveys can be requested from the World Bank Microdata Library or the national statistical offices under their own terms; the harmonized variables and the pooled analysis files would then need to be reconstructed.

Source 2. Climate Change Knowledge Portal (CCKP), observed climate data

- **Source and producer:** World Bank Group, Climate Change Knowledge Portal, observed climate data derived from the CRU TS gridded dataset (Climatic Research Unit, University of East Anglia), spatially aggregated by the CCKP to first-level administrative units (admin1).
- **URL / catalogue:** https://climateknowledgeportal.worldbank.org/download-data; World Bank Data Catalog: https://datacatalog.worldbank.org/search/dataset/0040276
- **Variables used:** annual mean temperature (°C) and annual accumulated precipitation (mm), admin1 level.
- **Access type:** public, free of charge, no registration required.
- **Licence / terms of use:** Creative Commons Attribution 4.0 (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/; World Bank Terms of Use for Datasets, https://www.worldbank.org/en/about/legal/terms-of-use-for-datasets
- **Access date:** June 2026.
- **Full citation:** World Bank Group (2026). Climate Change Knowledge Portal: Observed Climate Data, CRU TS4.08 0.5-Degree. https://climateknowledgeportal.worldbank.org. https://doi.org/10.57966/tw2k-9h36. Accessed June 2026. Underlying data: Harris et al. (2020).
- **Files:** climate variables merged with Source 1 by admin1 region and survey year in `data/raw/africa_raw_temperature.dta` and `data/raw/africa_raw_rainfall.dta`.
- **Contact:** data@worldbank.org

Construction of the pre-processed analysis files

This section describes the procedure we followed to build `data/raw/africa_raw_temperature.dta` and `data/raw/africa_raw_rainfall.dta` from Sources 1 and 2. It is intended to allow researchers who have access to the harmonized survey microdata to retrace each step and obtain the same two files. All operations described here were carried out outside this package; the operations performed by the code in the package are summarized at the end of the section.

1. Access to the harmonized surveys (datalibweb)

1. Go to https://datalibweb2.worldbank.org/my%20subscriptions and register.
2. Open *Directory*; in the *Select view* window choose "Server" and in *Select server* choose "SSA". Under *Classification*, select "Private": these are the collections that require the authorization of the Task Team Leader (TTL).
3. Tick the boxes of the relevant surveys and add them to the cart.
4. Press the *Subscription card* button: all requests are listed in this window. Provide a short description of the project and press *Submit*.
5. The request is confirmed by email, and a notification is sent once the TTL has approved it.
6. Once the request is approved, go back to https://datalibweb2.worldbank.org/my%20subscriptions, open the profile page and copy your token.
7. In Stata, type `datalibweb, token(<your token>)` and then `datalibweb_inventory, region(SSA)` to list the surveys available for the region.

2. Selection and download of the surveys

We downloaded all the household budget surveys for sub-Saharan Africa available from 2000 to 2025 that include both module P (poverty) and module H (household characteristics). For example, for the first survey of Angola:

```
dlw, coun(AGO) y(2008) t(SSAPOV) mod(P)
dlw, coun(AGO) y(2008) t(SSAPOV) mod(H)
```

where `coun()` is the country code, `y()` the survey year, `t()` the collection (SSAPOV, the harmonized data used for poverty measurement) and `mod()` the module. The same commands were run for all the countries in the region.

For each survey we also downloaded module I (individual characteristics) and module L (labour), which contain the characteristics of the household head, e.g. `dlw, coun(AGO) y(2018) t(SSAPOV) mod(I)` and `dlw, coun(AGO) y(2018) t(SSAPOV) mod(L)`. From these two modules we kept only the household head, so that each file has one record per household.

The two analysis files contain 40 surveys from 27 countries (Table D2): Angola (2008, 2018), Benin (2015, 2018, 2021), Botswana (2015), Burkina Faso (2003, 2018, 2021), Burundi (2013, 2020), Cameroon (2007, 2021), Central African Republic (2021), Chad (2003, 2018), Comoros (2020), Côte d'Ivoire (2015, 2021), Democratic Republic of the Congo (2004, 2013), The Gambia (2020), Ghana (2016), Guinea (2018), Guinea-Bissau (2021), Kenya (2021), Lesotho (2017), Mali (2021), Mozambique (2002, 2019), Niger (2005, 2021), Rwanda (2016), Senegal (2005, 2021), Sierra Leone (2018), Sudan (2014), Tanzania (2018), Togo (2018) and Uganda (2012). Of these, 35 surveys from 24 countries form the estimation sample used in the paper (Table D1); the other five drop out at step 1 of the package, as explained in point 11.

3. Household-level files

For each survey, modules P and H were merged one-to-one with modules I and L on the household identifier (`hid`). Some files contain two survey rounds (e.g. the 2018 and 2021 surveys of Benin): in these cases we first kept the round of interest (variable `year`) and then merged on both `hid` and `year`.

4. Welfare aggregate

The welfare variable is the per capita consumption aggregate of module P (`wel_PPP`), expressed in local currency at the prices of the survey year. We deflated it to 2017 prices with the national consumer price index (CPI) and converted it into 2017 PPP dollars:

consumption (2017 PPP $) = `wel_PPP` × CPI(2017) / CPI(survey year) / PPP(2017)

We used the CPI (2010 = 100) of the World Development Indicators and the 2017 PPP conversion factor reported in the harmonized data (variable `icp2017`). For Angola 2008, for example, the CPI is 76.8114 in 2008 and 271.0199 in 2017 and the PPP conversion factor is 107.27, so that consumption in 2017 PPP dollars is `wel_PPP` × 271.0199 / 76.8114 / 107.27. Two exceptions apply: for Benin 2015 the consumption aggregate is the variable `wel_PPPdr`; for Comoros 2020 the harmonized data already report the ratio between the CPI of the survey year and that of 2017 (variable `cpi2017`), so the consumption aggregate was divided by this ratio and by the PPP conversion factor.

5. Weights

We used the population weight of the harmonized data (`wta_pop`), equal to the household weight multiplied by household size, so that all statistics refer to individuals rather than households.

6. Covariates

The covariates refer to the household head and to the household. From the harmonized data we took the sex and age of the household head (`sex`, `ageyrs`), household size (`hhsize`) and area of residence, urban or rural (`rururb`). Education of the household head has four levels (`educat4`): no education, primary, secondary and tertiary; missing values were coded as no education.

The labour market status of the household head has six categories, built from the labour status (`lstatus`) and the sector of the main job (`industrycat4`): unemployed (`lstatus` = 2), not in the labour force (`lstatus` =

3) and, for employed heads, agriculture, industry, services and other (`industrycat4` = 1, 2, 3 and 4). Households with missing labour status were coded as unemployed. The only exception is Angola 2008, where the status is built from the variables `employed_WB`, `unemployed_WB` and `industrycat4_1`, and missing values were coded as not in the labour force.

7. Asset index

For each survey we built an asset index from the dummies on dwelling characteristics and durable goods available in the harmonized data: piped water, electricity, home ownership, improved sanitation, garbage disposal, radio, television, phone, fridge, computer, washing machine, stove, bicycle, car and internet (not all of them are available in every survey). Categorical variables were recoded as dummies (e.g. home ownership equals 1 if the household owns the dwelling and 0 otherwise) and missing values were set to 0. We then ran a principal component factor analysis on these dummies, weighted by the population weight (in Stata: `factor` followed by the list of dummies, `[aw=wta_pop], pcf`), predicted the score of the first factor and normalized it between 0 and 1 as (score − minimum) / (maximum − minimum). The index is computed separately for each survey.

8. Pooling, period and macro region

All the surveys were appended into a single dataset, and two variables were created. The first is the period (`period`), equal to 1 for the surveys fielded between 2002 and 2014 and to 2 for those fielded between 2015 and 2021; the variable `anno` labels the two periods as 2010 and 2020, as in Table D1. The second is the macro region (`area`), with four areas: Western Africa (Benin, Burkina Faso, Côte d'Ivoire, The Gambia, Ghana, Guinea, Guinea-Bissau, Mali, Niger, Senegal, Sierra Leone and Togo), Central Africa (Angola, Cameroon, Central African Republic, Chad, Democratic Republic of the Congo and Sudan), Eastern Africa (Burundi, Comoros, Kenya, Mozambique, Rwanda, Tanzania and Uganda) and Southern Africa (Botswana and Lesotho). In the analysis Eastern and Southern Africa are grouped together, because in the first period there are almost no observations for the Southern countries.

9. Climate variables

From the CCKP (https://climateknowledgeportal.worldbank.org/) we downloaded, for each country and in Excel format, the historical annual series of mean temperature (°C) and precipitation (mm) for the first-level administrative regions (admin1), based on the CRU TS data of the University of East Anglia (Source 2).

For each region we computed in Excel the percentage variation of temperature and precipitation in 2005 and in 2015 with respect to the average of the previous fifteen years:

- 2005: (value in 2005 − average 1990–2004) / average 1990–2004 × 100;
- 2015: (value in 2015 − average 2000–2014) / average 2000–2014 × 100.

The Excel files were imported into Stata and appended into a single file with one record per region, containing the country (the first three characters of the regional code are the ISO country code), the name of the region and the variations of temperature and precipitation for 2005 and 2015.

10. Matching of regions and merge with the household data

The climate data are merged with the household data on country, region and survey year; in the household data the region is the variable `region` of the harmonized data. Since some countries have more than one survey, and the regions used by the surveys can change from one survey to another, the climate records of each country were duplicated for each of its surveys (e.g. the regions of Senegal were duplicated for 2005 and 2021) and the survey year was added.

To allow the merge, the regions of the climate data were modified to match those used in each survey, merging or splitting them where necessary:

- **Different spelling:** the region was renamed as in the survey (e.g. Bié in the climate data and Bie in the survey for Angola; Ségou and Segou for Mali).
- **Larger regions or older boundaries in the survey:** the climate regions were merged into the survey region. In the Central African Republic the prefectures were grouped into the seven regions of the survey (e.g. Ombella M'Poko and Lobaye form the region Oubangui); in Chad 2003 the provinces were grouped as in

the survey (Barh el Gazel, Kanem and Lac form one region); in Mozambique 2002 the provinces were grouped into North, Central and South; in Ghana 2016 the new regions were brought back to the ten old ones (Ahafo and Bono East into Brong Ahafo); in Senegal 2005 the regions created after 2005 were brought back to the regions they were part of (Kaffrine into Kaolack, Sédhiou into Kolda and Kédougou into Tambacounda); in Botswana the districts were grouped into the strata of the survey.

- **Survey regions not present in the climate data:** the climate region was split. In Botswana, Francistown takes the values of Rural North-East, and "Other cities & towns" the values of Gaborone.
- **Tanzania:** the regions of Zanzibar were dropped.

When several climate regions are merged into one survey region, we used the average of their variations. The climate data were then merged with the household data (many-to-one merge on country, region and survey year). The 2005 variation was assigned to the households of the surveys fielded between 2002 and 2008, the 2015 variation to those of the surveys fielded from 2012 onwards; the two variations are stored in the variables `Var_2005` and `Var_2015`, so that for each household only one of them is non-missing. Households whose region has no match in the climate data have missing climate variables.

The resulting household-level dataset (379,283 households) was saved in two versions, which differ only in the climate variables: in `data/raw/africa_raw_temperature.dta` the variables `Var_2005` and `Var_2015` refer to temperature, in `data/raw/africa_raw_rainfall.dta` to precipitation. Both files contain the following variables:

Variable	Content
`country`	ISO 3166 alpha-3 country code
`surveyear`	survey year
`period`	period: 1 = surveys fielded between 2002 and 2014; 2 = surveys fielded between 2015 and 2021
`anno`	period label: 2010 (first period), 2020 (second period)
`area`	macro region: Western, Central, Eastern, Southern Africa
`regione`, `subregione`	regional identifiers from the harmonized data, used to match the climate data
`hid`	household identifier
`wta_pop`	population weight (household weight × household size)
`wel_PPP`	per capita consumption in 2017 PPP dollars (point 4)
`hhsize`	household size
`rururb`	area of residence: 0 = rural, 1 = urban
`sex`	sex of the household head: 0 = female, 1 = male
`age`	age of the household head (years)
`educat4`	education of the household head: no education, primary, secondary, tertiary
`lab_sector`	labour market status of the household head: unemployed, not in labor force, agriculture, industry, services, other
`asset_normalized`	asset index, normalized between 0 and 1 (point 7)
`Var_2005`, `Var_2015`	percentage variation of the climate variable in 2005 and 2015 with respect to the average of the previous fifteen years (point 9)

11. Steps performed by the package

All subsequent operations are carried out by the code in this package, starting from the two files above. `01_weight_recalibration.do` drops the households with missing values in consumption, in the covariates or in the climate variables, where temperature and precipitation are considered jointly so that the estimation sample is the same with the two climate variables, and recalibrates the weights so that each country

keeps, in each period, the same total weight it had before the deletion: within each country and period, the weights of the remaining households are multiplied by the ratio between the sum of the weights before and after the deletion. Since the first period uses `Var_2005` and the second `Var_2015`, five of the 40 surveys drop out entirely at this stage: Comoros 2020, for which consumption is missing for all households, and the four first-period surveys fielded between 2012 and 2014 (Burundi 2013, Democratic Republic of the Congo 2013, Sudan 2014 and Uganda 2012), which carry the 2015 rather than the 2005 climate variation. The resulting estimation sample includes 35 surveys from 24 countries (Table D1), with 75,047 households in the first period and 249,842 in the second (Table D2). `03_reldist_analysis.do` and `04_shapley_decomposition.do` then compute the relative distribution, the polarization indices (MRP, LRP and URP) and their symmetrized Shapley decomposition with `reldist` (Jann, 2021), always using the recalibrated weights, with the second period compared with the first (reference) period. Standard errors and 95% confidence intervals are obtained by bootstrap, resampling households within each period (see Computational requirements).

Computational requirements

- Stata 17 or later.
- Packages: `moremata` and `kdens` (installed from SSC by `main.do` if missing), and a patched `reldist` 1.3.1 (Jann, 2021) shipped in `code/ado/`, which `main.do` places at the front of the adopath.
- Disk space: about 300 MB.
- Runtime: most of it comes from the bootstrap in steps 3 and 4, and therefore varies with the number of replications and with the machine's performance; a full run is expected to take a long time. `main.do` records the time taken by each step in `outputs/main.log`.
- Bootstrap settings: the number of replications and the seed are set once in `main.do` (`global breps 200`, `global bseed 12345`) and read by both `03_reldist_analysis.do` and `04_shapley_decomposition.do`. These values reproduce the paper. For an integrity check of the workflow, set `global breps` to a small value (e.g. 5): point estimates are unaffected, while standard errors and confidence intervals will differ from those in the paper.

List of tables and figures

The figures were finalized in Stata's Graph Editor, so the .gph files match the paper in content but not in appearance. All paths are relative to `outputs/`.

Exhibit	Program	Output file
Table 1	02	`02_descriptive_statistics/descriptive_statistics.docx`
Figure 1	03	`figures/Figure_1.gph`
Table 2	04 (temperature, eb)	`04_shapley/temperature_eb/shapley_group_avg.smcl`
Table 3	04 (temperature, eb)	`04_shapley/temperature_eb/shapley_group_avg.smcl`
Table 4	04 (temperature, eb)	`04_shapley/temperature_eb/shapley_group_avg.smcl`
Figure 2	04 (temperature, eb)	`figures/Figure_2.gph`
Figure 3	04 (temperature, eb)	`figures/Figure_3.gph`
Figure 4	04 (temperature, eb)	`figures/Figure_4.gph`
Figure 5	04 (temperature, eb)	`figures/Figure_5.gph`
Table A1	04 (rainfall, eb)	`04_shapley/rainfall_eb/shapley_group_avg.smcl`
Table A2	04 (rainfall, eb)	`04_shapley/rainfall_eb/shapley_group_avg.smcl`
Figure A1	04 (rainfall, eb)	`figures/Figure_A1.gph`

Exhibit	Program	Output file
Figure A2	04 (rainfall, eb)	`figures/Figure_A2.gph`
Figure A3	04 (rainfall, eb)	`figures/Figure_A3.gph`
Figure A4	04 (rainfall, eb)	`figures/Figure_A4.gph`
Table B1	04 (temperature, ipw)	`04_shapley/temperature_ipw/shapley_group_avg.smcl`
Table B2	04 (temperature, ipw)	`04_shapley/temperature_ipw/shapley_group_avg.smcl`
Figure B1	04 (temperature, ipw)	`figures/Figure_B1.gph`
Figure B2	04 (temperature, ipw)	`figures/Figure_B2.gph`
Figure B3	04 (temperature, ipw)	`figures/Figure_B3.gph`
Figure B4	04 (temperature, ipw)	`figures/Figure_B4.gph`
Table C1	04 (rainfall, ipw)	`04_shapley/rainfall_ipw/shapley_group_avg.smcl`
Table C2	04 (rainfall, ipw)	`04_shapley/rainfall_ipw/shapley_group_avg.smcl`
Figure C1	04 (rainfall, ipw)	`figures/Figure_C1.gph`
Figure C2	04 (rainfall, ipw)	`figures/Figure_C2.gph`
Figure C3	04 (rainfall, ipw)	`figures/Figure_C3.gph`
Figure C4	04 (rainfall, ipw)	`figures/Figure_C4.gph`

Each file in `outputs/figures/` is a copy, made by `main.do`, of the corresponding graph in `03_reldist/` or `04_shapley/<climate>_<balance>/`; the mapping is listed in section 5 of `main.do`.

Table D1. Countries in the estimation sample

Countries and periods in the estimation sample (24 countries). Year 2010 identifies the reference period *t*0 (surveys fielded between 2002 and 2008), year 2020 the comparison period *t*1 (surveys fielded between 2015 and 2021).

Code	Country	Year
AGO	Angola	2010
AGO	Angola	2020
BDI	Burundi	2020
BEN	Benin	2020
BFA	Burkina Faso	2010
BFA	Burkina Faso	2020
BWA	Botswana	2020
CAF	Central African Republic	2020
CIV	Côte d'Ivoire	2020
CMR	Cameroon	2010
CMR	Cameroon	2020
COD	Congo, Dem. Rep.	2010
GHA	Ghana	2020
GIN	Guinea	2020
GMB	Gambia, The	2020

Code	Country	Year
GNB	Guinea-Bissau	2020
KEN	Kenya	2020
LSO	Lesotho	2020
MLI	Mali	2020
MOZ	Mozambique	2010
MOZ	Mozambique	2020
NER	Niger	2010
NER	Niger	2020
RWA	Rwanda	2020
SEN	Senegal	2010
SEN	Senegal	2020
SLE	Sierra Leone	2020
TCD	Chad	2010
TCD	Chad	2020
TGO	Togo	2020
TZA	Tanzania	2020

Table D2. Surveys in the pre-processed analysis files

All the surveys included in `data/raw/africa_raw_temperature.dta` and `data/raw/africa_raw_rainfall.dta` (40 surveys, 27 countries), with the number of households in the analysis files and in the estimation sample built by step 1 (– = survey not in the estimation sample; see "Construction of the pre-processed analysis files", point 11). Period 1 corresponds to *t*0 (2010) and period 2 to *t*1 (2020) in Table D1.

Code	Country	Survey year	Period	Households (analysis files)	Households (estimation sample)
AGO	Angola	2008	1	9,002	9,002
AGO	Angola	2018	2	11,934	11,934
BDI	Burundi	2013	1	6,681	–
BDI	Burundi	2020	2	7,993	7,855
BEN	Benin	2015	2	19,920	19,910
BEN	Benin	2018	2	7,941	7,941
BEN	Benin	2021	2	7,529	7,529
BFA	Burkina Faso	2003	1	8,237	8,237
BFA	Burkina Faso	2018	2	7,010	7,010
BFA	Burkina Faso	2021	2	7,176	7,176
BWA	Botswana	2015	2	7,058	7,058
CAF	Central African Republic	2021	2	6,411	4,609
CIV	Côte d'Ivoire	2015	2	12,899	12,899
CIV	Côte d'Ivoire	2021	2	12,879	12,879

Code	Country	Survey year	Period	Households (analysis files)	Households (estimation sample)
CMR	Cameroon	2007	1	10,353	10,353
CMR	Cameroon	2021	2	7,990	7,990
COD	Congo, Dem. Rep.	2004	1	11,903	11,903
COD	Congo, Dem. Rep.	2013	1	21,239	–
COM	Comoros	2020	2	5,568	–
GHA	Ghana	2016	2	14,009	14,009
GIN	Guinea	2018	2	8,243	8,240
GMB	Gambia, The	2020	2	13,488	13,488
GNB	Guinea-Bissau	2021	2	4,327	4,327
KEN	Kenya	2021	2	16,738	16,732
LSO	Lesotho	2017	2	4,264	4,263
MLI	Mali	2021	2	6,143	6,143
MOZ	Mozambique	2002	1	8,700	8,700
MOZ	Mozambique	2019	2	13,292	13,292
NER	Niger	2005	1	6,690	6,690
NER	Niger	2021	2	6,159	6,159
RWA	Rwanda	2016	2	14,579	14,579
SDN	Sudan	2014	1	11,953	–
SEN	Senegal	2005	1	13,542	13,467
SEN	Senegal	2021	2	6,307	6,307
SLE	Sierra Leone	2018	2	6,810	6,810
TCD	Chad	2003	1	6,695	6,695
TCD	Chad	2018	2	7,376	7,376
TGO	Togo	2018	2	5,087	5,085
TZA	Tanzania	2018	2	8,270	8,242
UGA	Uganda	2012	1	6,888	–
	Total			**379,283**	**324,889**

References

Harris, I., Osborn, T. J., Jones, P., & Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Scientific Data, 7, 109. https://doi.org/10.1038/s41597-020-0453-3

Jann, B. (2021). Relative distribution analysis in Stata. The Stata Journal, 21(4), 885–951.