

# Skills Profile Survey 2017, A Refugee and Host Community Survey

**Utz Johann Pape (IBRD - World Bank)**

Report generated on: March 17, 2022

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

## Overview

### Identification

#### ID NUMBER

WBG\_ETH\_2017\_SPS\_v01\_M

### Overview

#### ABSTRACT

The SPS 2017 was conducted in refugee camps and host communities in four regions in Ethiopia. The survey was used to draw a profile for skills and potential opportunities for refugees and host communities to design a better mix of approaches which could help the government in designing livelihood opportunities for these communities. The SPS 2017 contains information on employment, barriers to labor force participation, livelihood structures of refugees before displacement, education and economic conditions as well as access to services, and perceptions. The data combines detailed household questionnaire information with displacement-specific information including drivers of displacement, access to resettlement mechanisms, and return intentions. It also includes comprehensive information on assets and consumption, to allow estimation of poverty based on the Rapid Consumption methodology as detailed in Pape and Mistiaen (2018).

#### KIND OF DATA

Sample survey data [ssd]

#### UNITS OF ANALYSIS

Household and individual.

### Coverage

#### GEOGRAPHIC COVERAGE

The Skills Profile Survey 2017 covered refugees in camps, and surrounding host communities, in Ethiopia. Refugees of four nationalities were surveyed: Eritrean, Somali, South Sudanese, and Sudanese. Only the refugees living in camps were surveyed, because tracing households outside the camps was not feasible. However, 66 percent of all refugees in Ethiopia live in camps, while those that live outside camps are largely Eritrean. Host communities, defined as Ethiopian non-displaced households living within a 5km radius of a camp, were also surveyed.

### Producers and Sponsors

#### PRIMARY INVESTIGATOR(S)

Name	Affiliation
Utz Johann Pape (IBRD - World Bank)	

#### FUNDING

Name	Abbreviation	Role
DfID		

### Metadata Production

#### METADATA PRODUCED BY

Name	Abbreviation	Affiliation	Role
Development Economics Data Group	DECDG	The World Bank	Documentation of the DDI

DATE OF METADATA PRODUCTION  
2019-03-04

DDI DOCUMENT VERSION  
Version 01 (March 2019)

DDI DOCUMENT ID  
DDI\_WBG\_ETH\_2017\_SPS\_v01\_M\_WB

## Sampling

### Sampling Procedure

---

The SPS 2017 is a household survey with a multi-stage stratified random sample. The sampling frame was the list of refugee camps, sites and locations as of January 2017, provided by UNHCR-Ethiopia. The sample consisted of four strata based on four regions: Tigray Afar (Eritrean refugees), Gambella (South Sudanese refugees), Benishangul Gumuz (Sudanese refugees), and Somali (Somali refugees). Each region hosts predominantly one refugee nationality leading to an implicit stratification based on nationality. In each stratum, camps were divided into Enumeration Areas (EAs) of equal size using GIS technology, and 82 EAs were selected per stratum. Within the stratum, the number of EAs per camp was selected proportional to the size of the camp. Within camps, EAs were selected with equal probability. Within each selected EA, all households were listed and then 12 were randomly selected for interview.

The host community sample consisted of the same four regional strata. Within each stratum, areas within a 5km radius of a camp, were divided into EAs of equal size. Of these EAs, those classified as 'residential' by Open Street Maps were used as the sampling frame. 42 EAs were selected per stratum. Within a stratum, EAs were selected using probability proportional to size, with the probability of selection of an EA corresponding to the area of the EA. Within each selected EA, all households were listed and then 12 were randomly selected for interview.

### Deviations from Sample Design

---

Due to security concerns, revisions were made to the sample during fieldwork.

Enumerators in Gambella stratum (hosting South Sudanese refugees) faced repeated security threats and could survey only 439 of the intended 900 refugee households in the region. As the survey team was withdrawn from Gambella region, the host community in Gambella region was not surveyed. The remaining interviews with refugees in Gambella region were substituted by oversampling EAs in Benishangul Gumuz, as 25 percent of the refugee population in this region is South Sudanese. In September 2017, violent conflict in Oromia and Somali regions escalated, rendering some of the camps in Somali stratum inaccessible. The EAs of Jijiga sub-region were replaced by EAs in non-violent areas of Somali stratum. Further, as most refugee camps are in remote areas with sparse host populations, the final number of host households surveyed fell short of the original intended sample of 500 host households per stratum. However, despite the changes in sample, the survey captured roughly similar number of refugee households of the four main refugee nationalities.

### Response Rate

---

NA

### Weighting

---

The sampling weight is the inverse probability of selection. The selection probability for a household can be decomposed into the selection probability of the EA within the stratum, and the selection probability of the household within the EA. For refugees, the selection probability of an EA is calculated as the number of households within the EA divided by the number of households within the stratum, multiplied by the number of selected EAs in the stratum. For the host community, the selection probability of an EA is calculated as the number of EAs selected per stratum divided by the total number of EAs in the stratum. The selection probability for a household within an EA is constant across households, and is calculated as the number of households selected in the EA (usually 12) divided by the total number of households listed in the EA. Sampling weights for refugees were scaled to equal the number of households per strata as per the sampling frame provided by UNHCR. Due to changes in the sample during fieldwork, the number of EAs surveyed in each stratum differed from the original sample. The weights were also scaled to correct for this change.

# Questionnaires

## Overview

---

The questionnaire contains modules on Household Member Roster, Household Characteristics, Food Consumption, Non food consumption, Livestock, Durable Good, Wellbeing and Opinions and Forced Displacement, Movement and Return Intentions. The questionnaire is available for download with the dataset.

## Data Collection

### Data Collection Dates

---

<b>Start</b>	<b>End</b>	<b>Cycle</b>
2017-08-07	2017-11-29	N/A

### Data Collection Mode

---

Computer Assisted Personal Interview [capi]

### Data Collection Notes

---

Due to insecurities at the time of data collection, the sampling design had to be altered, see sampling section.

### Questionnaires

---

The questionnaire contains modules on Household Member Roster, Household Characteristics, Food Consumption, Non food consumption, Livestock, Durable Good, Wellbeing and Opinions and Forced Displacement, Movement and Return Intentions. The questionnaire is available for download with the dataset.

## Data Processing

### Data Editing

---

See accompanying Stata do-files, available under the related materials tab.

### Other Processing

---

The data in the 1-CleanInput folder is the processed and anonymized data, but it has not been cleaned yet. The set of do files provided does the cleaning and the cleaned data is then saved in the 1-CleanOutput folder.

Missing values are coded as:

“.” for missing data

“.a” for “don't know”

“.b” for “Refused to respond”

“.z” for “Not asked due to questionnaire skipping patterns”

# Data Appraisal

No content available