

Ethiopia - COVID-19 High Frequency Phone Survey of Households 2020

World Bank

Report generated on: September 23, 2020

Visit our data catalog at: <https://microdata.fao.org/index.php>

Overview

Identification

ID NUMBER

ETH_2020_HFPS_v01_EN_M_v01_A_OCS

Overview

ABSTRACT

The potential impacts of the COVID-19 pandemic in Ethiopia are expected to be severe on Ethiopian households' welfare. To monitor these impacts on households, the team selected a subsample of households that had been interviewed for the Living Standards Measurement Study (LSMS) in 2019, covering urban and rural areas in all regions of Ethiopia. The 15-minute questionnaire covers a series of topics, such as knowledge of COVID and mitigation measures, access to routine healthcare as public health systems are increasingly under stress, access to educational activities during school closures, employment dynamics, household income and livelihood, income loss and coping strategies, and external assistance.

The survey is implemented using Computer Assisted Telephone Interviewing, using a modular approach, which allows for modules to be dropped and/or added in different waves of the survey. Survey data collection started at the end of April 2020 and households are called back every three to four weeks for a total of seven survey rounds to track the impact of the pandemic as it unfolds and inform government action. This provides data to the government and development partners in near real-time, supporting an evidence-based response to the crisis.

The sample of households was drawn from the sample of households interviewed in the 2018/2019 round of the Ethiopia Socioeconomic Survey (ESS). The extensive information collected in the ESS, less than one year prior to the pandemic, provides a rich set of background information on the COVID-19 High Frequency Phone Survey of households which can be leveraged to assess the differential impacts of the pandemic in the country.

KIND OF DATA

Sample survey data [ssd]

UNITS OF ANALYSIS

Households

Scope

NOTES

The Ethiopia - COVID-19 High Frequency Phone Survey of Households covered the following topics:

- Household Roster (Rounds 1, 2, 3)
- Knowledge Regarding the Spread of COVID-19 (Round 1)
- Behaviour and Social Distancing (Rounds 1 and 3)
- Access to Basic Services (Rounds 1, 2, 3)
- Employment (Rounds 1, 2, 3)
- Income Loss and Coping (Rounds 1, 2, 3)
- Food Security (Rounds 1, 2, 3)
- Aid and Support/ Social Safety Nets (Rounds 1, 2, and 3)
- Agriculture (Round 3)

TOPICS

Topic	Vocabulary	URI
Food (production, crisis)	FAO	
Aid effectiveness	FAO	
Social protection	FAO	
Disaster Risk Management	FAO	

Coverage

GEOGRAPHIC COVERAGE

National

UNIVERSE

The survey covered all de jure households excluding prisons, hospitals, military barracks, and school dormitories.

Producers and Sponsors

PRIMARY INVESTIGATOR(S)

Name	Affiliation
World Bank	

OTHER PRODUCER(S)

Name	Affiliation	Role
Central Statistical Agency		Collaborator

FUNDING

Name	Abbreviation	Role
United States Agency for International Development	USAID	Funded the study
The World Bank Group	WB	Funded the study

Metadata Production

METADATA PRODUCED BY

Name	Abbreviation	Affiliation	Role
Office of Chief Statistician	OCS	Food and Agriculture Organization	Adoption of metadata for FAM
Development Economics Data Group	DECDG	The World Bank	Documentation of the DDI

DDI DOCUMENT VERSION

ETH_2020_HFPS_v01_EN_M_v01_A_OCS_v01

DDI DOCUMENT ID

DDI_ETH_2020_HFPS_v01_EN_M_v01_A_OCS_FAO

Sampling

Sampling Procedure

The sample of the HFPS-HH is a subsample of the 2018-2019 Ethiopia Socioeconomic Survey (ESS). The ESS is built on a nationally and regionally representative sample of households in Ethiopia. ESS 2018-2019 interviewed 6,770 households in urban and rural areas. In the ESS interview, households were asked to provide phone numbers either their own or that of a reference household (i.e. friends or neighbours) so that they can be contacted in the follow-up ESS surveys should they move from their sampled location. At least one valid phone number was obtained for 5,374 households (4,626 owning a phone and 995 with a reference phone number). These households established the sampling frame for the HFPS-HH.

To obtain representative strata at the national, urban, and rural level, the target sample size for the HFPS-HH is 3,300 households; 1,300 in rural and 2,000 households in urban areas. In rural areas, we attempt to call all phone numbers included in the ESS as only 1,413 households owned phones and another 771 households provided reference phone numbers. In urban areas, 3,213 households owned a phone and 224 households provided reference phone numbers. To account for non-response and attrition all the 5,374 households were called in round 1 of the HFPS-HH.

The total number of completed interviews in round one is 3,249 households (978 in rural areas, 2,271 in urban areas). The total number of completed interviews in round two is 3,107 households (940 in rural areas, 2,167 in urban areas). The total number of completed interviews in round three is 3,058 households (934 in rural areas, 2,124 in urban areas).

Weighting

To obtain unbiased estimates from the sample, the information reported by households needs to be adjusted by a sampling weight (or raising factor) “ w_h ”. To construct the sampling weights, we follow the steps outlined in Himelein, K. (2014), which outlines eight steps, of which we follow six, to construct the sampling weights for the HFPS-HH:

1. Begin with base weights from the Ethiopia Socioeconomic Survey ESS 2018/19 for each household
2. Incorporate probability of sub-selection of round 1 unit for each of the phone survey households. We calculate the probability of selection for each of the 20 strata in the ESS (urban and rural in each of the 11 regions except for Addis Ababa where we only have an urban stratum) by creating the numerators as the number of completed phone interviews and the denominator as the number of households in the ESS for each stratum.
3. Pool the weights in Steps 1 and 2.
4. Derive attrition-adjusted weights for all individuals by running a logistic response propensity model based on characteristics of the household head (i.e. education, labour force status, demographic characteristics), characteristics of the household (consumption, assets, financial characteristics), and characteristics of the dwelling (house ownership, overcrowding).
5. Trim weights by replacing the top two percent of observations with the 98th percentile cut-off point; and
6. Post-stratify weights to known population totals to correct for the imbalances across our urban and rural sample. In doing so, we ensure that the distribution in the survey matches the distribution in the ESS.

* Additional technical details and explanations on each of the steps briefly outlined above can be found in Himelein, K. (2014).

Questionnaires

No content available

Data Collection

Data Collection Dates

Start	End	Cycle
2020-04-22	2020-05-13	Round 1
2020-05-14	2020-06-03	Round 2
2020-06-04	2020-06-26	Round 3

Data Collection Mode

Computer Assisted Telephone Interview [cati]

Data Processing

Data Editing

DATA CLEANING

At the end of data collection, the raw dataset was cleaned by the Research team. This included formatting, and correcting results based on monitoring issues, enumerator feedback and survey changes. The details are as follows.

Variable naming and labelling:

- Variable names were changed to reflect the lowercase question name in the paper survey copy, and a word or two related to the question.
- Variables were labelled with longer descriptions of their contents and the full question text was stored in Notes for each variable.
- “Other, specify” variables were named similarly to their related question, with “_other” appended to the name.
- Value labels were assigned where relevant, with options shown in English for all variables, unless preloaded from the roster in Amharic.

Variable formatting:

- Variables were formatted as their object type (string, integer, decimal, time, date, or datetime).
- Multi-select variables were saved both in space-separated single-variables and as multiple binary variables showing the yes/no value of each possible response.
- Time and date variables were stored as POSIX timestamp values and formatted to show Gregorian dates.
- Location information was left in separate ID and Name variables, following the format of the incoming roster. IDs were formatted to include only the variable level digits, and not the higher-level prefixes (2-3 digits only.)
- Full Household and Enumeration Area ID variables were given leading 0s to match incoming roster format.

Observation and variable arrangement:

- Only consented surveys were kept in the dataset, and all personal information and internal survey variables were dropped from the clean dataset.
- Roster data is separated from the main data set and kept in long-form but can be merged on the key variable (key can also be used to merge with the raw data).
- In the main dataset, ii4_resp_id and cs7_hhh_id are the roster IDs of the respondent and household head respectively and can be merged with individual_id in the roster.
- The variables were arranged in the same order as the paper instrument, with observations arranged according to their submission time.

Data Appraisal

Other forms of Data Appraisal

Backcheck data review: Results of the backcheck survey are compared against the originally captured survey results using the `bcstats` command in Stata. This function delivers a comparison of variables and identifies any discrepancies. Any discrepancies identified are then examined individually to determine if they are within reason.