Somali High Frequency Survey Wave 1
Preliminary Results

30th June, 2016

GPV01
AFRICA
Somali High Frequency Survey Wave 1
Preliminary Summary

Overview

The historical civil war and political insecurity in Somalia has resulted in a lack of socioeconomic, perception and other key data in Somalia. The Somalia Socioeconomic Survey 2002 was the last Somalia-wide representative survey. This lack of data makes it difficult for the government and its development partners to plan and implement appropriate policies and programs that are needed to support economic growth and stability. Especially the lack of poverty numbers undermined the development of an interim poverty reduction strategy paper, which is required to apply for HIPC debt relief.

The Somali High Frequency Survey closes this crucial data gap. The first wave of the Somali High Frequency Survey was conducted as part of the Somalia Poverty TA program in February 2016. This summary document describes the methodology and presents a few preliminary findings of the first wave. A more comprehensive compilation of preliminary findings is available in the accompanying presentation. An in-depth analysis has not yet been conducted but will be proposed in the last section of this document.

Methodology

Data collection in Somalia is challenging due to insecurity in some areas. Traditional sampling methodologies require a full listing of enumeration areas, which is impossible in insecure areas. Also face-to-face time is limited to about 60 minutes while a full consumption questionnaire takes 90 to 120 minutes. Finally, limited field access makes monitoring of data quality difficult.

The poverty team developed solutions to overcome these challenges and allow household consumption data collection in Somalia. The new solutions were tested in a pilot survey in Mogadishu. The first challenge related to sampling was resolved by employing a segmentation approach instead of requiring a full listing in insecure areas. The second challenge of limited face-to-face time was overcome by a newly developed methodology to collect consumption data in 60 minutes. An ex-post simulation using data from Hargeisa showed that the methodology is able to provide accurate poverty estimates. The third challenge of monitoring limitations was tackled by the design of a remote real-time data monitoring system. Implementing these innovations in the Somali High Frequency Survey ensured high data quality despite limitations for field monitoring.

The sample design had to be adopted due to missing enumeration area maps. The survey was originally planned to rely on the sample framework provided by PESS. However, maps for a large number of rural enumeration areas were not available. Therefore, the sample design was altered for areas without existing maps. For those areas, settlement data from PESS as well as from UNDP (2005) were used to create a sample frame. The draft sample
frame was cleaned by merging duplicate enumeration areas and by splitting larger settlements into multiple enumeration areas. Boundaries of the enumeration areas were constructed as circles and then transformed into non-overlapping Theissen polygons.¹

The survey covered Somalis living in Mogadishu, in urban and rural areas in Puntland and Somaliland as well as in Internally Displaces Persons (IDPs) settlements. The survey did not include the nomadic population, which presents about one quarter of the Somali population (according to PESS). This ad hoc approach to create the missing sample frame aimed to ensure representativeness of the covered population but has technical limitations, which should be kept in mind when interpreting results.

Table 1: Sample Properties

<table>
<thead>
<tr>
<th></th>
<th>Overall</th>
<th>Mogadishu</th>
<th>Other Urban</th>
<th>Rural</th>
<th>IDP Settlements</th>
<th>Nomads</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Size (Households)</td>
<td>4,117</td>
<td>816</td>
<td>2,048</td>
<td>822</td>
<td>431</td>
<td>0</td>
</tr>
<tr>
<td>Covered Households</td>
<td>923,092</td>
<td>187,246</td>
<td>445,113</td>
<td>88,770</td>
<td>201,963</td>
<td>0</td>
</tr>
<tr>
<td>Sample Size (Individuals)</td>
<td>21,026</td>
<td>3,619</td>
<td>11,123</td>
<td>4,094</td>
<td>2,190</td>
<td>0</td>
</tr>
<tr>
<td>Covered Individuals</td>
<td>4,930,351</td>
<td>895,915</td>
<td>2,459,482</td>
<td>463,266</td>
<td>1,111,689</td>
<td>0</td>
</tr>
<tr>
<td>Population (PESS)</td>
<td>12,316,895</td>
<td>1,280,939</td>
<td>3,935,453</td>
<td>2,806,787</td>
<td>1,106,751</td>
<td>3,186,965</td>
</tr>
<tr>
<td>Population Covered</td>
<td>40%</td>
<td>70%</td>
<td>62%</td>
<td>17%</td>
<td>100%</td>
<td>0%</td>
</tr>
<tr>
<td>Number of Enumeration Areas</td>
<td>341</td>
<td>67</td>
<td>170</td>
<td>69</td>
<td>35</td>
<td>0</td>
</tr>
</tbody>
</table>

The questionnaire of the survey is focused on consumption. Consumption is measured using the Rapid Consumption Methodology. The approach partitions consumption items into core and optional modules. Only the core and one optional module is administered to each household. The missing consumption information is imputed within the survey using multiple imputation techniques. The questionnaire also covered livestock and perception data. In more secure areas, a long-form of the questionnaire was administered including modules for income / remittances, household enterprises and shocks.

Poverty is estimated using the international 1.90 USD 2011 PPP poverty line. As the poverty line is defined in USD 2011 PPP, it must be converted to the currency used to measure consumption in the survey. First, USD 2011 PPP are converted into Somali Shilling in 2011 using the regression-based PPP estimate for Somalia. Second, the change in purchasing

¹ Two overlapping circles become two polygons sharing as boundary the intersection of the two original circles.
power per Somali Shilling is considered by estimating inflation from 2011 to 2016. In the absence of a CPI, the consumption shares of the survey are used together with prices collected from 2011 to 2016 by the Food Security and Nutrition Analysis Unit (FSNAU) Somalia led by FAO. Third, the poverty line is converted back to USD using the current exchange rate. The resulting poverty line is 1.58 USD (2016) per day per person. The consumption estimates from the household are accordingly converted into USD using the region-specific exchange rates collected by the market price surveys. Food consumption is also spatially deflated using a Laspeyres deflator to ensure comparability.

**Preliminary Findings**

Poverty ranges from 35 to 71 percent in Somalia, across different parts of the population. The poverty rate in Mogadishu is similar to other urban areas while rural areas are poorer. Most people in IDP settlements are poor. Household receiving remittances are better off than household that do not receive remittances (Figure 1).

**Figure 1: Poverty incidence (% of population living on less than $1.9 per day in 2011 PPP terms).**

One in four working-age persons participate in the labor market. One in third of those is unemployed (Figure 2). Women are often not participating in the labor market because they engage in household work and/or are not allowed by their husbands to seek work. More than half of the youth pursues education.

**Figure 2: Education and Labor Status**
More than half of the population is literate. Literacy are higher in urban compared to rural areas (Figure 3). Women are slightly less often literate. Wealthier households have higher literacy rates than poorer households.

![Figure 3: Literacy](image)

Somalis are optimistic about the future. Asked about their outlook on living standards and employment opportunities, 4 in 10 people are optimistic while only 2 in 10 people are pessimistic (Figure 4). People living in IDP settlements are generally more pessimistic. Wealthier and female-headed households are more optimistic.

![Figure 4: Outlook on living standards and employment opportunities.](image)

**Next Steps**

The survey data offer the opportunity for an in-depth analysis of poverty and related socio-economic indicators to improve our understanding of poverty in Somalia. This work is planned for FY17 contributing to the preparation of the Systematic Country Diagnostic while also feeding into the Country Economic Memorandum.
Additional waves of the High Frequency Survey will help to understand poverty dynamics and can increase coverage, e.g. by including nomads. The first wave of the survey only gave a snapshot of poverty for the represented part of the population. Additional waves are planned including previously insecure areas as well as the nomadic population. The additional waves will be funded by the Somalia Multi-Partner Knowledge Fund.

The incomplete sampling frame for Somalia will negatively affect representativeness of any surveys in Somalia. Therefore, it is urgent to update the sampling frame for Somalia. Based on the experience gained in complementing the existing sample frame for the first wave of the Survey, it is proposed to support the Government of Somalia with technical assistance to construct a complete sample frame. Due to security constraints to conduct required field work for updating the sampling frame, an approach based on satellite images with selected verification will be considered. Given the reach beyond the High Frequency Survey and the strong capacity building component in this undertaking, the TFSCB could be the appropriate vehicle for this.

Selected References


Somali High Frequency Survey
Wave I (February 2016)

Overview and Preliminary Results

Utz Johann Pape & Johan A. Mistiaen
Global Poverty and Equity Practice
The World Bank
August 30, 2016

Agenda

1. Methodology
   a) Sample
   b) Questionnaire
   c) Rapid Consumption Methodology
   d) Fieldwork Monitoring
   e) Poverty Measurement

2. Results
   a) Demographics
   b) Poverty
   c) Labor
   d) Education
   e) Access to Services
   f) Perceptions

3. Recap of Main Findings

4. Discussion and Next Steps
Sample Design without EA maps

PESS maps were only available for some areas. In areas without PESS maps, we designed a sample frame based on settlements recorded in other data sources.

1. Removal of duplicates
Sample Design without EA maps

PESS maps were only available for some areas. In areas without PESS maps, we designed a sample frame based on settlements recorded in other data sources.

1. Removal of duplicates
2. Demarcate boundaries in urban areas using Thiessen polygons

Sample Design without EA maps

PESS maps were only available for some areas. In areas without PESS maps, we designed a sample frame based on settlements recorded in other data sources.

1. Removal of duplicates
2. Demarcate boundaries in urban areas using Thiessen polygons
3. Check demarcation
Full vs. Micro-listing

Traditionally, all households in an enumeration area are listed before households are selected randomly for interviews. A full listing can raise suspicion in some areas. Thus, we opted for a micro-listing approach.

The micro-listing approach splits an enumeration area into multiple segments. Each segment is further split into blocks. Within a selected block, the enumerator records all housing structures. The tablet selects randomly one structure of which the enumerator records all households. The tablet selected randomly the household to be interviewed.

This methodology provides unbiased estimates but reduces precision due to design effects introduced by the additional layers of hierarchy.

Wave I surveyed 4,117 household across rural and urban areas and IDP settlements, representing 40 percent of the population

Sample properties of the SHFS

<table>
<thead>
<tr>
<th></th>
<th>Overall</th>
<th>Mogadishu</th>
<th>Other Urban</th>
<th>Rural</th>
<th>IDP Settlements</th>
<th>Nomads</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Size (Households)</td>
<td>4,117</td>
<td>816</td>
<td>2,048</td>
<td>822</td>
<td>431</td>
<td>0</td>
</tr>
<tr>
<td>Population (Covered Households)</td>
<td>923,092</td>
<td>187,246</td>
<td>445,113</td>
<td>88,770</td>
<td>201,963</td>
<td>0</td>
</tr>
<tr>
<td>Sample Size (Individuals)</td>
<td>21,026</td>
<td>3,619</td>
<td>11,123</td>
<td>4,094</td>
<td>2,190</td>
<td>0</td>
</tr>
<tr>
<td>Population (Covered Individuals)</td>
<td>4,930,351</td>
<td>895,915</td>
<td>2,459,453</td>
<td>463,266</td>
<td>1,111,689</td>
<td>0</td>
</tr>
<tr>
<td>Population (Total PESS)</td>
<td>12,316,895</td>
<td>1,280,939</td>
<td>3,955,453</td>
<td>2,806,787</td>
<td>1,106,751</td>
<td>3,186,965</td>
</tr>
<tr>
<td>Percentage of Population Covered</td>
<td>40%</td>
<td>70%</td>
<td>62%</td>
<td>17%</td>
<td>100%</td>
<td>0%</td>
</tr>
<tr>
<td>Number of Enumeration Areas</td>
<td>341</td>
<td>67</td>
<td>170</td>
<td>69</td>
<td>35</td>
<td>0</td>
</tr>
</tbody>
</table>

Notes
1. Wave I of the SHFS covered the following pre-war regions: Awdal, Banadir, Bari, Mudug, Nugaal, Sanaag, Sool, Togdheer, and Woqooyi Galbeed. Not included were Bakool, Bay, Galgaduud, Gedo, Hiraan, Lower Juba, Lower Shabelle, Middle Juba, and Middle Shabelle.
2. ‘Covered’ population includes extrapolation to inaccessible areas within covered pre-war regions. It is assumed that inaccessible areas are similar to the bottom 25 percent of enumeration areas in the same analytical strata. IDP settlements are scaled to all regions.
3. Percentage of Population Covered is based on PESS population estimates.
Agenda

1. Methodology
   a) Sample
   b) Questionnaire
   c) Rapid Consumption Methodology
   d) Fieldwork Monitoring
   e) Poverty Measurement

2. Results
   a) Demographics
   b) Poverty
   c) Labor
   d) Education
   e) Access to Services
   f) Perceptions

3. Recap of Main Findings

4. Discussion and Next Steps

Questionnaire – Modules

- Household Roster (110 questions)
- Household Characteristics (38 questions)
- Consumption
  - Food (30 questions per item)
  - Non-Food (14 questions per item)
- Livestock (39 questions per item)
- Durables (16 questions per item)
- Perception (24 questions)
- Food Security* (24 questions)
- Income and Remittances* (14 questions)
- Household Enterprise* (172 questions)
- Shocks* (15 questions)

* Only administered in areas with full listing
Questionnaire – Dataset

- Household: 348 variables
- Household members: 148 variables
- Food: 33 variables
- Non-Food: 18 variables
- Durables: 30 variables
- Livestock: 33 variables
- Shocks: 16 variables

Agenda

1. Methodology
   a) Sample
   b) Questionnaire
   c) Rapid Consumption Methodology
   d) Fieldwork Monitoring
   e) Poverty Measurement
2. Results
   a) Demographics
   b) Poverty
   c) Labor
   d) Education
   e) Access to Services
   f) Perceptions
3. Recap of Main Findings
4. Discussion and Next Steps
Rapid Consumption Methodology

In traditional household surveys, consumption is measured using a long list of >300 items. This takes multiple hours or days.

- **Core**
  - Maize (flour)
  - Maize (grain)
  - Millet (flour)
  - Millet (grain)
  - Apples
  - Bananas
  - Melons
  - Pears
  - Camel milk
  - Cow milk
  - Goat milk
  - Milk powder

- **Optional Module**
  - Vegetables

Full Reduced

- Aggregated
  - Maize
  - Millet
  - Vegetables
  - Milk

- Reduced
  - Maize (flour)
  - Millet (grain)
  - Bananas
  - Melons
  - Cow milk
  - Milk powder

- Full Reduced Aggregated Poverty Line

- Bias

- Scaled BUT scale factor usually unknown

Rapid Consumption Methodology

- **Group 1**
  - Skip
  - 1
  - Core

- **Group 2**
  - 2
  - Skip
  - Core

- **Total Imputed Consumption**
  - 2
  - 1
  - Core

- Items are partitioned into a core and multiple optional modules
- Households are assigned to the core and one optional module
The Rapid Consumption Methodology performs well as simulation results indicate. The simulation uses household consumption survey data and compares indicators based on full consumption with indicators based on *ex ante* implemented Rapid Consumption.

Source: Somaliland Household Survey 2012

---

**Rapid Consumption Methodology – Mogadishu 2015 Pilot**

<table>
<thead>
<tr>
<th></th>
<th>Food Consumption</th>
<th></th>
<th>Non-Food Consumption</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of Items</td>
<td>Share Hargeisa</td>
<td>Share Mogadishu Imputed</td>
<td>Number of Items</td>
</tr>
<tr>
<td>Core</td>
<td>33</td>
<td>91%</td>
<td>64%</td>
<td>54%</td>
</tr>
<tr>
<td>Module 1</td>
<td>19</td>
<td>3%</td>
<td>9%</td>
<td>16%</td>
</tr>
<tr>
<td>Module 2</td>
<td>20</td>
<td>2%</td>
<td>14%</td>
<td>14%</td>
</tr>
<tr>
<td>Module 3</td>
<td>15</td>
<td>2%</td>
<td>5%</td>
<td>6%</td>
</tr>
<tr>
<td>Module 4</td>
<td>15</td>
<td>2%</td>
<td>8%</td>
<td>9%</td>
</tr>
</tbody>
</table>

Only recording consumption from ‘core’ items will result in severe under-estimation of consumption and, thus, over-estimation of poverty.

Source: Somaliland Household Survey 2012 and HFS Mogadishu 2015 Pilot
Agenda

1. Methodology
   a) Sample
   b) Questionnaire
   c) Rapid Consumption Methodology
   d) Fieldwork Monitoring
   e) Poverty Measurement

2. Results
   a) Demographics
   b) Poverty
   c) Labor
   d) Education
   e) Access to Services
   f) Perceptions

3. Recap of Main Findings

4. Discussion and Next Steps

Field Monitoring

Data collection was monitored daily using a real-time monitoring system.
Field Monitoring

Data collection was monitored daily using a real-time monitoring system.

2. Allocation of Optional Modules

<table>
<thead>
<tr>
<th>Sum of Nb of interviews with Treat=1</th>
<th>Sum of Nb of interviews with Treat=2</th>
<th>Sum of Nb of interviews with Treat=3</th>
<th>Sum of Nb of interviews with Treat=4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. Duration of interviews by enumerators in minutes

- Average of Mean duration of interviews - minutes
- Average of Median duration of interviews - minutes
Agenda

1. Methodology
   a) Sample
   b) Questionnaire
   c) Rapid Consumption Methodology
   d) Fieldwork Monitoring
   e) Poverty Measurement

2. Results
   a) Demographics
   b) Poverty
   c) Labor
   d) Education
   e) Access to Services
   f) Perceptions

3. Recap of Main Findings

4. Discussion and Next Steps

Poverty Measurement from the SHFS

1. Consumption Aggregate
   - For food and non-food items
   - For assets by estimating consumption flow
   - Impute ‘missing’ consumption values

2. Deflator
   - Laspeyres: calculate spatial price indices using a common food basket and spatial prices
   - Apply to food and nonfood consumption aggregate

3. Define a Poverty Line based on 1.90 USD PPP 2011
   - Converting 1.90 USD PPP to SSh in 2011
   - Estimate inflation of SSh from 2011 to 2016 by a CPI-like index based on estimated consumption shares and FSNAU price data (food and non-food)
   - Convert poverty line back to current USD using current exchange rate from SSh to USD
   - Resulting poverty line: 1.58 USD (2016)
Agenda

1. Methodology
   a) Sample
   b) Questionnaire
   c) Rapid Consumption Methodology
   d) Fieldwork Monitoring
   e) Poverty Measurement

2. Results
   a) Demographics
   b) Poverty
   c) Labor
   d) Education
   e) Access to Services
   f) Perceptions

3. Recap of Main Findings

4. Discussion and Next Steps

The population is predominantly young
Almost half of the population is less than 15 years old
Almost half of Somali households are headed by women. 2 in 3 households in Mogadishu and IDP Settlements are headed by men.

The average household size is 5.3. Household size decreases with income.
Agenda

1. Methodology
   a) Sample
   b) Questionnaire
   c) Rapid Consumption Methodology
   d) Fieldwork Monitoring
   e) Poverty Measurement

2. Results
   a) Demographics
   b) Poverty
   c) Labor
   d) Education
   e) Access to Services
   f) Perceptions

3. Recap of Main Findings

4. Discussion and Next Steps

The poverty headcount ranges from 35 to 71 percent

[Chart showing poverty headcount rates for different groups, with the overall average indicated by a dashed line.]
The population ranks among the poorest of the world

Poverty incidence (% of population living on less than $1.9 per day in 2011 PPP terms)

The poverty gap ranges from 14 to 39 percent

Poverty gap (% shortfall relative to poverty line)
The distribution of per capita consumption expenditures rises steeply to the poverty line => highly elastic (pro's and con's)
IDPs are the poorest, while urban areas outside Mogadishu are wealthiest along every point of the distribution.

The top 20 percent consume seven times more than the bottom 20 percent.

<table>
<thead>
<tr>
<th>Daily consumption expenditure per capita by consumption quintile (current US$)</th>
<th>Overall</th>
<th>Mogadishu</th>
<th>Other Urban</th>
<th>Rural</th>
<th>IDP Settlements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1 (bottom 20)</td>
<td>0.52</td>
<td>0.52</td>
<td>0.57</td>
<td>0.60</td>
<td>0.49</td>
</tr>
<tr>
<td>Q2</td>
<td>0.94</td>
<td>0.95</td>
<td>0.94</td>
<td>0.92</td>
<td>0.93</td>
</tr>
<tr>
<td>Q3</td>
<td>1.38</td>
<td>1.37</td>
<td>1.38</td>
<td>1.40</td>
<td>1.39</td>
</tr>
<tr>
<td>Q4</td>
<td>2.05</td>
<td>2.02</td>
<td>2.05</td>
<td>2.06</td>
<td>2.11</td>
</tr>
<tr>
<td>Q5 (top 20)</td>
<td>3.77</td>
<td>3.76</td>
<td>3.85</td>
<td>3.53</td>
<td>3.27</td>
</tr>
</tbody>
</table>
A majority of poor households are in urban areas
3 in 10 poor households are in IDP Settlements

Percentage breakdown of the poor population by region

- IDP Settlements: 32%
- Other Urban: 38%
- Rural: 9%
- Mogadishu: 21%

3 in 4 households did not experience hunger in February 2016
Households in IDP Settlements report hunger more often

Experience of hunger in the past 4 weeks

- Never
- Rarely (1-2 times)
- Sometimes (3-10 times)
- Often (more than 10 times)
Agenda

1. Methodology
   a) Sample
   b) Questionnaire
   c) Rapid Consumption Methodology
   d) Fieldwork Monitoring
   e) Poverty Measurement

2. Results
   a) Demographics
   b) Poverty
   c) Labor
   d) Education
   e) Access to Services
   f) Perceptions

3. Recap of Main Findings

4. Discussion and Next Steps

Labor Market Statistics: Key Concepts (I/II)

- The **working-age population** (15 to 64 years) is made up of people who are either inside ('active') or outside of the labor force ('inactive'). The **working-age youth** are those aged between 15 and 24 years.

- The **labor force** is made up of employed and unemployed people.

- **Employed** people are those who are of working-age (15 to 64 years) and engaged in activities producing goods or providing services for at least one hour during the last 7 days. This includes workers who contributed within the family establishment.

- **Unemployed** people are those who are not employed but are looking for work and are available to work.
  - **Long-term unemployed** are those who have been unemployed for at least 12 months.
  - **First-time job-searchers** are those who are currently unemployed looking for work, and have never worked before.

- Those **outside of the labor force** are called ‘inactive’: these are people who are not employed, not looking for work, and/or not available to work.
In contrast to a labor force survey, this survey only asked the main respondent about labor activities for all household members. This can result in under-reporting of activity status, employment and activities to look for work.

1 in 4 working-age persons participate in the labor market. More men than women are inside the labor force, inactivity highest among the youth.

* Youth are defined as the population aged older than 15 and younger than 25
3 in 10 working-aged persons are pursuing education
Among young people (15 – 24 years) more than half are pursuing education

4 in 10 ‘inactive’ women aged 15 and older work in the household
Almost 3 in 10 of ‘inactive’ men are in education
More than half of the labor force is looking for work
Unemployment highest in IDP settlements, long-term unemployment low
Many are looking for work for the first time

Employment and unemployment in detail

Young people (15 to 24 years) are unemployed more often than adults (25 to 64 years)

Adult unemployment and youth unemployment (percentage of the active adult/youth population)
Half of working adults are workers who receive a salary
Women are more often work as own-account workers or contributing family workers

Status in employment

Agenda

1. Methodology
   a) Sample
   b) Questionnaire
   c) Rapid Consumption Methodology
   d) Fieldwork Monitoring
   e) Poverty Measurement

2. Results
   a) Demographics
   b) Poverty
   c) Labor
   d) Education
   e) Access to Services
   f) Perceptions

3. Recap of Main Findings

4. Discussion and Next Steps
More than half of individuals can read and write
Wealthier individuals and residents of urban areas are literate more often

Less than half of the population has no education
Educational attainment is highest in urban areas and among wealthier households
The younger generation (15-29 years) is more educated than the older generations (30+ years)
More than half of children aged 6 to 17 are enrolled in school. Children in non-poor households are enrolled in school more often.

Enrollment increases between ages 6 and 11 indicating that children go to school delayed. 3 in 4 children between 11 and 17 go to school.
Agenda

1. Methodology
   a) Sample
   b) Questionnaire
   c) Rapid Consumption Methodology
   d) Fieldwork Monitoring
   e) Poverty Measurement

2. Results
   a) Demographics
   b) Poverty
   c) Labor
   d) Education
   e) Access to Services
   f) Perceptions

3. Recap of Main Findings

4. Discussion and Next Steps

Location matters more than income level for access to high quality amenities (I)
Location matters more than income level for access to high quality amenities (II)

![Bar chart showing drinking water quality by location and type of water source.]

Agenda

1. Methodology
   a) Sample
   b) Questionnaire
   c) Rapid Consumption Methodology
   d) Fieldwork Monitoring
   e) Poverty Measurement

2. Results
   a) Demographics
   b) Poverty
   c) Labor
   d) Education
   e) Access to Services
   f) Perceptions

3. Recap of Main Findings

4. Discussion and Next Steps
4 in 10 households are optimistic about the future
Households in IDP Settlements are more pessimistic, wealthier households are more optimistic

Most households feel safe
Households in Mogadishu and IDP Settlements feel least safe
Recap of Preliminary Findings

- The population is predominantly young
- 52% of the population covered by the SHFS live in poverty (below $1.9 per day in 2011 PPP terms) ranking as one of world’s poorest countries
- Households in IDP Settlements are most affected by poverty and unemployment
- More than half of the working-age population is ‘outside the labor force’
- Women are more often outside the labor force and working in the household
- More than half of people inside the labor force are unemployed but almost half are optimistic about their labor market prospects
- The youth is better educated but also more often unemployed than adults
- Many Somali children enroll in school delayed
Technical Appendix: Wave 1 Somali HFS

This technical appendix describes sample design, cleaning and construction of consumption aggregates for the Wave 1 Somali High Frequency Survey data.

Introduction
Estimating monetary poverty rates requires a sound, reproducible methodology. The methodology starts with the sample design, continues with questionnaire design and the construction of food and non-food consumption aggregates, selection of spatial price deflators and how to determine the consumption value derived from assets, and what process to use to construct the poverty lines. This appendix describes the methodology used to estimate poverty for the Wave 1 Somali High Frequency Survey.

The chosen methodology balances a trade-off between feasibility and accuracy. Somalia is a fragile country with severe security constraints for field work and widespread displacement. The sampling methodology was adapted to the context by excluding several inaccessible areas. The questionnaire design utilized the Rapid Consumption methodology that can be easily and quickly implemented. The choice of deflators and the poverty line were driven by data quality.

A household is defined as poor if the per-capita household consumption does not exceed a given threshold

\[ y_i \leq z \]

where \( y_i \) is the nominal per-capita household expenditure and \( z \) is the poverty line at the nominal level. In the following, we discuss the selection of households \( i \) as part of the sample design and present the construction of the consumption aggregate \( y_i \) before discussing the choice of the poverty line \( z \) and standard poverty measures.

Sample
The Population Estimation Survey of Somalia (PESS) was used as sample frame alongside a list of settlements from three different sources (UNDP 1997, UNDP 2006 and FSAU 2003) to complement missing rural and semi-urban settlements. The combined sample frame was cleaned and preprocessed before the number of enumeration areas per strata was calculated and enumeration areas selected proportional to size. Depending on the strata, different multi-stage clustering approaches were used to select households.

Sample Frame
Due to the combination of the different data sources, the resulting sample frame included enumeration areas as well as settlements. While enumeration areas are defined as geographical areas with about 50 to 200 households, settlements often are larger areas with a larger population. In fact, all rural and a large fraction of semi-urban enumeration areas and settlements did not have boundaries available but were only defined by a GPS position.

Since PESS is also partially based on the same data sources (especially UNDP 1997 and UNDP 2006) and since some PESS enumeration areas had the same GPS location, several GPS positions were very close of each other and, thus, considered duplicates (Figure 1). Technically, duplicates are defined where the distance between the GPS position is below 75m. In groups with multiple duplicates, the additional criteria was introduced that all GPS positions must have pair-wise distances below 200m to prevent large sequential areas of GPS positions. Duplicates were merged into one ‘hypothetical’ enumeration area with a tag of the number of duplicates. Those duplicate counts were used to position manually midpoints for new enumeration areas around the main duplicate GPS position to ensure that larger settlements have the appropriate number of surrounding enumeration areas.1

In a second step, boundaries of enumeration areas without corresponding shape files were drawn automatically. First, the GPS positions were used as midpoints of circles with a radius of 200m. Overlapping circles were transformed to Thiessen polygons where the line connecting the overlapping points becomes the new boundary. The algorithm was tested for areas where PESS shapefiles were available (Figure 2).

---

1 Note that this was only done for selected duplicate enumeration areas to reduce manual processing.
Sample Stratification and Size

The sample is designed based on predicted statistical precision of consumption as well as cost considerations. Without political implications, the survey stratifies the sample into four zones, A including Mogadishu, B including Garowe, C including Hergeiza and D for Sanaag, Sool and Togdheer. The sample is stratified for each zone into economic/political centers, urban centers, other urban settlements, rural settlements and – if existent – IDP camps. The result are 16 strata (star marks areas where a micro-listing approach was utilized; see below):

- A: Mogadishu*; IDPs*
- B: Garowe; Urban Centers; Other Urban; Rural; IDPs*
- C: Hergeiza; Urban Centers; Other Urban; Rural; IDPs*
- D: Sanaag Urban; Sanaag Rural; Sool Urban; Sool Rural; Togdheer Urban; Togdheer Rural
The sample employs a clustered design with the Primary Sampling Unit (PSU) being the enumeration area. Within each enumeration area, 12 households will be selected for interviews. A larger number of households per enumeration area would only marginally benefit the statistical estimation of indicators. A smaller number of households would result in less than 3 observations for each of the four optional modules capturing consumption data.

A total sample of about 3,800 households is sufficient to obtain consumption estimators with a relative standard error below 1 percent. After rounding the number of enumeration areas ensuring that 12 households per enumeration area, 324 enumeration areas were initially selected. The 324 enumeration areas are first distributed into the 16 strata. The number of enumeration areas per strata is determined by (i) the population of the strata, (ii) the variability of consumption within the strata, and (iii) the requirement of at least two enumeration areas per strata. Strata with larger population and larger variability will need a larger sample to retrieve the same relative standard error as a strata with smaller population and consumption variability (Table 12). Variability is estimated based on previous surveys and a pilot in Mogadishu. The strata for Mogadishu was later amended by an additional 20 enumeration areas to correct against a faulty optional module assignment in the first days of data collection.

**Household Selection**

Depending on the strata, different clustering approaches were used. In strata with more volatile security as well as for IDP camps, a multi-stage cluster design was employed called micro-listing. Each selected enumeration area was divided into multiple segments and each segment was further divided into blocks. A block is defined as a geographical area where an enumerator can see (and list) all households from one location in the center of the block. Within each enumeration area, one segment was randomly selected and within the segment 12 blocks were chosen. In each block, all structures were listed before selecting randomly one structure. Within the selected structure, all households were listed and one household randomly selected for interview. This multi-stage clustering approach reduces the time in the field substantially and contributes to a lower profile of enumerators, which is paramount in fragile areas. In strata less volatile, the complete enumeration area was listed before 12 households were randomly selected for interviews (called full-listing).

**Data Collection and Replacements**

The survey was implemented using tablets as survey devices (CAPI). The data collection system consisted of Samsung Smartphones equipped with SIM cards, mobile data plans, microSD cards (16 GB capacity), and external battery packs. The phones were secured with Android’s native encryption and protected by a password. GPS tracker helped to track all devices using a web interface (www.gps-server.net), Barcode Scanner allowed to use barcodes for the identification of enumerators and a parental control application provided a safe contained working environment for enumerators. Interviews were conducted using SurveyCTO Collect on the tablet with data transmitted to a secure SurveyCTO server in a cloud computing environment.

EAs were replaced if security rendered field work unfeasible (Table 12). Replacements were approved by the project manager. Replacement of households were approved by the supervisor after a total of three unsuccessful visits of the household.

Incoming data is processed to create a raw consistent data set. Interviews with wrongly entered EAs were manually corrected. Interviews conducted outside sampled EAs were discarded. For duplicate submissions, only one record is kept.\(^2\) Sampling weights are added to the final dataset and subsequently anonymized at the strata level. Missing values are recoded into four different types of missing values: (i) genuinely missing values coded as “.”; (ii) respondent indicated “don’t know” coded as “.a”; (iii) respondent refused to respond to the question coded as “.b”; and (iv) missing values due to the questionnaire skipping pattern because the question does not apply to the respondent coded as “.z”.

**Cleaning Process of Submissions**

The total number of interviews submitted through SurveyCTO was 4,590, and the breakdown by zone the following:

- A: 1,060
- B: 1,035
- C: 2,396

\(^2\) Two types of duplicate households are identified. Technical duplicates are defined as duplicate submission of the same interview. They are identified as households with identical GPS data (latitude, longitude and altitude coordinates). Manual duplicates are defined as two interviews conducted with the same household. They are identified by almost identical household rosters. The interview with more information is kept based on manual inspection.
The first step corresponds to a cleaning process identifying general issues and inconsistencies with submissions.

- B: 1 empty household record dropped
- C
  - 3 household records deleted as they were submitted through the web and they were part of a test to monitor scripts before fieldwork
  - 1 submission dropped as it corresponds to a test that a team leader made to check if the GPS of one of his enumerator's phone was working
  - 1 additional household record dropped as it corresponds to an interview completed by the enumerator to check he had the latest version of the questionnaire

Therefore, after making the described adjustment, the number of correct submissions became 4,584, with the following breakdown by region:

- A: 1,069
- B: 1,034
- C: 2,361
- D: 120

The second step excludes submissions from EAs and blocks that were not included in the final sample.

- A: 3 submissions were dropped as they belong to a block that was not included in the final sample
- B: 12 submissions dropped, as they correspond to an EA that was not included in the final sample, since it was a replacement EA that was never executed
- C: 3 interviews dropped because the enumerators selected a wrong EA that had been replaced

Therefore, after making the described adjustment, the number of correct submissions became 4,566, with the following breakdown by region:

- A: 1,066
- B: 1,022
- C: 2,358
- D: 120

The next step was to validate the acceptance of submissions, for which six criteria were defined and interviews were dropped that failed to meet at least one of them:

- The duration of the interview had to exceed a threshold of 30 minutes
  - 26 submissions were excluded because they were completed in 30 minutes or less
- Random sound bites check, including respondent and enumerator voices. This criterion will be assumed to hold if a specific interview was not checked on this criterion.
  - No interview was removed for this reason
- The interview has GPS coordinates and it was conducted within a buffer area of the correspondent EA
  - 5 interviews did not have GPS coordinates; and
  - 5 were also excluded as the GPS coordinates indicate the interview did not took place within the boundaries of the EA
- If the interview was not completed in the first visit, then the household record for the first visit must be valid using the previous criterions (except for the duration), and both household records must contain a matching GPS positions, with a margin of +/- 10 meters
  - 34 interviews were dropped as they corresponded to a second visit, and the record from the previous visit did not exist or was not valid
  - 26 additional submissions were not considered, as the GPS coordinates of the first visit did not match with those of the subsequent visit
- If the interview corresponds to a replacement household, the record of the original household must be valid, except for the duration of the interview: 67 submissions were not considered as the interview corresponded to a replacement household with an inexistent or invalid record for the original household.

- Finally, unsuccessful interviews were discarded; the ones where no one answered the door, there was not a knowledgeable adult present or the respondent did not give permission to continue: 282 submissions were not successful and thus were also excluded.

Therefore, at this point, the dataset had a total number of 4,121 submissions, with the following breakdown by region:

- A: 1,031
- B: 929
- C: 2,045
- D: 116

The final step excludes interviews that were incomplete, and thus have several sections without any single response. 4 households did not have any record in the sections corresponding to food consumption, assets and livestock, and thus they were excluded. Therefore, the final dataset includes a total number of 4,117 complete, valid and successful submissions from valid EA and blocks, with the following breakdown by region:

- A: 1,031
- B: 929
- C: 2,041
- D: 116

**Sampling weights**

This section describes calculation of sample weights for households in the dataset. The sample design was different for some strata due to security volatility. Thus, the methods differ between micro-listing and full-listing. After the sample weights were calculated as described below, they were scaled to the number of households accessible with GPS from the sample frame.

- Full listing: The sample was drawn in a two-stage process for strata 201-204, 301-304 and 1103-1304. Therefore, the weights were calculated based on the sampling probabilities for each sampling stage and for each cluster in the following way:

\[
P_{hij} = P_1 P_2 = \frac{E_{Aj} H_i}{H_j} \frac{HS_i}{HL_i}
\]

such that

\[
P_{hij}: \text{Probability of selecting household } h \text{ in EA } i \text{ of strata } j
\]

\[
P_1: \text{Probability of selecting the EA in stage 1}
\]

\[
P_2: \text{Probability of selecting the household in stage 2}
\]

\[
E_{Aj}: \text{Number of EAs selected in strata } j
\]

\[
H_i: \text{Number of households estimated in the sample frame for EA } i
\]

\[
H_j: \text{Number of households estimated in the sample frame in strata } j
\]

\[
HS_i: \text{Number of households selected in EA } i
\]

\[
HL_i: \text{Number of households listed in EA } i
\]

Therefore, the sample weight for each household corresponds to

\[
w = \frac{1}{P_{hij}}
\]
Micro-listing: In strata 101, 105, 205 and 305, the sample was segmented in blocks within EAs, in addition to the two-stage, stratified cluster sampling design. Therefore, the weights were calculated based on the sampling probabilities for each sampling stage and for each cluster in the following way:

\[ P_{hi} = \frac{EA_j}{H_j} \frac{BS_i}{B_i} \frac{HS_i}{H_i} \]

such that:

- \( P_{hi} \): Probability of selecting household \( h \) in EA \( i \) of strata \( j \)
- \( P_1 \): Probability of selecting the EA
- \( P_2 \): Probability of selecting the Block
- \( P_3 \): Probability of selecting the household
- \( EA_j \): Number of EAs selected in strata \( j \)
- \( H_j \): Number of households estimated in the sample frame for EA \( i \)
- \( H_j \): Number of households estimated in the sample frame in strata \( j \)
- \( BS_i \): Number of blocks selected in EA \( i \)
- \( B_i \): Number of blocks in EA \( i \)
- \( HS_i \): Number of households selected in EA \( i \)
- \( H_i \): Number of households in EA \( i \)

Therefore, the sample weight for each household corresponds to

\[ w = \frac{1}{P_{hi}} \]

Finally, three types of sampling weights were estimated:

1) Unadjusted weights: Considers all submissions (4,117) and scales the weights so that the sum of the sampling weights by analytical strata matches the total number of accessible households with GPS according to sample frame.

2) Adjusted weights: Considers all submissions (4,117) and scales the weights uniformly so that the sum of the weights by analytical strata matches the total number of households according to the PESS (Table 3).4

3) Adjusted weights for consumption and poverty variables: Considers only submissions with consumption data (excludes 53 submissions with missing values in the consumption of food, non-food and durables) and adjusts the weights of the remaining 4,064 submissions according to the following scenarios:

- If the number of accessible households with GPS (i.e. the sum of weights) is larger than the total number of households according to PESS by analytical strata, then the weights were scaled downwards uniformly to match the total number of households from PESS, which already reflects the re-allocation of the weights from the 53 submissions excluded.

- If the number of accessible households with GPS (i.e. the sum of weights) is smaller than the total number of households according to PESS, then the weights were scaled upwards in two steps: i) re-allocating uniformly the weights from the 53 households excluded across the 4,064 submissions; and then ii) assigning the additional weights needed to match the figures from PESS only to those households or submissions in the bottom 25 percent of the total consumption distribution for the respective analytical strata. The bottom 25 percent were taking up the weight of the additional households to reflect the fact that excluded enumeration areas were not randomly chosen but differed from other enumeration areas by

---

1 The segmentation step cancels out as exactly one segment is chosen.
2 Usually, the household number from the sample frame should reflect the number of households from the last Census. However, the incomplete sample frame necessitated using different (overlapping) data sources for the sample frame. While the probabilities for selection for duplicates are adjusted for already in the EA selection step, the total number of households did not automatically sum up to the number of households from PESS.
inaccessibility due to security and/or infrastructure. As those enumeration areas are expected to be more deprived than the average enumeration area, they were assumed to be similar to the bottom 25 percent.

<table>
<thead>
<tr>
<th>PESS Region and Analytical Strata</th>
<th>Number of Households</th>
</tr>
</thead>
<tbody>
<tr>
<td>All IDPs</td>
<td>201,963</td>
</tr>
<tr>
<td>Banadir Urban</td>
<td>187,246</td>
</tr>
<tr>
<td>Nugaal Urban</td>
<td>23,119</td>
</tr>
<tr>
<td>Bari and Mudug Urban</td>
<td>140,334</td>
</tr>
<tr>
<td>Woqooyi Galbeed Urban</td>
<td>123,390</td>
</tr>
<tr>
<td>Awdal, Sanaag, Sool and Togdheer Rural</td>
<td>158,279</td>
</tr>
<tr>
<td>Bari, Mudug and Nugaal Rural</td>
<td>27,684</td>
</tr>
<tr>
<td>Awdal, Sanaag, Sool, Togdheer and Woqooyi Galbeed Rural</td>
<td>61,086</td>
</tr>
</tbody>
</table>

### Consumption Aggregate

The nominal household consumption aggregate is the sum of three components, namely 1) expenditures on food items, 2) expenditures on non-food items, and 3) the value of the consumption flow from durable goods:

\[
y_t = y_t^f + y_t^n + y_t^d
\]

This section describes in detail the cleaning of the recorded data for each of the three components. Subsequently, the construction of the consumption aggregate using the Rapid Consumption Methodology is explained as well as the estimation of the consumption flow for durables and the details on the deflator used to calculate spatial price indices.

Moreover, 53 households were assigned a missing value in consumption since 52 of them reported not consuming any food items, and 1 household only reported consuming a non-core food item.

### Cleaning

#### Food

Food expenditure data is cleaned in a four-step process. First, units for reported quantities of consumption and purchase are corrected. Typical mistakes include recorded consumption of 100 kg of a product (like salt) where the correct quantity is grams. These mistakes are corrected using generic rules (Table 5). Then, we introduce a conversion factor to kg for some specific items and units. For example, we recognize that a small piece of bread must have a different weight than a small piece of garlic (Table 6: ). The third step consists of correcting issues with the exchange rate selected (Table 7). Finally, outliers are detected using the six cleaning rules below to correct quantities and prices.

- **Rule 1**
  - Consumption quantities with missing values for items reported as consumed were replaced with item-specific median consumption quantities.
  - Missing purchase quantities and missing prices for items consumed were replaced with item-specific median purchase quantity and item-specific median purchase price.

- **Rule 2:** Records where the respondent did not know or refused to respond if the household had consumed the item, were replaced with the mean value, including non-consumed records.

- **Rule 3:** Records with the same value for quantity consumed or quantity purchased and price are assumed to have a data entry error in the price or quantity and are replaced with the item-specific medians.

- **Rule 4:** Records that have the same value in quantity consumed and quantity purchased but different units are assumed to have a wrong unit either for consumption or purchase. For both quantities, the item-specific distribution of quantities in
kg is calculated to determine the deviation of the entered figure from the median of the distribution. The unit of the quantity that is further away from the median is corrected with the unit of the quantity closer to the median.

- Rule 5:
  - Missing and zero prices are replaced with item-specific medians
  - Outliers for unit prices were identified and replaced with the item-specific median. This includes unit prices in the top 10 percent of the overall cumulative distribution (considering all items), and unit prices below 0.07 USD.

- Rule 6: the consumption value in USD was truncated to the mean plus 3 times the standard deviation of the cumulative distribution for each item, if the record exceeded this threshold.

All medians are estimated at the EA level if a minimum of 5 observations are available excluding previously tagged records. If the minimum number of observations is not met, medians are estimated at the strata-level before proceeding to the survey level. In addition, medians greater than 20 kg and smaller than 0.02 kg were not considered for quantities, while medians greater than 20 USD and smaller than 0.005 USD were also excluded for unit prices.

Non-Food
The non-food dataset only contains values without quantities and units. First, we apply the same cleaning rules for currencies (Table 7), and then the following cleaning rules:

- Rule 1: Zero, missing prices and missing currency for purchased items are replaced with item-specific medians.
- Rule 2: Records where the respondent did not know or refused to respond if the household had purchased the item, were replaced with the mean value, including non-consumed records.
- Rule 3: Prices that are beyond a specific threshold for each recall period (Table 8) are replaced with item-specific medians.
- Rule 4: Prices below the 1 percent and above the 95 percent of the cumulative distribution for each item are replaced with item-specific medians.
- Rule 5: the purchase value in USD was truncated to the mean plus 3 times the standard deviation of the cumulative distribution for each item, if the record exceeded this threshold.

The item-specific medians were applied at the EA, strata and survey level as described above.

Durables
For durables, we also apply the same cleaning rules for currencies (Table 7), and then the following cleaning rules:

- Rule 1: Vintages with missing values and greater than 10 years are replaced with item-specific medians.
- Rule 2: Current and purchase prices equal to zero are replaced with item-specific medians.
- Rule 3: Records that have the same figure in current value and purchase price are incorrect. For both, the item-vintage-specific distribution is calculated to determine the deviation of the entered figure from the median. The one that is further away from that median is corrected with the item-year-specific median value.
- Rule 4: Depreciation rates are replaced by the item-specific medians in the following cases:
  - Negative records
  - Depreciation rates in the top 10 percent and vintage of one year
  - Depreciation rates in the bottom 10 percent and a vintage greater or equal to 3 years
- Rule 5: Records with 100 items or more, and those that reported to own a durable good but did not report the number were replaced with the item-specific medians of consumption in USD.

- Rule 6: Consumption in the top and bottom 1 percent of the overall distribution were replaced with item-specific medians.

- Rule 7: Records where the respondent did not know or refused to respond if the household owned the asset, were replaced with the mean consumption value, including non-consumed records.

- Rule 8: The consumption value in USD was truncated to the mean plus 3 times the standard deviation of the cumulative distribution for each item, if the record exceeded this threshold.

All medians are estimated at the EA level if a minimum of 3 observations are available excluding previously tagged records. If the minimum number of observations is not met, medians are estimated at the strata-level before proceeding to the survey level. Table 9 contains a general overview of consumption of durables, while Table 10 presents the details by item. Table 11 shows the median depreciation rate by durable good.

Rapid Consumption Methodology: Food and Non-Food Aggregates

The survey used the new Rapid Consumption methodology to estimate consumption. A detailed description including an ex post assessment of the methodology is available in a separate document. The rapid survey consumption methodology consists of five main steps. First, core items are selected based on their importance for consumption. Second, the remaining items are partitioned into optional modules. Third, optional modules are assigned to groups of households. After data collection, fourth, consumption of optional modules is imputed for all households. Fifth, the resulting consumption aggregate is used to estimate poverty indicators.

First, core consumption items are selected. Consumption in a country bears some variability but usually a small number of a few dozen items captures the majority of consumption. These items are assigned to the core module, which will be administered to all households. Important items can be identified by its average food share per household or across households. Previous consumption surveys in the same country or consumption shares of neighboring / similar countries can be used to estimate food shares. In the worst case, a random assignment results in a larger standard error but does not introduce a bias.

Second, non-core items are partitioned into optional modules. Different methods can be used for the partitioning into optional modules. In the simplest case, the remaining items are ordered according to their food share and assigned one-by-one while iterating over the optional module in each step. A more sophisticated method takes into account correlation between items and partitions them into orthogonal sets per module. This leads to high correlation between modules supporting the total consumption estimation. Conceptual division into core and optional items is not reflected in the layout of the questionnaire. Rather, all items per household will be grouped into categories of consumption items (like cereals) and different recall periods. Using CAPI, it is straightforward to hide the modular structure from the enumerator.

Third, optional modules will be assigned to groups of households. Assignment of optional modules will be performed randomly stratified by enumeration areas to ensure appropriate representation of optional modules in each enumeration area. This step is followed by the actual data collection.

Fourth, household consumption will be estimated by imputation. The average consumption of each optional module can be estimated based on the sub-sample of households assigned to the optional module. In the simplest case, a simple average can be estimated. More sophisticated techniques can employ a welfare model based on household characteristics and consumption of the core items. The results presented in this note uses a multiple imputation technique based on a multi-variate normal approximation.

Next, the methodology is formalized and assessed using an ex post simulation based on the consumption data from Hergeiza using the Somaliland 2012 household survey (SHS12). Food and non-food consumption for household $i$ are estimated by the sum of expenditures for a set of items.

---


6 As shown later, the assignment of items to modules is very robust and, thus, even rough estimates of consumption shares are sufficient to inform the assignment without requiring a baseline survey.
\[ y_i^f = \sum_{j=1}^{m} y_{ij}^f \quad \text{and} \quad y_i^n = \sum_{j=1}^{m} y_{ij}^n \]

where \( y_i^f \) and \( y_i^n \) denote the food and non-food consumption of item \( j \) in household \( i \). As the estimation for food and non-food consumption follows the same principles, we neglect the upper index \( f \) and \( n \) in the remainder of this section. The list of items can be partitioned into \( M+1 \) modules each with \( m_k \) items:

\[ y_i = \sum_{k=0}^{M} y_{i(k)} \quad \text{with} \quad y_{i(k)} = \sum_{j=1}^{m_k} y_{ikj} \]

For each household, only the core module \( y_{i(0)} \) and one additional optional module \( y_{i(k^*)} \) are collected.

The item assignment to the modules is based on the SHS12 survey with manual modifications especially to treat ‘other’ items correctly. The core module was designed to maximize its consumption share resulting in 91 percent and 76 percent of food respectively non-food consumption captures in the core modules (based on SHS12 consumption; Table 2). Optional modules are constructed using an algorithm to assign items iteratively to optional modules so that items are orthogonal within modules and correlated between modules. In each step, an unassigned item with highest consumption share is selected. For each module, total per capita consumption is regressed on household size, the consumption of all assigned items to this module as well as the new unassigned item. The item will be assigned to the module with the highest increase in the R2 relative to the regression excluding the new unassigned item. The sequenced assignment of items based on their consumption share can lead to considerable differences in the captured consumption share across optional modules. Therefore, a parameter is introduced ensuring that in each step of the assignment procedure the difference in the number of assigned items per module does not exceed \( d \). Using \( d=1 \) assigns items to modules (almost) maximizing equal consumption share across modules. Increasing \( d \) puts increasing weight on orthogonality within and correlation between modules. The parameter was set to \( d=3 \) balancing the two objectives.

In each enumeration area, 12 households were interviewed with an ideal partition of three items per optional module. The assignment of optional modules must ensure that a sufficient number of households are assigned to each optional module. Household consumption was then estimated using the core module, the assigned module and estimates for the remaining optional modules

\[ \hat{y}_i = y_{i(0)} + y_{i(k^*)} + \sum_{k \in K^*} \hat{y}_{i(k)} \]

where \( K^* := \{1, \ldots, k^* - 1, k^* + 1, \ldots, M\} \) denotes the set of non-assigned optional modules. Consumption of non-assigned optional modules is estimated using multiple imputation techniques taking into account the variation absorbed in the residual term.

Multiple imputation was implemented using multi-variate normal regression based on an EM-like algorithm to iteratively estimate model parameters and missing data. This technique is guaranteed to converge in distribution to the optimal values. An EM algorithm draws missing data from a prior (often non-informative) distribution and runs an OLS to estimate the coefficients. Iteratively, the coefficients are updated based on re-estimation using imputed values for missing data drawn from the posterior distribution of the model. The implemented technique employs a Data-Augmentation (DA) algorithm, which is similar to an EM algorithm but updates parameters in a non-deterministic fashion unlike the EM algorithm. Thus, coefficients are drawn from the parameter posterior distribution rather than chosen by likelihood maximization. Hence, the iterative process is a Monte-Carlo Markov–Chain (MCMC) in the parameter space with convergence to the stationary distribution that averages over the missing data. The distribution for the

---

7 Items ‘other’ are often found to capture remaining items for a food category. Using the Rapid Consumption Methodology, this creates problems as ‘other’ will include different items depending on which optional module is administered. This can lead to double-counting after the imputation. Therefore, ‘other’ items are re-formulated and carefully assigned so that double counting cannot occur.

8 Even with \( d=1 \), equal consumption share across modules is not maximized because among the modules with the same number of assigned items, the new item will be assigned to the module it’s most orthogonal to; rather than to the module with lowest consumption share.

9 Field work implementation aimed to achieve a balanced partition among optional modules but due to challenges in following the protocol exactly some enumeration areas are not completely balanced.
missing data stabilizes at the exact distribution to be drawn from to retrieve model estimates averaging over the missing value distribution. The DA algorithm usually converges considerably faster than using standard EM algorithms:

\[ \hat{y}_i^{(k)} = \beta_0^{(k)} y_i^{(0)} + x_i^T \beta^{(k)} + u_i^{(k)} \]

The performance of the estimation technique was assessed based on an *ex post* simulation using the Hergeiza data from SHS12 and mimicking the Rapid Consumption methodology by masking consumption of items that were not administered to households. The results of the simulation were compared with the estimates using the full consumption from SHS12 as reference. The simulation results distinguish between different levels of aggregation to estimate consumption.10 The methodology generally does not perform well at the household level (HH) but improves considerably already at the enumeration area level (EA) where the average of 12 households is estimated. At the national aggregation level, the Rapid Consumption methodology slightly over-estimates consumption by 0.3 percent. Assessing the standard poverty measures including poverty headcount (FGT0), poverty depth (FGT1) and poverty severity (FGT2), the simulation results show that the Rapid Consumption methodology retrieves estimates within 1.5 percent of the reference measure (Figure 3). Generally, the estimates are robust as suggested by the low standard errors (Figure 4).

![Figure 3: Relative bias of simulation results using Rapid Consumption estimation.](source)

![Figure 4: Relative standard error of simulation results using Rapid Consumption estimation.](source)

| Source: Authors' own calculations based on SHS12 data. | Source: Authors' own calculations based on SHS12 data. |

### Table 2: Item partitions based on SHS12 and pilot in Mogadishu.

<table>
<thead>
<tr>
<th>Food Items</th>
<th>Non-food Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of</td>
<td>Share</td>
</tr>
<tr>
<td>Core</td>
<td>Items</td>
</tr>
<tr>
<td>33</td>
<td>91%</td>
</tr>
<tr>
<td>Module 1</td>
<td>19</td>
</tr>
<tr>
<td>Module 2</td>
<td>20</td>
</tr>
<tr>
<td>Module 3</td>
<td>15</td>
</tr>
<tr>
<td>Module 4</td>
<td>15</td>
</tr>
</tbody>
</table>

Source: Authors' own calculations based on SHS12 and Mogadishu Pilot data.

**Durable consumption flow**

The consumption aggregate includes the consumption flow of durables calculated based on the user-cost approach. The consumption flow distributes the consumption value of the durable over multiple years. The user-cost principle defines the consumption flow of an item as the difference of selling the asset at the beginning and the end of the year as this is the opportunity

---

10 The performance of the estimation techniques is presented using the relative bias (mean of the error distribution) and the relative standard error. The relative error is defined as the percentage difference of the estimated consumption and the reference consumption (based on the full consumption module, averaged over all imputations). The relative bias is the average of the relative error. The relative standard error is the standard deviation of the relative error. The simulation is run over different household-module assignments while ensuring that each optional module is assigned equally often to a household per enumeration. The relative bias and the relative standard error are reported across all simulations.
cost of the household for keeping the item. The opportunity cost is composed of the difference in the sales price and the forgone earnings on interest if the asset is sold at the beginning of the year.

If the durable item is sold at the beginning of the year, the household would receive the market price $p_t$ for the item and the interest on the revenue for one year. With $i_t$ denoting the interest rate, the value of the item thus is $p_t (1 + i_t)$. If the item is sold at the end of the year, the household will receive the depreciated value of the item while considering inflation. With $\pi_t$ being the inflation rate during the year, the household would obtain $p_t (1 + \pi_t) (1 - \delta)$ with the annual physical or technological depreciation rate denoted as $\delta$ assumed constant over time. The difference between these two values is the cost that the household is willing to pay for using the durable good for one year. Hence, the consumption flow is:

$$(3) \quad y^d = p_t (1 + i_t) - p_t (1 + \pi_t) (1 - \delta)$$

By assuming that $\delta \times \pi_t \equiv 0$, the equation simplifies to

$$(4) \quad y^d = p_t (i_t - \pi_t + \delta) = p_t (r_t + \delta)$$

where $r_t$ is the real market interest rate in period $t$. Therefore, the consumption flow of an item can be estimated by the current market value $p_t$, the current real interest rate $r_t$, and the depreciation rate $\delta$. Assuming an average annual inflation rate $\pi$, the depreciation rates $\delta$ can be estimated utilizing its relationship to the market price $12$:

$$(5) \quad p_t = p_{t-k} (1 + \pi)^k (1 - \delta)^k$$

The equation can be solved for $\delta$ obtaining:

$$(6) \quad \delta = 1 - \left( \frac{p_t}{p_{t-k}} \right)^\frac{1}{k} \frac{1}{(1 + \pi)}$$

Based on this equation, item-specific median depreciation rates are estimated assuming an inflation rate of 0.5 percent, a nominal interest rate of 2.0 percent and, thus, a real interest rate of 1.5 percent (Table 11).

For all households owning a durable but did not report the current value of the durable, the item-specific median consumption flow is used. For households that own more than one of the durable, the consumption flow of the newest item is added to the item-specific median of the consumption flow times the number of those items without counting the newest item. $13$

**Deflator**

Prices fluctuate considerably between regions, thus we calculated spatial price indices using a common food basket and spatial prices to make consumption comparable across regions. The Laspeyres index is chosen as a deflator due to its moderate data requirements. The deflator is calculated by analytical strata areas based on the price data collected by the HFS.

The Laspeyres index reflects the item-weighted relative price differences across products. Item weights are estimated as household-weighted average consumption share across all households before imputation. Based on the democratic approach, consumption shares are calculated at the household level. Core items use total household core consumption as reference while items from optional modules use the total assigned optional module household consumption as reference. The shares are aggregated at the national level (using household weights) and then calibrated by average consumption per module to arrive at item-weights summing to 1. The item-weights are applied to the relative differences of median item prices for each analytical strata. Missing prices are replaced by the item-specific median over all households. A large Laspeyres indicates a high price level deflating consumption stronger than a lower Laspeyres index. The resulting indices show the fluctuation of prices across regions (Table 4).

**Table 3: Laspeyres Deflators by Analytical Strata**

<table>
<thead>
<tr>
<th>Analytical Strata</th>
<th>Deflator</th>
</tr>
</thead>
</table>

$11$ Assuming a constant depreciation rate is equivalent to assuming a “radioactive decay” of durable goods (see Deaton and Zaidi, 2002).

$12$ In particular $\pi$ solves the equation $\prod_{t=0}^{k} (1 + \pi_t) = (1 + \pi)^k$

$13$ The 2016 HFS questionnaire provides information on a) the year of purchase and b) the purchasing price only for the most recent durable owned by the household.
<table>
<thead>
<tr>
<th>Area</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All IDPs</td>
<td>0.923</td>
</tr>
<tr>
<td>Mogadishu</td>
<td>0.964</td>
</tr>
<tr>
<td>Garowe</td>
<td>0.862</td>
</tr>
<tr>
<td>Urban Bari and Mudug</td>
<td>1.107</td>
</tr>
<tr>
<td>Hergeiza</td>
<td>1.133</td>
</tr>
<tr>
<td>Urban Awdal, Sanaag, Sool and Togdheer</td>
<td>0.922</td>
</tr>
<tr>
<td>Rural Bari, Mudug and Nugaal</td>
<td>1.013</td>
</tr>
<tr>
<td>Rural Awdal, Sanaag, Sool, Togdheer and Woqooyi Galbeed</td>
<td>1.075</td>
</tr>
</tbody>
</table>
## Tables for Cleaning Rules

### Table 4: Summary of Unit Cleaning Rules for Food Items

<table>
<thead>
<tr>
<th>Unit</th>
<th>Condition</th>
<th>Correction</th>
<th>Affected Records$^{14}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>250 ml tin</td>
<td>&lt;=0.03</td>
<td>Multiply by 4</td>
<td>2; 39</td>
</tr>
<tr>
<td>Animal back, ribs, shoulder, thigh,</td>
<td>&gt;=7</td>
<td>Divide by 10</td>
<td>4; 35</td>
</tr>
<tr>
<td>head or leg</td>
<td>2; 39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basket or Dengu (2 kg)</td>
<td>&gt;=10</td>
<td>Divide by 10</td>
<td>1,004; 20</td>
</tr>
<tr>
<td>Bottle (1 kg)</td>
<td>&gt;=10</td>
<td>Divide by 10</td>
<td>473; 281</td>
</tr>
<tr>
<td>Cup (200 g)</td>
<td>&gt;=200</td>
<td>Divide by 2</td>
<td>447; 24</td>
</tr>
<tr>
<td>Farnasilad (12kg)</td>
<td>&gt;12</td>
<td>Divide by 12</td>
<td>544; 60</td>
</tr>
<tr>
<td>Gram (if item corresponds to a spice)</td>
<td>&lt;1</td>
<td>Multiply by 100</td>
<td>115; 5</td>
</tr>
<tr>
<td>Gram (if item does not corresponds to</td>
<td>&gt;25</td>
<td>Divide by 25</td>
<td>357; 921</td>
</tr>
<tr>
<td>a spice</td>
<td>&gt;=25</td>
<td>Divide by 25</td>
<td></td>
</tr>
<tr>
<td>Haa (25 kg)</td>
<td>&gt;25</td>
<td>Divide by 25</td>
<td>357; 921</td>
</tr>
<tr>
<td>Heap (700g)</td>
<td>&gt;=0.69</td>
<td>Divide by 7</td>
<td>182; 11</td>
</tr>
<tr>
<td>Kilogram</td>
<td>&gt;100</td>
<td>Divide by 1,000</td>
<td>68; 4</td>
</tr>
<tr>
<td>Large bag (50 kg)</td>
<td>&gt;=50</td>
<td>Divide by 50</td>
<td>1; 27</td>
</tr>
<tr>
<td>Liter</td>
<td>&gt;10</td>
<td>Divide by 10</td>
<td>3; 32</td>
</tr>
<tr>
<td>Madal/Nus kilo ruba (0.75kg)</td>
<td>&gt;=7.5</td>
<td>Divide by 10</td>
<td>849; 20</td>
</tr>
<tr>
<td>Meals (300 g)</td>
<td>&gt;2.1</td>
<td>Divide by 10</td>
<td>366; 208</td>
</tr>
<tr>
<td>Packet sealed box/container (500 g)</td>
<td>&gt;=5</td>
<td>Divide by 10</td>
<td>340; 16</td>
</tr>
<tr>
<td>Piece (large - 300g)</td>
<td>&gt;=3</td>
<td>Divide by 10</td>
<td>397; 43</td>
</tr>
<tr>
<td>Piece (small - 150g)</td>
<td>&gt;=1.5</td>
<td>Divide by 10</td>
<td>95; 5</td>
</tr>
<tr>
<td>Rufuc/Jodha (12.5kg)</td>
<td>&gt;=12.5</td>
<td>Divide by 10</td>
<td>37; 15</td>
</tr>
<tr>
<td>Saxarad (20kg)</td>
<td>&gt;20</td>
<td>Divide by 10</td>
<td>312; 793</td>
</tr>
<tr>
<td>Small bag (1 kg)</td>
<td>&gt;=10</td>
<td>Divide by 10</td>
<td>110; 8</td>
</tr>
<tr>
<td>Teaspoon (10 g)</td>
<td>&lt;0.009</td>
<td>Multiply by 10</td>
<td>45; 4</td>
</tr>
</tbody>
</table>

$^{14}$ The first number indicates the number of affected records reported for consumption while the second number states the number of affected records for purchases.
### Table 5: Conversion factor to Kg for specific units and items

<table>
<thead>
<tr>
<th>Item</th>
<th>Unit</th>
<th>Conversion to Kg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biscuits</td>
<td>Piece – large</td>
<td>0.030</td>
</tr>
<tr>
<td></td>
<td>Piece - small</td>
<td>0.010</td>
</tr>
<tr>
<td>Bread</td>
<td>Piece – large</td>
<td>0.400</td>
</tr>
<tr>
<td></td>
<td>Piece - small</td>
<td>0.100</td>
</tr>
<tr>
<td>Eggs</td>
<td>Piece – large</td>
<td>0.070</td>
</tr>
<tr>
<td></td>
<td>Piece - small</td>
<td>0.050</td>
</tr>
<tr>
<td>Canned fish/shellfish</td>
<td>Piece – large</td>
<td>0.420</td>
</tr>
<tr>
<td></td>
<td>Piece - small</td>
<td>0.140</td>
</tr>
<tr>
<td>Grapefruits, lemons, guavas, limes</td>
<td>Piece – large</td>
<td>0.350</td>
</tr>
<tr>
<td></td>
<td>Piece - small</td>
<td>0.100</td>
</tr>
<tr>
<td>Milk</td>
<td>Piece – large</td>
<td>0.750</td>
</tr>
<tr>
<td></td>
<td>Piece - small</td>
<td>0.250</td>
</tr>
<tr>
<td>Milk powder</td>
<td>Piece – large</td>
<td>0.450</td>
</tr>
<tr>
<td></td>
<td>Piece - small</td>
<td>0.100</td>
</tr>
<tr>
<td></td>
<td>Small bag</td>
<td>1.00</td>
</tr>
<tr>
<td>Garlic</td>
<td>Piece – large</td>
<td>0.065</td>
</tr>
<tr>
<td></td>
<td>Piece - small</td>
<td>0.040</td>
</tr>
<tr>
<td>Onion</td>
<td>Piece – large</td>
<td>0.150</td>
</tr>
<tr>
<td></td>
<td>Piece - small</td>
<td>0.095</td>
</tr>
<tr>
<td>Tomatoes</td>
<td>Piece – large</td>
<td>0.200</td>
</tr>
<tr>
<td></td>
<td>Piece - small</td>
<td>0.110</td>
</tr>
<tr>
<td>Bell-pepper</td>
<td>Piece – large</td>
<td>0.150</td>
</tr>
<tr>
<td></td>
<td>Piece - small</td>
<td>0.080</td>
</tr>
<tr>
<td>Sweet/ripe bananas</td>
<td>Piece – large</td>
<td>0.110</td>
</tr>
<tr>
<td></td>
<td>Piece - small</td>
<td>0.070</td>
</tr>
<tr>
<td>Canned vegetables</td>
<td>Piece – large</td>
<td>0.400</td>
</tr>
<tr>
<td></td>
<td>Piece - small</td>
<td>0.200</td>
</tr>
<tr>
<td>Sorghum, flour</td>
<td>Cup</td>
<td>0.200</td>
</tr>
<tr>
<td>Cooking oats, corn flakes</td>
<td>Cup</td>
<td>0.200</td>
</tr>
<tr>
<td>Other cooked foods from vendors</td>
<td>Small bag</td>
<td>1.00</td>
</tr>
<tr>
<td>Purchased/prepared tea/coffee consumed at home</td>
<td>Small bag</td>
<td>0.400</td>
</tr>
<tr>
<td>Other spices</td>
<td>Small bag</td>
<td>0.400</td>
</tr>
</tbody>
</table>

### Table 6: Summary Cleaning Rules for Currency

<table>
<thead>
<tr>
<th>Currency</th>
<th>Condition</th>
<th>Correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Somaliland shillings</td>
<td>Entry in Somaliland shilling</td>
<td>Replace currency to Somali shillings</td>
</tr>
<tr>
<td>Price &lt;=500</td>
<td>Replace currency to Somaliland shillings</td>
<td></td>
</tr>
<tr>
<td>Price&gt;=500,000</td>
<td>Divide by 10</td>
<td></td>
</tr>
<tr>
<td>Somali shillings</td>
<td>Entry in Somali shilling</td>
<td>Replace currency to Somaliland shillings</td>
</tr>
<tr>
<td>Price &lt;=500</td>
<td>Replace currency to Somali shillings</td>
<td></td>
</tr>
<tr>
<td>Price&gt;=500,000</td>
<td>Divide by 10</td>
<td></td>
</tr>
<tr>
<td>USD</td>
<td>Price &gt;1,000</td>
<td>Replace currency to Somaliland shillings</td>
</tr>
</tbody>
</table>

### Table 7: Threshold for Non-Food Item Expenditure (USD)

<table>
<thead>
<tr>
<th>Recall period</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Week</td>
<td>0.05</td>
<td>30</td>
</tr>
<tr>
<td>1 Month</td>
<td>0.20</td>
<td>95</td>
</tr>
<tr>
<td>3 Months</td>
<td>0.45</td>
<td>200</td>
</tr>
<tr>
<td>1 Year</td>
<td>0.80</td>
<td>1,200</td>
</tr>
</tbody>
</table>
Table 8: Consumption of durable goods (per week in current USD)

<table>
<thead>
<tr>
<th>Item</th>
<th>SOM Wave 1 All regions</th>
<th>SOM Wave 1 Mogadishu</th>
<th>Pilot Mogadishu</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>0.74</td>
<td>1.17</td>
<td>1.01</td>
</tr>
<tr>
<td>Mean</td>
<td>1.24</td>
<td>1.52</td>
<td>1.91</td>
</tr>
<tr>
<td>Sd</td>
<td>1.51</td>
<td>1.49</td>
<td>2.62</td>
</tr>
</tbody>
</table>

Table 9: Median consumption of durable goods (per week in current USD)

<table>
<thead>
<tr>
<th>Item</th>
<th>SOM Wave 1 All regions</th>
<th>SOM Wave 1 Mogadishu</th>
<th>Pilot Mogadishu</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air conditioner</td>
<td>0.005</td>
<td>0.005</td>
<td>0.041</td>
</tr>
<tr>
<td>Bed</td>
<td>N/A</td>
<td>N/A</td>
<td>0.861</td>
</tr>
<tr>
<td>Bed with mattress</td>
<td>0.700</td>
<td>0.746</td>
<td>N/A</td>
</tr>
<tr>
<td>Car</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Cell phone</td>
<td>0.361</td>
<td>0.413</td>
<td>0.430</td>
</tr>
<tr>
<td>Chair</td>
<td>0.073</td>
<td>0.072</td>
<td>0.253</td>
</tr>
<tr>
<td>Clock</td>
<td>0.028</td>
<td>0.003</td>
<td>0.046</td>
</tr>
<tr>
<td>Coffee table</td>
<td>0.005</td>
<td>0.005</td>
<td>0.106</td>
</tr>
<tr>
<td>Computer equipment &amp; accessories</td>
<td>0.020</td>
<td>0.020</td>
<td>2.837</td>
</tr>
<tr>
<td>Cupboard, drawers, bureau</td>
<td>0.240</td>
<td>0.240</td>
<td>1.099</td>
</tr>
<tr>
<td>Desk</td>
<td>0.047</td>
<td>0.005</td>
<td>0.429</td>
</tr>
<tr>
<td>Electric stove or hot plate</td>
<td>0.001</td>
<td>0.001</td>
<td>N/A</td>
</tr>
<tr>
<td>Electric or gas stove; hot plate</td>
<td>N/A</td>
<td>N/A</td>
<td>0.012</td>
</tr>
<tr>
<td>Electric stove</td>
<td>N/A</td>
<td>N/A</td>
<td>0.004</td>
</tr>
<tr>
<td>Fan</td>
<td>0.069</td>
<td>0.064</td>
<td>0.101</td>
</tr>
<tr>
<td>Gas stove</td>
<td>0.007</td>
<td>0.007</td>
<td>0.275</td>
</tr>
<tr>
<td>Generator</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Iron</td>
<td>0.043</td>
<td>0.035</td>
<td>N/A</td>
</tr>
<tr>
<td>Kerosene/paraffin stove</td>
<td>0.024</td>
<td>0.007</td>
<td>0.009</td>
</tr>
<tr>
<td>Kitchen furniture</td>
<td>0.023</td>
<td>0.015</td>
<td>1.112</td>
</tr>
<tr>
<td>Lantern (paraffin)</td>
<td>0.000</td>
<td>0.000</td>
<td>0.002</td>
</tr>
<tr>
<td>Lorry</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Mattress without bed</td>
<td>0.217</td>
<td>0.212</td>
<td>N/A</td>
</tr>
<tr>
<td>Mini-bus</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>Mortar/pestle</td>
<td>0.016</td>
<td>0.009</td>
<td>0.112</td>
</tr>
<tr>
<td>Motorcycle/scooter</td>
<td>0.002</td>
<td>0.002</td>
<td>0.006</td>
</tr>
<tr>
<td>Photo camera</td>
<td>0.001</td>
<td>0.001</td>
<td>0.595</td>
</tr>
<tr>
<td>Radio ('wireless')</td>
<td>0.021</td>
<td>0.001</td>
<td>0.016</td>
</tr>
<tr>
<td>Refrigerator</td>
<td>0.282</td>
<td>0.018</td>
<td>0.267</td>
</tr>
<tr>
<td>Satellite dish</td>
<td>0.117</td>
<td>0.008</td>
<td>0.265</td>
</tr>
<tr>
<td>Sewing machine</td>
<td>0.002</td>
<td>0.002</td>
<td>0.732</td>
</tr>
<tr>
<td>Small solar light</td>
<td>0.003</td>
<td>0.003</td>
<td>N/A</td>
</tr>
<tr>
<td>Solar panel</td>
<td>0.000</td>
<td>0.000</td>
<td>0.018</td>
</tr>
<tr>
<td>Stove for charcoal</td>
<td>0.032</td>
<td>0.023</td>
<td>0.020</td>
</tr>
<tr>
<td>Table</td>
<td>0.042</td>
<td>0.042</td>
<td>0.092</td>
</tr>
<tr>
<td>Tape or CD/DVD player; HiFi</td>
<td>0.001</td>
<td>0.001</td>
<td>0.092</td>
</tr>
<tr>
<td>Television</td>
<td>0.330</td>
<td>0.278</td>
<td>0.417</td>
</tr>
<tr>
<td>Upholstered chair, sofa set</td>
<td>0.019</td>
<td>0.019</td>
<td>2.657</td>
</tr>
<tr>
<td>VCR</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Washing machine</td>
<td>0.405</td>
<td>0.368</td>
<td>0.557</td>
</tr>
<tr>
<td>Item</td>
<td>SOM Wave 1 All</td>
<td>SOM Wave 1 Mogadishu</td>
<td>Pilot Wave 1: Awdal, Sanaag, Sool, Togdheer and Woqooyi Galbeed</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>----------------</td>
<td>----------------------</td>
<td>----------------------------------------------------------------</td>
</tr>
<tr>
<td>Air conditioner</td>
<td>0.278</td>
<td>0.241</td>
<td>0.210</td>
</tr>
<tr>
<td>Bed</td>
<td>N/A</td>
<td>N/A</td>
<td>0.210</td>
</tr>
<tr>
<td>Bed with mattress</td>
<td>0.172</td>
<td>0.172</td>
<td>0.111</td>
</tr>
<tr>
<td>Car</td>
<td>0.118</td>
<td>0.111</td>
<td>0.111</td>
</tr>
<tr>
<td>Cell phone</td>
<td>0.188</td>
<td>0.296</td>
<td>0.188</td>
</tr>
<tr>
<td>Chair</td>
<td>0.149</td>
<td>0.371</td>
<td>0.149</td>
</tr>
<tr>
<td>Clock</td>
<td>0.204</td>
<td>0.228</td>
<td>0.204</td>
</tr>
<tr>
<td>Coffee table (for sitting room)</td>
<td>0.279</td>
<td>0.296</td>
<td>0.279</td>
</tr>
<tr>
<td>Computer equipment &amp; accessories</td>
<td>0.182</td>
<td>0.364</td>
<td>0.150</td>
</tr>
<tr>
<td>Cupboard, drawers, bureau</td>
<td>0.150</td>
<td>0.296</td>
<td>0.150</td>
</tr>
<tr>
<td>Desk</td>
<td>0.134</td>
<td>0.502</td>
<td>0.134</td>
</tr>
<tr>
<td>Electric stove or hot plate</td>
<td>0.262</td>
<td>0.005</td>
<td>0.252</td>
</tr>
<tr>
<td>Electric stove</td>
<td>N/A</td>
<td>0.296</td>
<td>N/A</td>
</tr>
<tr>
<td>Fan</td>
<td>0.131</td>
<td>0.235</td>
<td>0.131</td>
</tr>
<tr>
<td>Gas stove</td>
<td>0.174</td>
<td>0.296</td>
<td>0.174</td>
</tr>
<tr>
<td>Generator</td>
<td>N/A</td>
<td>0.296</td>
<td>N/A</td>
</tr>
<tr>
<td>Iron</td>
<td>0.161</td>
<td>0.161</td>
<td>0.161</td>
</tr>
<tr>
<td>Kerosene/paraffin stove</td>
<td>0.224</td>
<td>0.296</td>
<td>0.224</td>
</tr>
<tr>
<td>Kitchen furniture</td>
<td>0.188</td>
<td>0.393</td>
<td>0.188</td>
</tr>
<tr>
<td>Lantern (paraffin)</td>
<td>0.064</td>
<td>0.067</td>
<td>0.064</td>
</tr>
<tr>
<td>Lorry</td>
<td>0.154</td>
<td>0.296</td>
<td>0.154</td>
</tr>
<tr>
<td>Mattress without bed</td>
<td>0.185</td>
<td>0.185</td>
<td>0.185</td>
</tr>
<tr>
<td>Mini-bus</td>
<td>0.153</td>
<td>0.296</td>
<td>0.153</td>
</tr>
<tr>
<td>Mortar/pestle</td>
<td>0.210</td>
<td>0.254</td>
<td>0.210</td>
</tr>
<tr>
<td>Motorcycle/sooter</td>
<td>0.172</td>
<td>0.138</td>
<td>N/A</td>
</tr>
<tr>
<td>Photo camera</td>
<td>0.134</td>
<td>0.296</td>
<td>0.122</td>
</tr>
<tr>
<td>Radio ('wireless')</td>
<td>0.210</td>
<td>0.337</td>
<td>0.210</td>
</tr>
<tr>
<td>Refrigerator</td>
<td>0.133</td>
<td>0.065</td>
<td>0.133</td>
</tr>
<tr>
<td>Satellite dish</td>
<td>0.110</td>
<td>0.303</td>
<td>0.110</td>
</tr>
<tr>
<td>Sewing machine</td>
<td>0.138</td>
<td>0.296</td>
<td>0.138</td>
</tr>
<tr>
<td>Small solar light</td>
<td>0.296</td>
<td>N/A</td>
<td>0.296</td>
</tr>
<tr>
<td>Solar panel</td>
<td>0.005</td>
<td>0.296</td>
<td>0.005</td>
</tr>
<tr>
<td>Stove for charcoal</td>
<td>0.226</td>
<td>0.254</td>
<td>0.254</td>
</tr>
<tr>
<td>Table</td>
<td>0.157</td>
<td>0.296</td>
<td>0.160</td>
</tr>
<tr>
<td>Tape or CD/DVD player; HiFi</td>
<td>0.172</td>
<td>0.138</td>
<td>0.172</td>
</tr>
<tr>
<td>Television</td>
<td>0.131</td>
<td>0.240</td>
<td>0.131</td>
</tr>
<tr>
<td>Upholstered chair, sofa set</td>
<td>0.168</td>
<td>0.289</td>
<td>0.168</td>
</tr>
<tr>
<td>VCR</td>
<td>0.166</td>
<td>0.296</td>
<td>0.130</td>
</tr>
<tr>
<td>Washing machine</td>
<td>0.138</td>
<td>0.171</td>
<td>0.138</td>
</tr>
</tbody>
</table>
Note that the number of (accessible) households does not resemble necessarily the number of PESS households due to the merging of multiple data sources. Therefore, sample weights were adjusted accordingly to scale with PESS household estimates.