

A harmonized dataset of forcibly displaced populations and their hosts

Data Dictionary

1. HOUSEHOLD-LEVEL DATA

1.1 Sample

countrycode

Variable Label: Country code

countrycode is a string variable that specifies the 3-character country code used by the World Bank to identify each country. These are the ISO-standard country codes (ISO 3166-1 alpha-3). Although there are different naming conventions, it is necessary to use those specified to ensure that the data for each country is appropriately labeled.

countryname

Variable Label: Country name

countryname is a string variable that specifies the country name used by the World Bank to identify each country.

hhid

Variable Label: Household identifier

hhid specifies the unique household identification number in the harmonized data file within each country (ex. hhid = 2381 is unique within JOR_2016_SRHCS, but by coincidence a different household in IRQ_2015_SRHCS also has hhid = 2381). The original format, string or numeric, of original data should be kept.

int_month

Variable Label: Interview month

int_month is a numeric variable that specifies the month when the survey questionnaire was administered to the household. Categories after harmonization:

- 1 - January
- 2 - February
- 3 - March
- 4 - April
- 5 - May
- 6 - June
- 7 - July
- 8 - August

Commented [1]: I believe hhid is unique by country. For example, there is a hhid==2381 for both Jordan and Iraq. Could you please double-check and adapt this description if necessary?

Commented [2]: Thanks you are correct - I fixed the wording. I worked with the country-specific data in isolation so often I forgot this "harmonized data" includes all countries.

9 - September
10 - October
11 - November
12 - December

int_year

Variable Label: Interview year

int_year is a numeric variable that specifies the year when the survey questionnaire was administered to the household.

rep_year

Variable Label: Reporting year where majority of survey took place

rep_year is a numeric variable that specifies the reporting year where the majority of the survey took place. This is most commonly the year used in the name of the survey.

strata

Variable Label: Strata of the survey

strata refer to the division of the target population – typically the census sample frame -- into subpopulations based on auxiliary information that is known about the full population. Sampling is conducted separately for each strata. The strata should be mutually exclusive: every element in the population must be assigned to only one stratum. The strata should also be collectively exhaustive: no population element can be excluded. Sampling strata need to be considered when constructing the variance (or confidence intervals) of population estimates. strata is needed for the correct calculation of standard deviation for each sample design. A unique identifier is created for each stratum. In Stata, users are advised to specify strata through the svyset command. The variable is in numeric format.

survey_name

Variable Label: Abbreviated survey name (CCC_YYYY_NNNN)

survey_name is a string variable that specifies that abbreviated survey name in the format CCC_YYYY_NNNN, where C = countrycode, Y = rep_year, and N = abbreviated survey name.

surveyprd

Variable Label: Survey period (start-end years)

surveyprd is a string variable that specifies the survey period start and end years. For several surveys, survey period covers two years and this should reflect this.

weight

Variable Label: Household weight

weight contains household weights, typically inversely proportional to the probability of the household being selected for the sample, that should be applied to make the results representative of the population. [For details on the construction of weights, please refer to the Harmonized database's description and the documentation accompanying each survey.](#)

Commented [3]: I agree with the "documentation accompanying each survey", but what is the "Harmonized database's description"? Is that something you are creating?

1.2 Current Assets

ac

Variable Label: Ownership of air conditioner

ac is a dummy variable indicating whether the household owns a central or wall air conditioner. Air conditioner ownership does not depend on who owns the asset within the household, nor on its condition. Two categories after harmonization:

1 - Yes

0 - No

agrilandownti

Variable Label: Type of ownership document for agricultural land

agrilandownti is a categorical variable that specifies the type of document that a household has to prove agricultural land ownership. The two customary rights categories (3 and 4) differentiate whether issued by plot or as a joined group title. Customary groups and cooperatives are differentiated, as well. Customary groups not required to have formal membership declared, while cooperative members have formalized status. If household has several plots under several types of ownership, the harmonizer should collapse the plots by area and then pick the type of ownership for the largest area. Categories after harmonization:

1 - Title; deed

2 - Leasehold (govt issued)

3 - Customary land certificate/plot level

4 - Customary based / group right

5 - Cooperative

6 - Other

boiler

Variable Label: Ownership of boiler/water heater

boiler is a dummy variable indicating whether the household owns a boiler. Boiler ownership does not depend on who owns the asset within the household, nor on its condition. Two categories after harmonization:

1 - Yes

0 - No

cellphone

Variable Label: Ownership of a cellular/mobile phone

cellphone is a dummy variable indicating whether anyone in the household owns a cell phone. Cell phone ownership does not depend on who owns the cellphone is within the household, nor on its condition. Two categories after harmonization:

1 - Yes

0 - No

computer

Variable Label: Ownership of a computer

computer is a dummy variable indicating whether the household owns a computer, including desktop and laptop computer. Computer ownership does not depend on who owns the computer within the household, nor on its condition. Two categories after harmonization:

1 - Yes

0 - No

cookstove

Variable Label: Ownership of improved cookstove

cookstove is a dummy variable indicating whether the household owns an improved stove.

Stove generally refers to a portable or fixed apparatus that burns fuel or uses electricity to provide heat for cooking or heating purposes and includes a cooker (stove). Improved stove ownership does not depend on who owns the asset within the household. Two categories after harmonization:

1 - Yes

0 - No

mvehicle

Variable Label: Ownership of a motorized vehicle

mvehicle is a dummy variable indicating whether the household owns a motorized vehicle (i.e. car or truck). Motorized vehicle ownership does not depend on who owns the asset within the household, nor on its condition. Two categories after harmonization:

1 - Yes

0 - No

nmvehicle

Variable Label: Ownership of a non-motorized vehicle

nmvehicle is a dummy variable indicating whether the household owns a non-motorized vehicle (i.e. bicycle). Non-motorized vehicle ownership does not depend on who owns the asset within the household, nor on its condition. Two categories after harmonization:

1 - Yes

0 - No

oven

Variable Label: Ownership of oven

oven is a dummy variable indicating whether the household owns a oven. Oven ownership does not depend on who owns the asset within the household, nor on its condition. Two categories after harmonization:

1 - Yes

0 - No

ownagriland

Variable Label: Ownership of agriculture land

ownagriland is a dummy variable that specifies whether a household owns agricultural land (yes/no). Owned land can be by freehold, deed, customary, or government leasehold. Only those households that declared using agricultural land. Two categories after harmonization:

1 - Yes

0 - No

ownhouse

Variable Label: Ownership of house

ownhouse is a dummy variable that specifies whether a household owns their house.

Ownership (1) includes ownership or other equivalent of secure tenure, whether or not full-payment has been made yet. Two categories after harmonization:

1 - Yes

0 - No

ownland

Variable Label: Ownership of land

ownland is a dummy variable that specifies whether a household owns residential land (yes/no).

Ownership for property versus residential land on which property is constructed can be different in certain jurisdictions (land vested in a state or municipality). Two categories after harmonization:

1 - Yes

0 - No

ownresland

Variable Label: Ownership of residential land

ownresland is a dummy variable indicating whether the household owns residential land.

Residential land ownership does not depend on who owns the asset within the household, nor on its condition. Two categories after harmonization:

1 - Yes

0 - No

radio

Variable Label: Ownership of a radio

radio is a dummy variable indicating whether the household owns a radio (i.e. radio, radio cassette, and 3-in-1 radio cassette player (radio). Radio ownership does not depend on who owns the radio within the household, nor on its condition. Two categories after harmonization:

1 - Yes

0 - No

refrigerator

Variable Label: Ownership of refrigerator

refrigerator is a dummy variable indicating whether the household owns a refrigerator (i.e. refrigerator or freezer). It does not include cooler, icebox or ice chest. Refrigerator ownership

does not depend on who owns the asset within the household, nor on its condition. Two categories after harmonization:

1 - Yes

0 - No

tv

Variable Label: Ownership of a television

tv is a dummy variable indicating whether the household owns a TV set. This includes both color and black and white TVs. TV set ownership does not depend on who owns the TV set within the household, nor on its condition. Two categories after harmonization:

1 - Yes

0 - No

washingm

Variable Label: Ownership of washing machine

washingm is a dummy variable indicating whether the household owns a machine for washing clothes and household linen; but does not include non-electric washing machine. Washing machine ownership does not depend on who owns the asset within the household, nor on its condition. Two categories after harmonization:

1 - Yes

0 - No

1.3 Previous Assets

previous_ac

Variable Label: Previous: Ownership of air conditioner

previous_ac is a dummy variable that specifies whether the household previously owned a central or wall air conditioner. Previous air conditioner ownership does not depend on who owns the asset within the household, nor on its condition. Date before which previous ownership is asked about is specified by previous_date. Two categories after harmonization:

1 - Yes

0 - No

previous_boiler

Variable Label: Previous: Ownership of a boiler/water heater

previous_boiler is a dummy variable that specifies whether the household previously owned a boiler. Boiler ownership does not depend on who owns the asset within the household, nor on its condition. Date before which previous ownership is asked about is specified by previous_date. Two categories after harmonization:

1 - Yes

0 - No

previous_cellphone

Variable Label: Previous: Ownership of a cell phone

previous_cellphone is a dummy variable that specifies whether the household previously owned a cell phone. Cell phone ownership does not depend on who owns the cellphone is within the household, nor on its condition. Date before which previous ownership is asked about is specified by previous_date. Two categories after harmonization:

1 - Yes

0 - No

previous_computer

Variable Label: Previous: Ownership of a computer

previous_computer is a dummy variable that specifies whether the household previously owned a computer, including desktop and laptop computer. Computer ownership does not depend on who owns the computer within the household, nor on its condition. Date before which previous ownership is asked about is specified by previous_date. Two categories after harmonization:

1 - Yes

0 - No

previous_cookstove

Variable Label: Previous: Ownership of improved cookstove

previous_cookstove is a dummy variable that specifies whether the household previously owned an improved stove. Stove generally refers to a portable or fixed apparatus that burns fuel or uses electricity to provide heat for cooking or heating purposes and includes a cooker (stove). Improved stove ownership does not depend on who owns the asset within the household. Date before which previous ownership is asked about is specified by previous_date. Two categories after harmonization:

1 - Yes

0 - No

previous_date

Variable Label: Date before which previous asset questions are asked about

previous_date is a string variable that specifies the date before which previous asset questions are asked about. For example: "July 1, 2017" means previous asset questions were asked about June 30, 2017 and before.

previous_mvehicle

Variable Label: Previous: Ownership of a motorized vehicle

previous_mvehicle is a dummy variable that specifies whether the household previously owned a motorized vehicle (i.e. car or truck). Motorized vehicle ownership does not depend on who owns the asset within the household, nor on its condition. Date before which previous ownership is asked about is specified by previous_date. Two categories after harmonization:

1 - Yes

0 - No

previous_nmvehicle

Variable Label: Previous: Ownership of a non-motorized vehicle

previous_nmvehicle is a dummy variable that specifies whether the household previously owned a non-motorized vehicle (i.e. bicycle). Non-motorized vehicle ownership does not depend on who owns the asset within the household, nor on its condition. Date before which previous ownership is asked about is specified by previous_date. Two categories after harmonization:

1 - Yes

0 - No

previous_oven

Variable Label: Previous: Ownership of oven

previous_oven is a dummy variable that specifies whether the household previously owned a oven. Oven ownership does not depend on who owns the asset within the household, nor on its condition. Date before which previous ownership is asked about is specified by previous_date. Two categories after harmonization:

1 - Yes

0 - No

previous_ownagriland

Variable Label: Previous: Ownership of agriculture land

previous_ownagriland is a dummy variable that specifies whether the household previously owned agricultural land (yes/no). Owned land can be by freehold, deed, customary, or government leasehold. Only those households that declared using agricultural land. Date before which previous ownership is asked about is specified by previous_date. Two categories after harmonization:

1 - Yes

0 - No

previous_ownhouse

Variable Label: Previous: Ownership of house

previous_ownhouse is a dummy variable that specifies whether the household previously owned their house. Ownership (1) includes ownership or other equivalent of secure tenure, whether or not full-payment has been made yet. Date before which previous ownership is asked about is specified by previous_date. Two categories after harmonization:

1 - Yes

0 - No

previous_ownland

Variable Label: Previous: Ownership of land

previous_ownland is a dummy variable that specifies whether the household previously owned residential land (yes/no). Ownership for property versus residential land on which property is constructed can be different in certain jurisdictions (land vested in a state or municipality). Date before which previous ownership is asked about is specified by previous_date. Two categories after harmonization:

1 - Yes

0 - No

previous_ownresland

Variable Label: Ownership of residential land

previous_ownresland is a dummy variable that specifies whether the household previously owned residential land. Residential land ownership does not depend on who owns the asset within the household, nor on its condition. Date before which previous ownership is asked about is specified by previous_date. Two categories after harmonization:

1 - Yes

0 - No

previous_radio

Variable Label: Previous: Ownership of a radio

previous_radio is a dummy variable that specifies whether the household previously owned a radio (i.e. radio, radio cassette, and 3-in-1 radio cassette player (radio). Radio ownership does not depend on who owns the radio within the household, nor on its condition. Date before which previous ownership is asked about is specified by previous_date. Two categories after harmonization:

1 - Yes

0 - No

previous_refrigerator

Variable Label: Previous: Ownership of refrigerator

previous_refrigerator is a dummy variable that specifies whether the household previously owned a refrigerator (i.e. refrigerator or freezer). It does not include cooler, icebox or ice chest. Refrigerator ownership does not depend on who owns the asset within the household, nor on its condition. Date before which previous ownership is asked about is specified by previous_date. Two categories after harmonization:

Two categories after harmonization:

1 - Yes

0 - No

previous_tv

Variable Label: Previous: Ownership of a television

previous_tv is a dummy variable that specifies whether the household previously owned a TV set. This includes both color and black and white TVs. TV set ownership does not depend on who owns the TV set within the household, nor on its condition. Date before which previous ownership is asked about is specified by previous_date. Two categories after harmonization:

1 - Yes

0 - No

previous_washingm

Variable Label: Previous: Ownership of washing machine

previous_washingm is a dummy variable that specifies whether the household previously owned a machine for washing clothes and household linen; but does not include non-electric washing

machine. Washing machine ownership does not depend on who owns the asset within the household, nor on its condition. Date before which previous ownership is asked about is specified by previous_date. Two categories after harmonization:

1 - Yes

0 - No

1.4 Demographics

hh_type

Variable Label: Household is from a refugee/host/IDP sample

hh_type is a categorical variable that specifies whether the household is from a refugee, host, or IDP sample. The category "National sample" is used only for samples of the local population that are representative at the national level. Otherwise, samples of local populations are categorized as Host.

1 - Refugee

2 - Host

3 - IDP

4 - National sample

5 - Venezuelan migrant

hhsz

Variable Label: Household size (# of people in household)

hhsz is a integer variable that specifies the number of people who live in the household. This does not include former members.

live_in_camp

Variable Label: Live in camp or outside camp

live_in_camp is a dummy variable indicating whether the household lives in a camp. All households should be categorized, even if they are not refugees or other displaced households. Categories after harmonization:

0 - No, household lives outside camp

1 - Yes, household lives in camp

origin_country

Variable Label: Household country of origin (if asked at HH-level)

origin_country is a string variable that specifies the country of origin of the household. It is only filled when this question is asked at the household level. It should be filled with the 3-digit ISO-standard country codes (ISO 3166-1 alpha-3).

urban

Variable Label: 1 is urban, 0 is rural

urban is a dummy variable indicating whether the household is in an urban area.

0 - rural

1 - urban

1.5 Housing

areaspace

Variable Label: Area dwelling in square meters

areaspace is a continuous variable that refers to the total floor area (in square meters) of all rooms and auxiliary premises (kitchen, vestibule, cloakroom, hallway, toilet room, sauna that is within the dwelling, pantry, interstice, bathroom, storeroom, porch, integrated wall closets) in the whole household dwelling unit.

central_acc

Variable Label: Access to central heating

central_acc is a dummy variable that indicates the access to central heating in the dwelling.

Categories after harmonization:

1 - Yes

0 - No

cooksource

Variable Label: Main cooking fuel

cooksource is a categorical variable that identifies the source of cooking. The categories after harmonization are:

1 - Firewood

2 - Kerosene

3 - Charcoal

4 - Electricity

5 - Gas

9 - Other

10 - No cook source

dweltyp

Variable Label: Types of Dwelling

dweltyp is a categorical variable that specifies the type of dwelling. Categories after harmonization are:

1 - Detached house

2 - Multi-family house

3 - Separate apartment

4 - Communal apartment

5 - Room in a larger dwelling

6 - Several buildings connected

7 - Several separate buildings

8 - Improvised housing unit

9 - Other

elec_acc

Variable Label: Connection to electricity in dwelling

elec_acc is a categorical variable that identifies type of connection to electricity. For instance, access to electricity ('Yes') may be public/quasi-public referring to mains electricity (i.e. the term used to refer to the electricity supply from power stations to households) or private referring to electricity from generator or solar or private company. The quality of electricity is assessed by other Tier 3 variables, such as number of electricity hours per day (elechr_acc).

Access to electricity, categories after harmonization:

- 1 - Yes, public/quasi-public
- 2 - Yes, private generator, including shared
- 3 - Yes, solar
- 4 - Yes, source unstated
- 5 - No

elechr_acc

Variable Label: Electricity availability (hr/day)

elechr_acc is a numeric continuous variable that specifies the access to electricity in hours per day.

electricity

Variable Label: Access to electricity in dwelling

electricity is a dummy variable that specifies the access to electricity in the household.

Categories after harmonization:

- 1 - Yes
- 0 - No

electtyp

Variable Label: Type of lighting and/or electricity

electtyp is a categorical variable that specifies the source of energy when cooksource and lightsource variables are not available and there is only one question about the type of energy source in the household; when cooksource and lightsource are available this variable has to be created prioritizing electricity, then Gas, then Lamp. Categories after harmonization:

- 1 - Electricity
- 2 - Gas
- 3 - Lamp
- 4 - Others
- 10 - No cook and light source

floor

Variable Label: Main material used for floor

floor is a categorical variable that indicates type of material used for floors. The floor material is categorized into 3 broad categories namely: Natural, rudimentary and finished. For cases that cannot be covered in the above three categories, please use code 14 = Other – "Specific".

Main source of material used for floors, 14 categories after harmonization:

- 1 - Natural – Earth/sand
- 2 - Natural – Dung
- 3 - Natural – Other
- 4 - Rudimentary – Wood planks
- 5 - Rudimentary – Palm/bamboo
- 6 - Rudimentary – Other
- 7 - Finished – Parquet or polished wood
- 8 - Finished – Vinyl or asphalt strips
- 9 - Finished – Ceramic/marble/granite
- 10 - Finished – Floor tiles/teraso
- 11 - Finished – Cement/red bricks
- 12 - Finished – Carpet
- 13 - Finished – Other
- 14 - Other – Specific

gas

Variable Label: Connection to gas/Usage of gas

gas is a categorical variable that identifies type of gas usage. The categories after harmonization are:

- 0 - No
- 1 - Yes, piped gas (LNG)
- 2 - Yes, bottled gas (LPG)
- 3 - Yes, but don't know

heatsource

Variable Label: Main source of heating

heatsource is a categorical variable that indicates the main source of heating. Main source of heating refers to the type of system used to provide heating for most of the space. It may be central heating covering all or parts of living quarters, or it may not be central, in which case the heating will be provided separately within the living quarters by a stove, fireplace or some other heating body.

As for the energy used for heating purposes, it is closely related to the type of heating and refers to the predominant source of energy, such as solid fuels (coal, lignite, and products of coal and lignite, wood), oils, gaseous fuels (natural or liquefied gas), or electricity.

Main sources of heating, seven categories after harmonization:

- 1 - Firewood
- 2 - Kerosene
- 3 - Charcoal
- 4 - Electricity
- 5 - Gas
- 6 - Central
- 9 - Other
- 10 - No heating

improved_roof

Variable Label: Used improved material for roof

improved_roof is a dummy variable that specifies whether the household has a roof built with improved materials. This is a simplified version of the variable roof in the Global Monitoring Database (GMD) guideline, which has 15 categories:

- 1 = Natural – Thatch/palm leaf
- 2 = Natural – Sod
- 3 = Natural – Other
- 4 = Rudimentary – Rustic mat
- 5 = Rudimentary – Palm/bamboo
- 6 = Rudimentary – Wood planks
- 7 = Rudimentary – Other
- 8 = Finished – Wood
- 9 = Finished – Asbestos
- 10 = Finished – Tile
- 11 = Finished – Concrete
- 12 = Finished – Metal
- 13 = Finished – Roofing shingles
- 14 = Finished – Other
- 15 = Other

We consider an improved_roof = 1 to be the GMD roof Finished categories 8-14, and improved_roof = 0 to be the GMD roof categories 1-7 and 15. Two categories after harmonization:

- 0 - Unimproved roof material
- 1 - Improved roof material

improved_stove

Variable Label: Household has clean cook stove

improved_stove is a dummy variable that specifies whether the household has a clean cook stove. Two categories after harmonization:

- 1 - Yes
- 0 - No

kitchen

Variable Label: Separate kitchen in the dwelling

kitchen is a dummy variable indicating whether the household has a separate kitchen in the dwelling, implying an independent space is set aside for cooking inside the dwelling (kitchen). Any other space reserved for cooking, such as kitchenette or an outer space for kitchen, is not considered as a kitchen. The unit of enumeration for this topic is the housing unit. However, some countries may find it useful to collect information on the availability of kitchen facilities for the use of occupants in collective living quarters, such as hotels, lodging houses, institutions camps and workers' quarters, though people living in these places are generally not captured in a household survey. Two categories after harmonization:

- 1 - Yes

0 - No

lightsource

Variable Label: Main source of lighting

lightsource is a categorical variable that identifies the source of light. The categories after harmonization are:

- 1 - Electricity
- 2 - Kerosene
- 3 - Candles
- 4 - Gas
- 5 - Solar lamps
- 6 - Battery
- 9 - Other
- 10 - No light source

rentin_agriland

Variable Label: Rent in agri land

rentin_agriland is a dummy variable that specifies whether any of the agricultural land a household uses is rented-in land or sharecropped (yes/no). This refers land owned by others (not members of the household) but cultivated or used by the household under fixed rental, sharecropped or similar arrangements. We agree that this question should apply to all households using agricultural land. Two categories after harmonization:

- 1 - Yes
- 0 - No

rooms

Variable Label: Number of habitable rooms

rooms is a integer variable that refers to the number of habitable rooms in the whole household dwelling unit. It may consist of one or more structure(s) (rooms), including all rooms used for living, sleeping and eating. It excludes storerooms, bathrooms, kitchens and rooms used for business or professional purposes. In the case of a one-room dwelling this variable will have the value of one.

wall

Variable Label: Used improved material for external walls

wall is a categorical variable that indicates whether the household used improved material for external walls. This is a simplified version of the variable wall in the Global Monitoring Database (GMD) guideline, which has 19 categories:

- 1 = Natural – Cane/palm/trunks
- 2 = Natural – Dirt
- 3 = Natural – Other
- 4 = Rudimentary – Bamboo with mud
- 5 = Rudimentary – Stone with mud
- 6 = Rudimentary – Uncovered adobe

- 7 = Rudimentary – Plywood
- 8 = Rudimentary – Cardboard
- 9 = Rudimentary – Reused wood
- 10 = Rudimentary – Other
- 11 = Finished – Woven Bamboo
- 12 = Finished – Stone with lime/cement
- 13 = Finished – Cement blocks
- 14 = Finished – Covered adobe
- 15 = Finished – Wood planks/shingles
- 16 = Finished – Plaster wire
- 17 = Finished – GRC/Gypsum/Asbestos
- 18 = Finished – Other
- 19 = Other

We consider an improved_wall = 1 to be the GMD roof Finished categories 11-18, and improved_wall = 0 to be the GMD wall categories 1-10 and 19. Two categories after harmonization:

- 0 - Unimproved wall material
- 1 - Improved wall material

1.6 Non-Labor Income

cash_assistance

Variable Label: Household received cash assistance from government, international agencies, NGOs, religious community, etc. in the last 30 days

cash_assistance is a dummy variable that specifies whether the household received cash assistance from government, international agencies, NGOs, religious community, etc. in the last 30 days. Two categories after harmonization:

- 1 - Yes
- 0 - No

inkind_assistance

Variable Label: Household received in kind assistance from government, international agencies, NGOs, religious community, etc. in the last 30 days

inkind_assistance is a dummy variable that specifies whether the household received in kind assistance from government, international agencies, NGOs, religious community, etc. in the last 30 days. Two categories after harmonization:

- 1 - Yes
- 0 - No

pension

Variable Label: Household received pension benefits in the last 30 days

pension is a dummy variable that specifies whether the household received pension benefits in the last 30 days. Two categories after harmonization:

- 1 - Yes

0 - No

remittances

Variable Label: Household received remittances in the last 30 days

remittances is a dummy variable that specifies whether the household received remittances in the last 30 days. Two categories after harmonization:

1 - Yes

0 - No

1.7 Water and Sanitation-Hygiene (WASH)

imp_san_rec

Variable Label: Improved sanitation facility

imp_san_rec is a dummy variable that categorizes access to improved sanitation facilities based on country-specific estimates or how evidence suggests that the expected error might be minimized. If the relevant survey was on file in the SDG calculations, this would be considered 1 if the majority of the problematic category was estimated therein to be of an improved type at the rural level, and otherwise considered 0. If the survey was not already in the SDG calculations, recommendations are based on the standard international classifications plus any relevant insights from other surveys on file for the specific country. In the few instances where there was no evidence, 0 is used. The recommended access, two categories after harmonization:

1 - Yes

0 - No

imp_wat_rec

Variable Label: Improved water

imp_wat_rec is a dummy variable that categorizes access to improved water sources based on country-specific estimates or how evidence suggests that the expected error might be minimized. If the relevant survey was on file in the SDG calculations, this would be considered 1 if the majority of the problematic category was estimated therein to be of an improved type at the rural level, and otherwise considered 0. If the survey was not already in the SDG calculations, recommendations are based on the standard international classifications plus any relevant insights from other surveys on file for the specific country. In the few instances where there was no evidence, 0 is used. The recommended access, two categories after harmonization:

1 - Yes

0 - No

open_def

Variable Label: Open defecation

open_def is a categorical variable that specifies whether a household has access to any sanitation facility. Access to any sanitation facility, two categories after harmonization:

0 - availability of any facility
1 - no facility, or bush, or field

pipd

Variable Label: Access to piped water

pipd is a categorical variable that indicates whether the household has access to piped water. There are two major types of water supply – within premises and outside premises. 'Within premises' refers to water service piped connection to own tap. It includes both household connection (in-house plumbing) and yard connection (yard or plot outside the house plumbing). Conversely, outside premise refers to a public water point from which people can collect water, shared among houses. It includes public tap and standpipe or a public fountain. Access to piped water, two categories after harmonization:

1 - Yes
0 - No

sanitation_source

Variable Label: Main sanitation facility

sanitation_source is a categorical variable that specifies the source of sanitation facilities. The best possible match is sought, but in many cases the correspondence between country-specific values and these standardized codes is imperfect. Harmonizers should refer to the survey questionnaire to assess the best matches.

Main sanitation source, fourteen categories after harmonization:

1 - A flush toilet
2 - A piped sewer system
3 - A septic tank
4 - Pit latrine
5 - Ventilated improved pit latrine (VIP)
6 - Pit latrine with slab
7 - Composting toilet
8 - Special case
9 - A flush/pour flush to elsewhere
10 - A pit latrine without slab
11 - Bucket
12 - Hanging toilet or hanging latrine
13 - No facilities or bush or field
14 - Other

shared_san_source

Variable Label: Sanitation source is shared with other households

shared_san_source is a dummy variable that specifies if the household sanitation source is shared with other households. Two categories after harmonization:

1 - Yes
0 - No

shared_water_source

Variable Label: Water source is shared with other households

shared_water_source is a dummy variable that specifies if the household water source is shared with other households. Two categories after harmonization:

1 - Yes

0 - No

waste

Variable Label: Main types of solid waste disposal

waste is a categorical variable that indicates the type of solid waste disposal. This variable contains information on the usual manner of collection and disposal of solid waste or garbage generated by occupants of the housing unit. Type of solid waste disposal is categorized by the manner of disposal, such as collection, disposal, burial or compost and by the administrator of the waste disposal, such as authorized collectors, self-appointed collectors, and dump supervised by authorities.

Main types of sewage disposal system, ten categories after harmonization:

- 1 - Solid waste collected on a regular basis by authorized collectors
- 2 - Solid waste collected on an irregular basis by authorized collectors
- 3 - Solid waste collected by self-appointed collectors
- 4 - Occupants dispose of solid waste in a local dump supervised by authorities
- 5 - Occupants dispose of solid waste in a local dump not supervised by authorities
- 6 - Occupants burn solid waste
- 7 - Occupants bury solid waste
- 8 - Occupants dispose solid waste into river, sea, creek, pond
- 9 - Occupants compost solid waste
- 10 - Other arrangement

water_source

Variable Label: Sources of drinking water (14 categories)

water_source is a categorical variable that indicates the main source of drinking water for the household. If the main source of water differs between the wet and dry season, the lowest quality water source across both seasons is preferred. The best possible match is sought, but in many cases the correspondence between country-specific values and these standardized codes is imperfect. Harmonizers should refer to the survey questionnaire to assess the best matches.

Main source of drinking water, fourteen categories after harmonization:

- 1 - Piped water into dwelling
- 2 - Piped water to yard/plot
- 3 - Public tap or standpipe
- 4 - Tube well or borehole
- 5 - Protected dug well
- 6 - Protected spring
- 7 - Bottled water
- 8 - Rainwater
- 9 - Unprotected spring

- 10 - Unprotected dug well
- 11 - Cart with small tank/drum
- 12 - Tanker-truck
- 13 - Surface water
- 14 - Other

watertype_quest

Variable Label: Type of water questions used in the survey

watertype_quest is a categorical variable that specifies the type of water questions used in the survey. This variable is to record the type of question(s) asked about access to water in the survey, for example, the survey had a specific question on the water source on drinking water, or on water source on general water or both.

Type of water question, four categories after harmonization:

- 1 - Drinking water
- 2 - General water
- 3 - Both
- 4 - Other

2. INDIVIDUAL-LEVEL DATA

2.1 Sample

countrycode

Variable Label: Country code

countrycode is a string variable that specifies the 3-character country code used by the World Bank to identify each country. Although there are different naming conventions, it is necessary to use those specified to ensure that the data for each country is appropriately labeled.

countryname

Variable Label: Country name

countryname is a string variable that specifies the country name used by the World Bank to identify each country.

hhid

Variable Label: Household identifier

hhid specifies the unique household identification number in the data file. The original format, string or numeric, of original data should be kept.

int_month

Variable Label: Interview month

int_month is a numeric variable that specifies the month when the survey questionnaire was administered to the household. Categories after harmonization:

1 - January
2 - February
3 - March
4 - April
5 - May
6 - June
7 - July
8 - August
9 - September
10 - October
11 - November
12 - December

int_year

Variable Label: Interview year

int_year is a numeric variable that specifies the year when the survey questionnaire was administered to the household.

pid

Variable Label: Personal identifier

pid allows identification of individuals. Variable will vary in length depending on how the identification code was constructed in each country. Depending on individual countries, this variable may be a concatenation of several variables in the raw data file. Keep format (string or numeric) of original data

rep_year

Variable Label: Reporting year where majority of survey took place

rep_year is a numeric variable that specifies the reporting year where the majority of the survey took place. This is most commonly the year used in the name of the survey.

strata

Variable Label: Strata of the survey

strata refer to the division of the target population – typically the census sample frame -- into subpopulations based on auxiliary information that is known about the full population. Sampling is conducted separately for each strata. The strata should be mutually exclusive: every element in the population must be assigned to only one stratum. The strata should also be collectively exhaustive: no population element can be excluded. Sampling strata need to be considered when constructing the variance (or confidence intervals) of population estimates. strata is needed for the correct calculation of standard deviation for each sample design. A unique identifier is created for each stratum. In Stata, users are advised to specify strata through the svyset command. The variable is in numeric format.

survey_name

Variable Label: Abbreviated survey name (CCC_YYYY_NNNN)

survey_name is a string variable that specifies that abbreviated survey name in the format CCC_YYYY_NNNN, where C = countrycode, Y = rep_year, and N = abbreviated survey name.

surveyprd

Variable Label: Survey period (start-end years)

surveyprd is a string variable that specifies the survey period start and end years. For several surveys, survey period covers two years and this should reflect this.

weight

Variable Label: Household weight

weight contains household weights, typically inversely proportional to the probability of the household being selected for the sample, that should be applied to all analysis to make the results representative of the population.

2.2 Demographics

age

Variable Label: Age of individual (continuous)

age is a numeric variable that refers to the interval of time between the date of birth and the date of the survey.

agecat

Variable Label: Age of individual (categorical)

agecat is a string variable that refers to age groups defined in the survey if information on age is only available in age categories rather than in years. The categories are survey-specific.

birth_country

Variable Label: Country of birth

birth_country is a string variable that specifies the country of birth of the individual. It is only filled when this question is asked at the individual level. It should be filled with the 3-digit ISO-standard country codes (ISO 3166-1 alpha-3).

displaced_forcibly

Variable Label: Person is forcibly displaced

displaced_forcibly is a dummy variable that specifies if the individual is forcibly displaced. Two categories after harmonization:

0 - no

1 - yes

displaced_month

Variable Label: Displacement - month

displaced_month is a categorical variable that specifies the month that the individuals was displaced. Categories after harmonization:

- 1 - January
- 2 - February
- 3 - March
- 4 - April
- 5 - May
- 6 - June
- 7 - July
- 8 - August
- 9 - September
- 10 - October
- 11 - November
- 12 - December

displaced_origin

Variable Label: Displacement - origin

displaced_origin is a categorical variable that specifies whether the individual was displaced from within the country or from another country. Categories after harmonization:

- 1 - within country
- 2 - from another country

displaced_reason

Variable Label: Reason for displacement

displaced_reason is a string variable that specifies why the individual was displaced.

displaced_year

Variable Label: Displacement - year

displaced_year is a numeric variable that specifies the year that the individual was displaced.

female

Variable Label: 1 is female, 0 is male

female is a dummy variable that specifies the sex – male or female – of an individual within a household. While constructing this variable, it is important to make sure that all relevant values are included. Variable values coded as '98' or other numeric characters should be excluded from the values of the 'female' variable. Sex of household member, two categories after harmonization:

- 0 - male
- 1 - female

hasID

Variable Label: Person has some form of ID

hasID is a dummy variable that specifies whether the person has some form of ID. Two categories after harmonization:

- 0 - no
- 1 - yes

idp

Variable Label: Person is displaced within country from locality of origin

idp is a dummy variable that specifies whether the individual is displaced within the country from their locality of origin. Two categories after harmonization:

0 - no

1 - yes

marstat

Variable Label: Marital status

marstat is a categorical variable that refers to the personal status of each individual in relation to the marriage laws or customs of the country. This variable should include at least the following: (a) married; (b) single; (c) living together; (d) divorced/separated; (e) widowed. Categories after harmonization:

1 - Married

2 - Single

3 - living together

4 - divorced/separated

5 - Widow/er

migrant

Variable Label: Person is a migrant - lived in another locality

migrant is a dummy variable that specifies whether the individual previously lived in another locality. Two categories after harmonization:

0 - no

1 - yes

nationality

Variable Label: Citizenship

nationality is a string variable that specifies the country of citizenship of the individual. It is only filled when this question is asked at the individual level. It should be filled with the 3-digit ISO-standard country codes (ISO 3166-1 alpha-3).

reltohead

Variable Label: Relationship to the head of household

reltohead is a categorical variable that indicates a relationship to the reference person of household (usually the head of household). Variable values coded as '98' or other numeric characters should be excluded from the values of reltohead variable. Relationship to head of household, six categories after harmonization:

1 - head

2 - spouse

3 - children

4 - parents

5 - other relative

6 - other non-relative

2.3 Education

dropout_age

Variable Label: Age person dropped out of school

dropout_age is a numeric variable that records the age at which the person dropped out of school. We consider someone a dropout only when they have started at least primary school, are not currently enrolled in school, and have completed less than secondary. It is missing if the person completed secondary school or are currently enrolled.

dropout_primary

Variable Label: Person dropped out of primary school

dropout_primary is a dummy variable that specifies whether the person dropped out of primary school. We consider someone a dropout only when they have started at least primary school, are not currently in school, and have completed less than primary. This variable should be missing for children too young for primary school. Categories after harmonization:

0 - No, completed or currently studying primary school

1 - Yes, started and dropped out of primary school

dropout_reason

Variable Label: Reason for dropping out of school

dropout_reason is a categorical variable specifying the reason for dropping out of school. We consider someone a dropout only when they have started at least primary school, are not currently in school, and have completed less than secondary. Categories after harmonization:

0 - Attending school at any level, or completed secondary school

1 - Considered too young

2 - No money

3 - Has to help at home or family farm or business

4 - No schools or available seats nearby

5 - No documents to enroll, curriculum mismatch, or language barriers

6 - Conflict or outbreak of violence

7 - Safety concerns

8 - Family and social norms

9 - Illness or disability

10 - Other

dropout_secondary

Variable Label: Person dropped out of secondary school

dropout_secondary is a dummy variable that specifies whether the person dropped out of secondary school. The "yes" response included people who are old enough but never attended secondary school. This variable should be missing for children too young for secondary school. Categories after harmonization:

0 - No, completed or currently studying secondary school

1 - Yes, started and dropped out of secondary school

edlev

Variable Label: Highest level of education attained

edlev is a categorical variable that specifies the highest completed education level of the person. This should include currently-enrolled students. Note and use country-specific conditions, such as age>15. Categories after harmonization:

- 1 - Less than primary and illiterate
- 2 - Primary
- 3 - Basic
- 4 - Secondary
- 5 - Tertiary (vocational and university)
- 6 - Other

educ_religious

Variable Label: Has some religious education

educ_religious is a dummy variable that specifies whether the person has had any religious education. Two categories after harmonization:

- 0 - no
- 1 - yes

educat4

Variable Label: Highest level of education completed (4 categories)

educat4 is a categorical variable that specifies the highest completed education level of the person. This should include currently-enrolled students. Secondary is everything from the end of primary to before tertiary (for example, grade 7 through 12). Vocational training is country-specific and will be defined by each region. Note and use country-specific conditions, such as age>15. Categories after harmonization:

- 1 - No education
- 2 - Primary (complete or incomplete)
- 3 - Secondary (complete or incomplete)
- 4 - Tertiary (complete or incomplete)

educat5

Variable Label: Highest level of education completed (5 categories)

educat5 is a categorical variable that specifies the highest completed education level of the person. This should include currently-enrolled students. Secondary is everything from the end of primary to before tertiary (for example, grade 7 through 12). Vocational training is country-specific and will be defined by each region. Note and use country-specific conditions, such as age>15. Categories after harmonization:

- 1 - No education
- 2 - Primary incomplete
- 3 - Primary complete but Secondary incomplete
- 4 - Secondary complete

5 - Tertiary (completed or incomplete)

educat7

Variable Label: Highest level of education completed (7 categories)

educat7 is a categorical variable that specifies the highest completed education level of the person. This should include currently-enrolled students. Secondary is everything from the end of primary to before tertiary (for example, grade 7 through 12). Vocational training is country-specific and will be defined by each region. Note and use country-specific conditions, such as age>15. Categories after harmonization:

- 1 - No education
- 2 - Primary incomplete
- 3 - Primary complete
- 4 - Secondary incomplete
- 5 - Secondary complete
- 6 - Higher than secondary but not university
- 7 - University incomplete or complete

educy

Variable Label: Years of completed education

educy is a numeric variable that records the years of completed education of the individual. This should include currently-enrolled students.

noschool_reason

Variable Label: Reason for never attending school

noschool_reason is a categorical variable that records the reason why the person never attended school. Categories after harmonization:

- 0 - Attended/attending school
- 1 - No money
- 2 - Considered too young

read

Variable Label: Individual can read

read is a dummy variable that specifies whether the individual can read. Each region defines this as they see fit. Two categories after harmonization:

- 1 - Yes
- 0 - No

school

Variable Label: Currently enrolled in or attending school

school is a dummy variable that specifies whether the individual is currently in school. Individuals attending kindergarten or below should be coded as 0 "no". Two categories after harmonization:

- 1 - Yes
- 0 - No

write

Variable Label: Individual can write

read is a dummy variable that specifies whether the individual can write. Each region defines this as they see fit. Two categories after harmonization:

1 - Yes

0 - No

2.4 Labor

empstat

Variable Label: Employment status, primary job (7-day ref period)

empstat is a categorical variable that specifies the main employment status in the last 7 days of any individual with a job (lstatus=1) and is missing otherwise. For this reason, the lower age cutoff (and perhaps upper age cutoff) at which information is collected will vary from country to country.

The definitions are taken from the International Labor Organization's Classification of Status in Employment with some revisions to take into account the data available.

Five categories after harmonization:

1 - Paid Employee

2 - Non-Paid Employee

3 - Employer

4 - Self-employed

5 - Other, workers not classifiable by status

healthins

Variable Label: Health insurance (7-day ref period)

healthins is a dummy variable that classifies the health insurance status (yes/no) of any individual with a job (lstatus=1) and is missing otherwise. Variable is constructed for all persons administered this module in each questionnaire. However, this variable is only constructed if there is an explicit question about health insurance provided by the job. Two categories after harmonization:

1 - Yes

0 - No

industrycat10

Variable Label: 1 digit industry classification, primary job (7-day ref period)

industrycat10 is a categorical variable that specifies the 1-digit industry classification in the last 7 days for the main job of any individual with a job (lstatus=1) and is missing otherwise. The variable is constructed for all individuals that respond to this question, even if they are below the working age. The codes for the main job are given here based on the UN International Standard Industrial Classification. It classifies the main job of any individual with a job (lstatus=1) and is missing otherwise.

Ten categories after harmonization:

- 1 - Agriculture, Hunting, Fishing, etc.
- 2 - Mining
- 3 - Manufacturing
- 4 - Public Utility Services
- 5 - Construction
- 6 - Commerce
- 7 - Transport and Communications
- 8 - Financial and Business Services
- 9 - Public Administration
- 10 - Others Services, Unspecified

industrycat4

Variable Label: 4-category industry classification, primary job (7-day ref period)

industrycat4 is a categorical variable that specifies the 1-digit industry classification in the last 7 days for the main job for Broad Economic Activities. This variable is either created directly from the data (if industry classification does not exist for ten categories) or created from industrycat10.

Four categories after harmonization:

- 1 - Agriculture
- 2 - Industry
- 3 - Services
- 4 - Other

job_2

Variable Label: Has second job

job_2 is a dummy variable that specifies whether the individual has a second job in the last 7 days. It should be created for any individual with a job (lstatus=1) and is missing otherwise. Two categories after harmonization:

- 1 - Yes
- 0 - No

laborincome

Variable Label: Total annual individual labor income in all jobs, incl. bonuses, etc.

laborincome is a continuous variable that specifies the total annualized individual labor income in all jobs including primary, secondary and others regardless of their period of reference. This income includes tips, compensations such as bonuses, dwellings or clothes, and other payments. This variable should be used as the total annual labor income of an individual.

lfstat

Variable Label: Labor force status

lfstat is a categorical variable that specifies the labor force status of the individual. Note and use country-specific conditions, such as age>15. Should be missing for all individuals in a survey if insufficient responses to categorize into codes 1-5. Categories after harmonization:

- 1 - employed

- 2 - unemployed
- 3 - retired
- 4 - student
- 5 - OLF

linc_nc

Variable Label: Total annual wage income in all jobs, excl. bonuses, etc.

linc_nc is a continuous variable that specifies the total annualized wage income in all the jobs including primary, secondary and others regardless of their period of reference. This excludes tips, bonuses, other compensation such as dwellings or clothes, and other payments.

Note: Use gross wages when available and net wages only when gross wages are not available. This is done to make it easy to compare earnings in formal and informal sectors.

lstatus

Variable Label: Labor status (7-day ref period)

lstatus is an individual's labor status in the last 7 days. The value must be missing for individuals less than the required age (minlaborage).

Three categories are used after harmonization:

- 1 = Employed
- 2 = Unemployed
- 3 = Not-in-labor force

All persons are considered active in the labor force if they presently have a job (formal or informal, i.e., employed) or do not have a job but are actively seeking work (i.e., unemployed).

1 = Employed

Employed is defined as anyone who worked during the last 7 days or reference week, regardless of whether the employment was formal or informal, paid or unpaid, for a minimum of 1 hour. Individuals who had a job, but for any reason did not work in the last 7 days are considered employed.

2 = Unemployed

A person is defined as unemployed if he or she is, presently not working but is actively seeking a job. The formal definition of unemployed usually includes being 'able to accept a job.' This last question was asked in a minority of surveys and is, thus, not incorporated in the present definition. A person presently not working but waiting for the start of a new job is considered unemployed.

3 = Not-in-labor force

A person is defined as not-in-labor force if he or she is, presently not working and it is not actively seeking a job during the last 7 days or reference week.

minlaborage

Variable Label: Labor module application age (7-day ref period)

This is the lowest age for which the labor module is implemented in the survey or the minimum working age in the country. For this reason, the lower age cutoff at which information is collected will vary from country to country.

ngo

Variable Label: Employed at an NGO

ngo is a dummy variable that specifies whether the individual is employed at an NGO in the last 7 days for any individual with a job (lstatus=1) and is missing otherwise. Two categories after harmonization:

- 1 - Yes
- 0 - No

njobs

Variable Label: Total number of jobs

njobs is a numeric variable that specifies the total number of jobs for any individual with a job (lstatus = 1).

nlfreason

Variable Label: Reason not in the labor force (7-day ref period)

nlfreason is the reason an individual was not in the labor force in the last 7 days. This variable is constructed for all those who are not presently employed and are not looking for work (lstatus=3) and missing otherwise. Fill this information for all people interviewed in the labor section of the questionnaire regardless of their age.

- 1 - Student
- 2 - Housewife
- 3 - Retired
- 4 - Disabled
- 5 - Don't want to work
- 6 - Lack of resources
- 7 - Moving
- 8 - No jobs available
- 9 - Refugee status
- 10 - Others

occup

Variable Label: 1 digit occupational classification, primary job (7-day ref period)

occup is a categorical variable that specifies the 1-digit occupational classification for the main job in the last 7 days of any individual with a job (lstatus=1) and is missing otherwise. The variable is constructed for all individuals that respond to this question, even if they are below the working age. For this reason, the lower age cutoff (and perhaps upper age cutoff) at which information is collected will vary from country to country. Most surveys collect detailed information and then code it, without keeping the original data, no attempt has been made to correct or check the original coding. The classification is based on the International Standard Classification of Occupations (ISCO). It classifies the main job of any individual with a job (lstatus=1) and is missing otherwise.

Eleven categories after harmonization:

- 1 - Managers
- 2 - Professionals

- 3 - Technicians and associate professionals
- 4 - Clerical support workers
- 5 - Service and sales workers
- 6 - Skilled agricultural, forestry and fishery workers
- 7 - Craft and related trades workers
- 8 - Plant and machine operators, and assemblers
- 9 - Elementary occupations
- 10 - Armed forces occupations
- 99 - Other/unspecified

occup_2

Variable Label: 1 digit occupational classification, secondary job (7-day ref period)
 occup_2 is a categorical variable that specifies the 1-digit occupation classification. It classifies the second job of any individual with a job (lstatus=1) and is missing otherwise. The variable is constructed for all individuals that respond to this question, even if they are below the working age. Most surveys collect detailed information and then code it, without keeping the original data. No attempt has been made to correct or check the original coding. The classification is based on the International Standard Classification of Occupations (ISCO). In the case of different classifications, re-coding has been done to best match the ISCO.

Eleven categories after harmonization:

- 1 - Managers
- 2 - Professionals
- 3 - Technicians and associate professionals
- 4 - Clerical support workers
- 5 - Service and sales workers
- 6 - Skilled agricultural, forestry and fishery workers
- 7 - Craft and related trades workers
- 8 - Plant and machine operators, and assemblers
- 9 - Elementary occupations
- 10 - Armed forces occupations
- 99 - Other/unspecified

ocusec

Variable Label: Sector of activity, primary job (7-day ref period)
 ocusec is a categorical variable that specifies the sector of activity in the last 7 days. It classifies the main job's sector of activity of any individual with a job (lstatus=1) and is missing otherwise. The variable is constructed for all persons administered this module in each questionnaire.

Four categories after harmonization:

- 1. Public Sector, Central Government, Army (including armed forces)
- Public sector is the part of economy run by the government.
- 2 = Private, NGO

Private sector is that part of the economy which is both run for private profit and is not controlled by the state, it also includes non-governmental organizations

- 3 = State-owned enterprises

State-owned includes para-state firms and all others in which the government has control (participation over 50%).

4 = Public or State-owned, but cannot distinguish

Select this option if the questionnaire does not ask for State-owned enterprises, and only for Public sector.

ocusec_2

Variable Label: Sector of activity, secondary job (7-day ref period)

ocusec_2 is a categorical variable that specifies the sector of activity in the last 7 days. It classifies the secondary job's sector of activity of any individual with a job (lstatus=1) and is missing otherwise. The variable is constructed for all persons administered this module in each questionnaire.

Four categories after harmonization:

1. Public Sector, Central Government, Army (including armed forces)

Public sector is the part of economy run by the government.

2 = Private, NGO

Private sector is that part of the economy which is both run for private profit and is not controlled by the state, it also includes non-governmental organizations

3 = State-owned enterprises

State-owned includes para-state firms and all others in which the government has control (participation over 50%).

4 = Public or State-owned, but cannot distinguish

Select this option if the questionnaire does not ask for State-owned enterprises, and only for Public sector.

t_wage_total

Variable Label: Annualized total wage for all jobs (7-day ref period)

t_wage_total is a continuous variable that specifies the total annualized wage income in all jobs including primary, secondary and others. This income includes tips, compensations such as bonuses, dwellings or clothes, and other payments. t_wage_total should be equal to t_wage_nc_total in case there are no bonuses, tips etc. offered as part of any of the jobs. If the number of months worked in this job is missing the harmonizer can assume that the person worked the whole year in this job.

unempl_reason

Variable Label: Reason for challenge in finding work

unempl_reason is a categorical variable that specifies the reason the individual is finding it challenging to find work. This variable should be constructed for those who are unemployed (lstatus = 2) and missing otherwise. Categories after harmonization:

1 - Age

2 - Bribery

3 - Family restrictions

4 - Low wages

5 - No jobs available

- 6 - Underqualified
- 7 - Lack of connections

unempldur_l

Variable Label: Unemployment duration (months) lower bracket (7-day ref period)

unempldur_l is a continuous variable specifying the duration of unemployment in months (lower bracket).

The variable is constructed for all unemployed persons (lstatus=2, otherwise missing). If it is specified as continuous in the survey, it records the numbers of months in unemployment. If the variable is categorical it records the lower boundary of the bracket (ex. "3" if the person is in a category "3-6 months unemployed").

Missing values are allowed for everyone who is not unemployed.

unempldur_u

Variable Label: Unemployment duration (months) upper bracket (7-day ref period)

unempldur_u is a continuous variable specifying the duration of unemployment in months (upper bracket).

The variable is constructed for all unemployed persons (lstatus=2, otherwise missing). If it is specified as continuous in the survey, it records the numbers of months in unemployment. If the variable is categorical it records the upper boundary of the bracket (ex. "6" if the person is in a category "3-6 months unemployed").

Missing values are allowed for everyone who is not unemployed.

If the duration of unemployment is not reported as a range, but as continuous variables, the unempldur_l and unempldur_u variables will have the same value. If the high range is open-ended the unempldur_u variable will be missing (ex. "." if the person is in a category "2 or more years").

want_to_work

Variable Label: Individual is interested in working or setting up a business

want_to_work is a dummy variable that specifies whether an unemployed individual (lstatus = 2) is interested in working or setting up a business. This variable should be constructed only for those who are unemployed (lstatus = 2) and missing otherwise. Two categories after harmonization:

- 1 - Yes
- 0 - No

whours

Variable Label: Hours of work in last week, primary job only (all ref periods)

whours is a continuous variable that specifies the hours of work last week for the main job of any individual with a job (lstatus=1) and is missing otherwise. The main job is defined as that occupation that the person dedicated more time to over the past week (most preferred), month (2nd most preferred), or longer such as a year (3rd most preferred). The variable is constructed for all persons administered this module in each questionnaire. Notes:

- If the respondent was absent from the job in the week preceeding the survey due to holidays, vacation, or sick leave, then record the time worked in the previous 7 days that the person worked.
- Sometimes the questions are phrased as, “on average, how many hours a week do you work?”.
- For individuals who only give information on how many hours they work per day and no information on number of days worked a week, multiply the hours by 5 days.
- In the case of a question that has hours worked per month, divide by 4.3 to get weekly hours.

whours_all

Variable Label: Hours of work in last week, all jobs (all ref periods)

whours_all is a continuous variable that specifies the hours of work last week for all jobs of any individual with a job (lstatus=1) and is missing otherwise. The variable is constructed for all persons administered this module in each questionnaire. Notes:

- If the respondent was absent from the job in the week preceeding the survey due to holidays, vacation, or sick leave, then record the time worked in the previous 7 days that the person worked.
- Sometimes the questions are phrased as, “on average, how many hours a week do you work?”.
- For individuals who only give information on how many hours they work per day and no information on number of days worked a week, multiply the hours by 5 days.
- In the case of a question that has hours worked per month, divide by 4.3 to get weekly hours.