**Living Standards Measurement Study-
Integrated Surveys on Agriculture (LSMS-ISA):
Ethiopia Rural Socioeconomic Survey (ERSS)**

Basic Information Document

**Central Statistical Agency & the World Bank
December 2012[1]**

ACRONYMS

AgSS	Annual Agricultural Sample Survey
BMGF	Bill and Melinda Gates Foundation
CAPI	Computer Assisted Personal Interviewing
CSA	Central Statistical Agency (Ethiopia)
DHS	Demographic and Health Survey
EA	Enumeration Area
EIAR	Ethiopia Institute of Agricultural Research
ERSS	Ethiopia Rural Socioeconomic Survey
IFPRI	International Food Policy Research Institute
LSMS-ISA	Living Standards Measurement Study – Integrated Surveys on Agriculture
MOA	Ministry of Agriculture
MOFED	Ministry of Finance and Economic Development
NSDS	National Strategy for the Development of Statistics
PAPI	Paper and Pen Interviewing
WB	World Bank
WFP	World Food Program

Table of Contents

1. Introduction

The Ethiopian Rural Socioeconomic Survey (ERSS) is a collaborative project between the Central Statistics Agency of Ethiopia (CSA) and the World Bank Living Standards Measurement Study-Integrated Surveys of Agriculture (LSMS-ISA) team. The objective of the LSMS-ISA is to collect multi-topic panel household level data with a special focus on improving agriculture statistics and the link between agriculture and other household income activities. The idea of the ERSS is rooted in the need to develop and implement a multi-topic survey that meets Ethiopia's data demands and gaps, is of high quality, accessible to the public, and aligned with the National Strategy for the Development of Statistics (NSDS). The project also aims to build capacity, share knowledge across countries, and improve survey methodologies and technology. In implementing the ERSS, the CSA has established a framework to facilitate consultations with national and international organizations that would provide technical guidance and later utilize the data for policy decision-making and research purposes. These institutions include the Ministry of Agriculture (MOA), Ministry of Finance and Economic Development (MOFED), Agricultural Transformation Agency, World Food Program (WFP), International Food Policy Research Institute (IFPRI), and Ethiopia Institute of Agricultural Research (EIAR).

Ethiopia is one of seven countries being supported by the World Bank, through funding from the Bill and Melinda Gates Foundation (BMGF), to strengthen the production of household-level data on agriculture. This regional project, the LSMS-ISA[2], has the over-arching objective of improving our understanding of agriculture in Sub-Saharan Africa – specifically, its role in household welfare and poverty reduction. The implementation will boost the data collection capacity of the national statistical organizations and the quality of household-level agriculture statistics. Also, the data will provide the bases of analyses looking for insights into how innovation and efficiency can be fostered in the agriculture sector.

The ERSS covered all regional states except the capital, Addis Ababa. It primarily collected information on rural areas. It was implemented in 290 rural and 43 small town enumeration areas (EAs). It is the first round of a long-term project to collect panel data on rural and small town households, their characteristics, welfare and their agricultural activities. ERSS responds to the needs of the country, given the dependence of a high percentage of households on agriculture activities in the country, for information on household agricultural activities along with other information on the households like human capital, other economic activities, access to services and resources. The ability to follow the same households over time makes the ERSS a new and powerful tool for studying and understanding the role of agriculture in household welfare over time as it allows analyses to be made of how households add to their human and physical capital, how education affects earnings, and the role of government policies and programs on poverty, *inter alia*. The ERSS is the first panel survey to be carried out by the CSA that links a multi-topic household questionnaire with detailed data on agriculture.

---

[2] For more information on the LSMS-ISA go to www.worldbank.org/lsms-isa

The ERSS contains several innovative features.

- ♦ Integration of household welfare data with agricultural data;
- ♦ Creation of a panel dataset that can be used to study poverty dynamics, the role of agriculture in development and the changes over time in health, education and labor activities, *inter alia.*
- ♦ Collection of information on the network of buyers and sellers of goods with which the household interacts;
- ♦ Expanding the use of GPS units for measuring agricultural land areas;
- ♦ Involvement of multiple actors in government, academia and the donor community in the development of the survey and its contents as well as its implementation and analysis;
- ♦ Implementation of a Computer Assisted Personal Interviewing (CAPI) application;
- ♦ Active dissemination of agriculture statistics.

The purpose of this document is to provide detailed information on the first wave of the ERSS fielded by the CSA in 2011-12. The survey consists of three rounds of visits to the household. The first round was carried out in September and October 2011 and collected information on post-planting agriculture activities. The second round was conducted in November-December 2011 and fielded the livestock questionnaire to collect information on ownership, production and utilization of livestock, and livestock by products. The third round took place in January- March 2012 to collect information included in the post-harvest agriculture, household, and community questionnaires.

2. The Survey Instruments

The survey consisted of five questionnaires; The **household questionnaire** was administered to all households in the sample. The **community questionnaire** was administered to the community to collect information on the socio-economic indicators of the enumeration areas where the sample households reside.[3] The three **agriculture questionnaires including post-planting agriculture questionnaire, post-harvest agriculture questionnaire** and **livestock questionnaire** were administered to all households engaged in agriculture activities.

Household questionnaire: The household questionnaire provides information on basic demographics; education; health (including anthropometric measurement for children); labor and time use; partial food and non-food expenditure; household nonfarm income-generating activities; food security and shocks; safety nets; housing conditions; assets; credit; and other sources of household income (Table 2.1). Household location is geo-referenced in order to be able to later link the ERSS data to other available geographic data sets (See Appendix 1 for discussion of the geo-data provided with the ERSS).

---

[3] The community questionnaire does not collect information from communities in the sociological sense. The data cannot be used to represent communities in Ethiopia. The data collected at the community level represent information that is common to the households selected for inclusion in the selected sample EAs.

Community questionnaire: The community questionnaire solicits information on infrastructure; community organizations; resource management; changes in the community; key events; community needs, actions and achievements; and local retail price information (Table 2.2).

Agriculture questionnaire: The post-planting and post-harvest agriculture questionnaire focus on farming activities and solicit information on land ownership and use; farm labor; inputs use; GPS land area measurement and coordinates of household fields; agriculture capital; irrigation; and crop harvest and utilization. The livestock questionnaire collects information on animal holdings and costs; and production, cost and sales of livestock by products (Table 2.3).

Table 2.1: ERSS Household Questionnaire

Section	Topic	Respondent	Description
Cover	Cover	Field staff	Household location identification; household size and head's name; field staff identification; and observation notes by enumerator regarding the interview.
1	Roster	Household head or spouse.	Roster of individuals living in the household and basic demographics; for members less than 18, parental education and occupation.
2	Education	Individuals 5 years and above	Educational attainment, enrollment, attendance, school characteristics, and expenditures for the 2011-12 academic year.[4]
3	Health	Individuals	Health problems, types of injury/illness, medical assistance/consultation, disabilities, and anthropometrics (children under five years).
4	Labor and Time Use	Individuals 7 years and above	Labor market participation during the last seven days, participation in food for work/ public works programs, and participation in domestic activities.
5A	Food Consumption Last 7 days	Person responsible for household purchases	Food consumption (quantity and value) by any household member in the last 7 days by source for a subset list of food items.
5B	Food Aggregate Last 7 days	Person responsible for household purchases	Summary on consumption of food in the last 7 days. Meal sharing with non-household members.
6	Non-food Expenditure	Person responsible for household purchases	Household expenditures on non-food items.
7	Food Security	Household head or eligible adult	Food security status of households in during the past 7 days/12 months.
8	Shocks	Household head or eligible adult	Shocks during the last 12 months and their impact on income, assets, food production, stock and purchase.
9	Housing	Household head or eligible adult	Dwelling ownership, and characteristics of the dwelling and utilities.

---

[4] The 2011-12 school started in September 2011 and ended in July 2012.

Section	Topic	Respondent	Description
10	Household assets	Household head or eligible adult	Household ownership of assets.
11	Non-farm Enterprises	Owner or manager of enterprise	Characteristics of enterprises owned by the household.
12	Other Income	Household head or eligible adult	Others sources of household income during the last 12 months.
13	Assistance	Household head or eligible adult	Assistance provided to the household by governmental and non-governmental agencies.
14	Credit	Household head or eligible adult	Loans or credit received by the household.
15	Contact information	Household head or eligible adult	Contact information.

Table 2.2: ERSS Community Questionnaire

Section	Topic	Respondent	Description
Cover (1.1&1.2)	Cover	Field staff	Community location identification; field staff identification; date and time of interviews.
Cover (1.3)	Cover	Direct Observation by the Field Staff	Community characteristics.
2	Roster of Informants	Informants	Respondent characteristics.
3	Community Basic Information	Informants	Mobility, population, religion, marriage types, common land use.
4	Access to Basic Services	Informants	Transportation, markets, proximity to the nearest town and major urban centers, electrification, bank and microfinance institutions, piped water.
5	Economic Activities	Informants	Main sources of employment, migration to and from the locality for work, cooperatives and microenterprises.
6	Agriculture	Informants	Agricultural activities, including major crops, main planting and harvesting seasons, rain seasons, input use, agricultural extension, and irrigation.
7	Changes	Informants	Important events in the community in the last five years.
8	Community Needs and Actions	Informants	Initiation, participation and mobilization of resources for community projects including roads, school, health facility, water, natural resource management, public transport, agriculture, law enforcement, etc.
9	Productive Safety Nets Program	Informants	Participation in the productive safety nets program. Management and performance of the program in the community.
10	Market Prices	Sellers in the nearby market- Market center 1	Market prices in the first closest market center.
11	Market Prices	Sellers in the nearby market- Market center 2	Market prices in the other closest market center.

Table 2.3: ERSS Post-planting Agriculture Questionnaire[5]

Section	Topic	Respondent	Description
Cover	Cover	Field Staff	Holder[6] location identification; household head name, holder name, household size, agriculture holding type (farming, livestock, or both); field staff identification.
1	Household Roster	Household head or eligible adult	Name, age, and gender of each household member and holding type (farming, livestock or both).

[5] No Section 6 in the printed version.

[6] A holder is a person who exercises management control over the operations of the agricultural holdings and makes the major decisions regarding the utilization of the available resources. S/he has technical and economic responsibility for the holding. S/he may operate the holding directly as an owner or as a manager.

2	Parcel Roster	Owner or manager of the parcel	Information on all parcels owned and/or managed by the holder:
3	Field Roster	Field Staff (for field measurements using GPS or rope and compass). Manager of field for other questions.	Information on all fields (sub-parcels) owned and/or managed including holder reported self-reported area, GPS or rope and compass measured area, and other details on the fields.
4	Crop Field Roster	Manager of field	Crop planting/management information for each crop on each field.
5	Seeds Roster	Manager of field	Seed related information for each crop planted on each field.
7	Miscellaneous	Manager of field	Information on holder characteristics including chemical fertilizer use, and access and use of credit, extension/advisory services.
NR	Network Roster	Manager of field	Roster of individuals, places or businesses from/where the holder purchases agricultural inputs.

Table 2.4: ERSS Post-harvest Agriculture Questionnaire

Section	Topic	Respondent	Description
Cover	Cover	To be completed by the field staff	Holder location identification; household head name, holder name, household size, agriculture holding type (farming, livestock, or both). Field staff identification.
1	Household Roster	Household head or eligible adult.	Name, age, and gender of each household member and holding type (farming, livestock or both).
9	Crop cut/ Crop Harvest by Field	Field staff (for the 2mX2m crop cut) / Holder/ Manager of field	Crop cut information for selected fields including fresh and dry weight (from a 2mX2m crop cut). Information on crop use, area harvested, amount harvested and damage to crops. This section excludes permanent, tree and root crops.
10	Harvest Labor	Holder/ Manager of field	Hired and household member labor used in harvest activities for each crop on each field. This section excludes permanent, tree and root crops.
11	Crop Disposition/ Sales	Holder/ Manager of field	Crop disposition/sale information. This section excludes permanent, tree and root crops.
12	Crop Disposition/ Sales (Tree / Permanent and Root Crops)	Holder/ Manager of field	Harvest information on permanent, tree, and root crops.
NR	Network Roster	Manager of field	Roster of individuals, places or businesses to/where the holder sells crops.

Table 2.5: ERSS Livestock Questionnaire

Section	Topic	Respondent	Description
Cover	Cover	To be completed by field Staff.	Holder location identification; household head name, holder name, household size, agriculture holding type (farming, livestock, or both). Field staff identification.
1	Household Roster	Household head or eligible adult	Name, age, and gender of each household member and holding type (farming, livestock or both).
8A	Livestock Population and Products	Holder or manager/owner of livestock	Characteristics of livestock owned.
8B	Livestock Feed Utilization	Holder or manager/owner of livestock	Type and source of feed used and participation in livestock improvement related extension services.
8C	Livestock Byproduct	Holder or manager/owner of livestock	Production and utilization of livestock byproducts.

3. Sample Design

The ERSS sample is designed to be representative of rural and small town areas of Ethiopia.[7] The ERSS rural sample is a sub-sample of the 2011/2012 Agricultural Sample Survey (AgSS). The small town sample comes from the universe of small town EAs. The ERSS sample size provides estimates at the national level for rural and small town households. At the regional level, it provides estimates for four regions including Amhara, Oromiya, SNNP, and Tigray.

The sample is a two-stage probability sample.

The first stage of sampling entailed selecting primary sampling units, which are a sample of the CSA enumeration areas (EAs). For the rural sample, 290 EAs were selected from the AgSS EAs. The AgSS EAs were selected based on probability proportional to size of the total EAs in each region. For small town EAs, a total of 43 EAs were selected.

In order to ensure sufficient sample in the most populous regions (Amhara, Oromiya, SNNP, and Tigray), quotas were set for the number of EAs in each region. The sample is not representative for each of the small regions including Afar, Benshangul Gumuz, Dire Dawa, Gambella, Harari, and Somalie regions. However, estimates can be produced for a combination of all smaller regions as one "other region" category. This is presented in table 3.1.

---

[7] The CSA defines small towns based on population estimates from the 2007 Population Census; a town with the population of less than 10,000 is a small town.

Table 3.1: First Stage Sample Design

	Population share	Rural EAs	Small town EAs
National	100.0%	290	43
Regions			
Tigray	6.6%	30	4
Afar	1.7%	10	2
Amhara	26.6%	61	10
Oromia	37.6%	55	10
Somali	4.5%	20	3
Benishangul-Gumuz	1.0%	10	1
SNNP	20.8%	74	10
Gambela	0.4%	10	1
Harari	0.3%	10	1
Dire Dawa	0.5%	10	1

The second stage of sampling was the selection of households to be interviewed in each EA.

For rural EAs, a total of 12 households are sampled in each EA. Of these, 10 households were randomly selected from the sample of 30 AgSS households. The AgSS households are households which are involved in farming or livestock activities. Another 2 households were randomly selected from all other households in the rural EA (those not involved in agriculture or livestock). In some EAs, there is only one or no such households, in which case, less than two non-agricultural households were surveyed and more agricultural households were interviewed instead so that the total number of households per EA remains the same.

In the small town EAs, 12 households are selected randomly from the listing of each EA, with no stratification as to whether the household is engaged in agriculture/livestock.

Households were not selected using replacement. Thus the final number of household interviewed was slightly less than the 3,996 as planned in the design. A total of 3,969 households were interviewed with a response rate of 99.3 percent[8].

The actual distribution of the ERSS final sample is shown in Table 3.2.

---

[8] Post-harvest interviews were interrupted for security reasons in one enumeration area in Liben zone, Somali region. All other interviews (Post-planting agriculture, livestock, household and community questionnaires) were completed earlier.

Table 3.2: Distribution of Final Sample for ERSS
by Region and Zone in Rural and Small town areas

Region Code	Region Name	Zone Code	Zone Name	Rural		Small Town	
				EA	HH	EA	HH
1	Tigray	1	North Western	5	60	1	12
		2	Central	7	84	1	12
		3	East	6	72	1	12
		4	South	6	72	1	12
		5	West	6	72	0	0
2	Afar	1	Zone-1	7	84	1	12
		3	Zone-3	3	36	1	12
3	Amhara	1	North Gonder	7	83	2	24
		2	South Gonder	6	72	1	11
		3	North Wollo	8	96	1	12
		4	South Wollo	6	72	1	12
		5	North Shewa	7	84	1	12
		6	East Gojam	6	71	1	10
		7	West Gojam	7	83	1	12
		8	Wag Himra	3	36	1	12
		9	Awi	5	59	1	10
		10	Ormiya	3	36	1	12
		12	Argoba Special Wereda	3	36	0	0
4	Oromiya	1	West Wellega	4	48	1	12
		2	East Wellega	3	35	1	12
		3	Ilu Aba Bora	3	36	1	12
		4	Jimma	4	48	1	10
		5	West Shewa	3	36	1	12
		6	North Shewa	3	36	1	10
		7	East Shewa	4	48	0	0
		8	Arsi	3	36	1	12
		9	West Hararge	4	48	1	12
		10	East Hararge	3	36	1	12
		11	Bale	4	47	1	12
		12	Borena	3	36	0	0
		13	South West Shewa	3	36	1	9
		14	Guji	3	34	0	0
		17	West Arsi	3	36	0	0
		18	Qelleme Wollega	2	24	0	0
		19	Horo Gudru Wellega	3	36	0	0
5	Somali	1	Shinille	5	60	1	12
		2	Jijiga	8	96	1	12
		9	Liben	7	81	1	12

Region Code	Region Name	Zone Code	Zone Name	Rural		Small Town	
				EA	HH	EA	HH
6	Benshangul Gumuz	2	Metekel	3	36	0	0
		3	Assosa	3	36	1	12
		4	Kamashi	2	24	0	0
		6	Mao Komo	2	24	0	0
7	SNNP	1	Gurage	5	60	1	12
		2	Hadiya	4	48	1	12
		3	Kembata Timbaro	3	36	0	0
		4	Sidama	6	72	1	12
		5	Gedeo	5	60	0	0
		6	Wolayita	4	48	1	12
		7	South Omo	3	36	1	12
		8	Sheka	3	36	0	0
		9	Kefa	4	48	1	12
		10	Gamu Gofa	5	60	1	12
		11	Bench Maji	5	57	1	11
		12	Yem	3	36	0	0
		13	Amaro Special	2	24	0	0
		14	Burji Special	2	24	0	0
		15	Konso Special	3	36	0	0
		16	Derashe	2	24	0	0
		17	Dauro	4	48	1	12
		18	Basketo	3	36	0	0
		19	Konta	2	24	0	0
		20	Silti	4	48	1	12
		21	Alaba	2	24	0	0
12	Gambella	1	Anuak	3	36	1	12
		2	Nuer	2	24	0	0
		3	Mejenger	4	48	0	0
		4	Etang Special	1	12	0	0
13	Harari	1	Zone 1	10	120	0	0
15	Dire Dawa	1	Zone 1	10	120	0	0
				290	**3466**	**43**	**503**

4. Field Work

Both paper assisted personal interviews (PAPI) and computer assisted personal interview (CAPI) methods were used. Most of the interviews were conducted using PAPI. Household and community questionnaires were fully PAPI. CAPI was implemented using Surveybe software as a pilot in 43 of the 290 rural EAs and was applied only to the agriculture questionnaires. These 43 EAs were purposively selected among the random sample of 290 EAs, based on electricity availability for re-charging the CAPI tablets during the field work.

4.1 Training of Field Staff and Data Entry Operators for the Survey

Four training sessions were held for trainers and field staff. These sessions included technical training for the survey design and implementation, and content and practical sessions for the field workers.

The first training session was for the CAPI software development, Surveybe Designer. It took place in Addis Ababa at the CSA headquarters from June 4-10, 2011. It was provided to CSA staff who were converting the agriculture questionnaires from PAPI to CAPI. The training focused on how to design a questionnaire using Surveybe Designer software and discussed topics on installing the software, response list, lookup tables, routing and enablement, validation techniques, and exporting the questionnaire.

The second training was training of trainers (TOT). The TOT was carried out in July 2011 in Hawassa. Trainers were trained on the content of the questionnaires and manuals. The training mainly focused on the three agriculture questionnaires: post planting, post-harvest and livestock. In addition, specific trainers were trained on the Surveybe implementer software. The Surveybe Implementer training covered both hardware and software issues. On the hardware part, the field staff were advised how to handle and manage the CAPI machine (ASUS portable computer) and its battery. On the data management and software techniques, sessions covered were: installing the Surveybe Implementer software, importing questionnaires, entering data/responses, navigating between different tabs, validating responses and error checks, and backing up file.

The third training was conducted for enumerators and supervisors in August 2011.Enumerators and supervisors were trained on the content of agriculture questionnaires and manuals. In addition, Surveybe Implementer training was given to the field staff assigned to CAPI EAs. The CAPI enumerators and supervisors were trained in one center in Hawassa, while the PAPI enumerators and supervisors were trained in their respective branch offices.

A re-fresher training took place in the middle of field work from January 9-14, 2012. It gathered all the field staff at five locations: Adama, Bahir Dar, Hawassa, Mekelle, and Nekemte. Interviewers from two or more branches were trained at each center.[9] CSA head office experts,

---

[9] (1) Adama, Asayita, Asebe Teferi, Debreberhan, Dire Dawa, Harar and Jijiga in Adama town, (2) Bahirdar, Dessie and Gondar branches in Bahirdar town; (3) Arba Minch, Goba, Hawassa, Hosaena, Negelle and Sodo branches in Hawassa town, (4) Mekelle and Shire branches in Mekelle town and (5) Assossa, Gambella, Jimma and Nekemte branches in Nekemte town.

and World Bank staff and consultants were present as trainers and facilitators. The training focused on the household and community questionnaires (which were fielded in the third phase of the field work) as well as a refresher on the post-harvest agriculture questionnaire.

4.2 Field Work Organization & Data Collection

The survey was designed to be implemented in three rounds following the AgSS field schedule. The first round took place between September and October 2011. In this round, the post-planting agriculture questionnaire was administered. The second round took place between November and December 2011 when the livestock questionnaire was administered. The third round took place from January through March 2012 when the household, community and post-harvest agriculture questionnaires were administered.

Table 4.1: Implementation Timeline

ERSS Questionnaire	2011				2012		
	Sep	Oct	Nov	Dec	Jan	Feb	Mar
Post-planting agriculture questionnaire	███	███					
Livestock questionnaire			███	███			
Household, community, and post-harvest agriculture questionnaire					███	███	███

The post-planting and post-harvest questionnaires collected information on the same fields and crops in two different rounds. While the post-planting questionnaire collected field and crop level information such as on area measurement, inputs, and other farming practices, the post-harvest questionnaire captured harvest, harvest inputs, crop damage, and utilization.

The specific section on harvest amount (Section 9) which includes crop cutting and self-reported harvest for crops by field (excluding tree/root/permanent crops) was fielded with either the post-planting and post-harvest agriculture questionnaires – depending on the timing of the harvest for that crop/field. For some crops, in some places, crop cutting started earlier than the date of the scheduled post-harvest round. Therefore, enumerators captured the crop cutting data while conducting the post-planting questionnaire because it needed to be done when the farmer actually harvests the crop. In those households where the crop cutting information was collected during the post-planting round, harvest information was collected during the post-harvest round and collection of crop cutting information was not repeated.

Most of the questions in the post-planting, post-harvest and livestock questionnaires were completed by asking the holder (Table 2.3-2.5). However, some questions were directly measured by the enumerator as per the specific instructions given in the questionnaires and field manuals. In the Post-planting questionnaire, the enumerator measured all the fields (sub-parcels) managed by the holder using GPS or, if a small field, by compass and rope. The enumerator also conducted the crop cutting and completed all the applicable questions in the crop cutting section.

Similarly, questions in the household questionnaire were collected from the most knowledgeable person. For some sections, individual members were asked (see Table 2.1). The enumerator conducted the anthropometric measurements for children 6-59 months.

The community questionnaire was collected from focus groups of community informants and by direct observation (Table 2.2). The survey was conducted by the field supervisors. The community informants in each EA were formed based on the instructions in the community questionnaire and the manual. The community questionnaire also collected commodity price information from up to two nearby markets. The price data was collected by asking the sellers in the market centers.

Resident enumerators were used to administer the household, agriculture (post-planting and post-harvest), and livestock questionnaires. The CSA assigned one resident enumerator for each EA. The enumerator lived in the EA for the entire survey period from September 2011 to March 2012. Daily laborers/ field guides were also hired locally for a few days to assist the enumerators in parcel/ field measurement and crop cutting activities. They also helped when child anthropometrics were taken.

One field supervisor was assigned to supervise the work of three enumerators in three EAs. As mentioned above, the field supervisors also administered community questionnaires in their respective three EAs.

4.3 Fieldwork Monitoring and Evaluation

Routine supervision by CSA's field supervisors entailed the field level coordination by 24 of 25 CSA branch offices.[10] Branch level statisticians and supervisors who were assigned to this project conducted the routine supervision. The branch supervisors made three visits to the EAs between September 2011-March 2012. As noted above, one field supervisor checked the work of three enumerators in three EAs. The last visit was combined with community interviews that were conducted by the supervisors themselves. Up to two branch statisticians were also in the field to check the work of the supervisors and enumerators.

Additional supervision was done by teams composed of the CSA head office experts and Bank staff and consultants, who made two supervision visits. Each visit took about 2 weeks. The first visit was held in September-October 2011 when interviews with the Post-planting questionnaire were being conducted. The second visit was in January-February 2011 when the household, community, and post-harvest agriculture data were being collected. A total of 7 teams were dispatched. Each team comprised two experts. Before leaving for the field, the team convened for a half day at the head office and discussed the supervision strategy. Once in the field, the teams went to different directions and updated each other on their daily findings. The teams covered all regions and checked samples of completed questionnaires in 105 EAs in total.

5. Data Management and Description of Datasets

---

[10] The Addis Ababa branch office is not part of ERSS as all EAs in Addis are urban.

5.1 Data Management

5.1.1 PAPI Data Entry and CAPI Data Export

The completed paper (PAPI) questionnaires were sent to the CSA headquarters in Addis Ababa. The completed questionnaires were first checked by manual editors. For each ERSS questionnaire, half day training was organized for data editors using the questionnaires and the manual. The editors checked completeness (taking inventory) and cross-checked the questionnaires with the EA codebook. Questionnaires with inconsistent responses or with errors were corrected by contacting the branch offices or, in some cases, by sending the questionnaires back to the field.

Checked questionnaires were keyed by data entry clerks at the head office. The CSPro software was used to design the data entry programs. The paper questionnaires that went through this process include agriculture questionnaires from 300 EAs, and all household and community questionnaires from all the 333 EAs.

The 33 EAs with agriculture questionnaires conducted using CAPI sent electronic data files back to the CSA headquarters where they were exported into Stata.

5.1.2 Data Cleaning

The data cleaning process was done in two stages. The first stage was at the CSA head office using the CSA's data cleaning staff. The CSA data cleaning staff used the CSpro data cleaning application to capture out of range values, outliers, and skip inconsistencies from the batch error reports. Once the errors were flagged in the batch error reports, the hard copy of the original questionnaire was retrieved and checked if the errors were at the data collection, editing, or entry level. Editing and entry level errors were corrected at the head office. Field level errors were communicated with the branch offices in the regions for correction.

The second level of data cleaning entailed checking the *Stata* data sets for outliers and completeness and referring to the paper questionnaire where relevant.

5.1.3 Weighting of Data

The ERSS data needs to be weighted to represent the national-level population of rural and small town households. A population weight was calculated for the households and this weight variable is included in all the datasets. When applied, this weight will raise the sample households to national values for rural areas and small towns.

5.2 Description of Public Datasets

The electronic datasets are organized by questionnaire with the following labels on file names in paraentheses: household (*hh*), community (*co*), post-planting agriculture (*pp*), post-harvest agriculture (*ph*), and livestock (*ls*).

Within each questionnaire type, the data file naming scheme is a combination of the prefix '*sect*', followed by section number, and then followed by suffix '*hh_w1*' for household wave 1 data, and '*co_w1*' for community wave 1 data. Similarly, the suffixes for post planting and post-harvest as well as livestock wave 1 data are '*pp_w1*', '*ph_w1*', and '*ls_w1*' respectively.

For example, the data set that corresponds with the section 1 of the household questionnaire is in the data file '*sect1_hh_w1*'. The exception to this rule are sections where the files are broken down even further due to different reference period or different levels of recording the data. An example is section 5 of the household questionnaire on consumption where the section is split into 4 files with each file corresponding with the reference period collected in the section. In this case, the name of the corresponding files will be '*sect5a_hh_w1*', '*sect5b_hh_w1*', sect5c1_hh_w1' and sect5c2_hh_w1'.

Each dataset has identification variables, a rural area/small town indicator variable, and sampling weight.

For purposes of maintaining the confidentiality of the data all names and addresses including contact addresses and field descriptions in the post planting agriculture questionnaire have been removed from the datasets. In addition, the GPS coordinates have also been removed as these could be used to locate households and fields with accuracy. However, as a courtesy to users, a set of geo variables are provided with the data as described in Appendix 1.

5.2.1 Household Data Files

The household data are organized in 22 data files (Table 5.1).

Table 5.1: ERSS Household Data Files

Section	Section Name/ Content	Dataset Filename
Cover	Cover	*sect_cover_hh_w1*
1	Roster	*sect1_hh_w1*
2	Education	*sect2_hh_w1*
3	Health	*sect3_hh_w1*
4	Labor and Time Use	*sect4_hh_w1*
5A	Food Consumption Last 7 days	*sect5a_hh_w1*
5B	Food Consumption Last 7 days	*sect5b_hh_w1*
5C1	Meal Sharing Last 7 days	*sect5c1_hh_w1*
5C2	Number of Days Meal shared Last 7 days	*sect5c2_hh_w1*
6A	Non-Food Expenditure (one month)	*sect6a_hh_w1*
6B	Non-Food Expenditure (one year)	*sect6b_hh_w1*
7	Food Security	*sect7_hh_w1*

8	Shocks	*sect8_hh_w1*
9	Housing	*sect9_hh_w1*
10	Household assets	*sect10_hh_w1*
11A	Non-Farm Enterprises Filter	*sect11a_hh_w1*
11B	Non-Farm Enterprises (at the household-enterprise level)	*sect11b_hh_w1*
11C	Non-Farm Enterprises (questions 18-19 at the household level)	*sect11c_hh_w1*
12	Other Income	*sect12_hh_w1*
13	Assistance	*sect13_hh_w1*
14A	Credit (Household Level)	*sect14a_hh_w1*
14 B	Credit (Loan Level Information)	*sect14b_hh_w1*

5.2.2 Community Data

The community data files are organized in 14 data files (Table 5.2).

Table 5.2: ERSS Community Data Files

Section	Topic	Dataset Filename
1A	Cover/ Identification	*sect1a_com_w1*
1B	Cover/ Community Overview/ Observation	*sect1b_com_w1*
2	Roster of Informants	*sect2_com_w1*
3	Community Basic Information/ Demographics	*sect3_com_w1*
4	Access to Basic Services/ Infrastructure	*sect4_com_w1*
5	Economic Activities/ Employment	*sect5_com_w1*
6	Agriculture	*sect6_com_w1*
7	Changes/Events	*sect7_com_w1*
8	Community Needs	*sect8_com_w1*
9	Productive Safety Nets Program	*sect9_com_w1*
10A1	Market Prices: Market 1 Location	*sect10a1_com_w1*
10A2	Market Prices: Market 2 Location	*sect10a2_com_w1*
10B1	Market Prices in Market 1	*sect10b1_com_w1*
10B2	Market Prices in Market 2	*sect10b2_com_w1*

5.2.3 Agriculture Data

The agriculture data is in three folders: post-planting, post-harvest and livestock. The sections and the file names are presented in Tables 5.3, 5.4 and 5.5 below.

Table 5.3: ERSS Post-planting Agriculture Data Files[11]

Section	Section Name	Dataset Filename
Cover	Cover	*sect_cover_pp_w1*
1	Household Roster	*sect1_pp_w1*
2	Parcel Roster	*sect2_pp_w1*
3	Field Roster	*sect3_pp_w1*
3RCA	Field Roster/Area Measurement	*sect3_rca_pp_w1*
3RCB	Field Roster/Area Measurement	*sect3_rcb_pp_w1*
4	Crop Field Roster	*sect4_pp_w1*
5	Seed acquisition	*sect5_pp_w1*
7	Planted Field Crops	*sect7_pp_w1*
NR	Network Roster	*sect_nr_pp_w1*

Table 5.4: ERSS Post-harvest Agriculture Data Files

Section	Section Name	Dataset Filename
Cover	Cover	*sect_cover_ph_w1*
1	Household Roster	*sect1_ph_w1*
9	Crop Cut and Harvest	*sect9_ph_w1*
10	Harvest Labor	*sect10_ph_w1*
11	Crop Utilization (cereals, pulses, oil seeds)	*sect11_ph_w1*
12	Crop Utilization (permanent tree and root crops)	*sect12_ph_w1*
NR	Network Roster	*sect_nr_ph_w1*

Table 5.5: ERSS Livestock Data Files

Section	Section Name	Dataset Filename
Cover	Cover	*sect_cover_ls_w1*
1	Household Roster	*sect1_ls_w1*
8A	Livestock Inventory	*sect8a_ls_w1*
8B	Livestock Utilization/Sales	*sect8b_ls_w1*

[11] There is no section 6 in the questionnaire.

Section	Section Name	Dataset Filename
8C	Livestock byproduct	*sect11_ls_w1*

5.2.4 Geospatial Data

The ERSS data files also include additional geospatial data computed for data users. The geovariables are stored in two data files: household-field-level data (Pub_ETH_PlotGeovariables_Y1), and household-level data (Pub_ETH_HouseholdGeovariables_Y1). Information on the ERSS geospatial data is presented in Appendix 1.

5.2.5 Land area conversion factors for local units

The agricultural questionnaire collects land size as reported by holders, in addition to GPS-device measures of field size. As a courtesy to data users, a set of conversion factors for local units have been compiled from several sources. The data file (ET_local_unit_area_conversion) is provided with the ERSS data to facilitate converting non-standard units into square meters. The provision of these conversion factors should not be interpreted as an endorsement of these factors by the CSA or the LSMS-ISA team. Neither the CSA nor the LSMS-ISA has validated these conversion factors under the ERSS program. The conversion factors were sourced primarily from a report completed by the CSA[12] and a study on land certification in Ethiopia[13], and supplemented by other internal CSA sources.

6. Using the ERSS Public Data

6.1 File Structure

The data should always be used in conjunction with the questionnaire and the interviewer's instruction manuals.

---

[12] Wigton, William, Arun Srivastava, Samia Zekaria, Yakob Mudesir, Yasin Mossa, Girma Tadesse, Biratu Yigezu, Aberash Tariku, Habekristos Beyene , and Jemal Abdi. 2009. "Agriculture Statistical Methodology in Ethiopia – Main Text." Central Statistical Agency, Addis Ababa.

[13] Deininger, Klaus, Daniel Ayalew Ali, Stein Holden, and Jaap Zevenbergen. 2010. "Rural Land Certification in Ethiopia: Process, Initial Impact, and Implications for Other African Countries," World Development, 36 (10) pp 1786-1812.

6.2 Merging Datasets

6.2.1 Household and Agriculture Datasets

All households are identified by the variable *household_id* in the household and agriculture data. This variable is used as the unique key variable in the merging of all household-level data files. In some of the other datasets, additional key variables may be required. In the case of individual level files, the variable that uniquely identifies the individual within the household is *individual_id*. So in order to merge any two individual type files, both the variables *household_id* and *individual_id* would be used. In the agriculture datasets, parcel files are merged using *holder_id* and *parcel_id* while crop files are merged using *holder_id, parcel_id, field_id,* and *crop_code*.

6.2.2 Community Datasets

The community questionnaire is administered at the EA level. A unique EA identifier, *ea_id,* is in every data file. This variable is simply the concatenation of the variables *region, zone, wereda,* and *ea*.

6.3 Network Roster

A network roster is included in the post-planting and post-harvest agriculture questionnaires. The network roster keeps a record of the list of places/persons (businesses, markets, friends, etc.) with which the agricultural holder engages in agricultural trading activities. Each person/place is assigned the network code of the line in which it is in that section and is recorded only once so we have for example, network codes N1, N2 etc. which are just a serialization of the persons/places in the holder's network. An entry in the network roster is similar to an individual in the household roster where an individual acquires the individual code of the line in which the person's name and its relationship to the household head and other individual level identifiers/ variables are recorded. Likewise, in the network roster, the N1, N2, N3 etc. are IDs. Other two variables recorded in the network roster are codes for the type of the network (relative, friend, neighbor, government agency, etc.,) and code for the location of the person/place.

7. Overall Problems and Challenges Faced During the First Wave of the Survey and Recommendations for the Second Wave of the Survey

Designing and implementing a complex survey such as the ERSS presents various challenges. In this section we outline some key issues that arose, lessons learned and make recommendations for the next wave of the survey.

7.1 Survey Instruments

The questionnaires used in the field were written in Amharic. Because the questionnaires were being edited simultaneously in Amharic and English, the Amharic questionnaires missed some of the latest revisions that had been made to the English versions. These errors were corrected in the field.

7.2 Challenges in implementing CAPI

CAPI EAs were selected based on electricity availability in the area so that the enumerators can charge the batteries in the ASUS portable computers. However, even in these areas, frequent power interruption delayed fieldwork in some places. In addition, the screens of the machines had visibility problems when used outdoors. The interviews had to be conducted indoors. The enumerators had to arrange a convenient time with the respondent to conduct the interviews in the respondent's house.

CAPI was also constrained by data exporting problems. In some cases, the data exported from Surveybe *yml* files to *csv* and *stata* formats were found to be incomplete. This needed the software company's intervention and the CSA had to contact the company each time there was a problem with the CAPI.

7.3 Challenges in implementing PAPI

The paper questionnaires were designed to have a fold out portion (flap) where parcel, field and crop identification information was written that could be matched to all pages in the questionnaire. The questionnaires were printed without this fold out section. This created problem in transferring the ids for parcel, field and crop rosters across pages and from one section to another.

7.4 Data Entry and Cleaning

The CSA has not yet introduced an approach of data entry in the field for its surveys.[14] The questionnaires instead are shipped to the branch office centers or head office in Addis Ababa for further processing. Therefore, some time elapsed between completion of interviews and data entry

---

[14] In this model of field work, there is a data entry person with the field staff in the EA who enters the questionnaire while the team is in the EA. The system is difficult to integrate with the AgSS as they use resident enumerators who live in the EA working only on surveys implemented in the area.

and checking. In addition to data cleaning procedures carried out in the field, more rigorous data cleaning desk edits were made to the household food consumption section (see Appendix 2).

7.5 Consumption data and seasonality

The household questionnaire was implemented in the January-March period during the post-harvest round. This is the period where food is relatively abundant in most parts of the country. Therefore, consumption-related analysis needs to consider that this reflects consumption only in this season.

7.6 Self-reported crop production

In the AgSS, production is estimated on the basis of crop-cutting of randomly selected crops/fields in each EA (see separate documentation on the method of crop-cutting). This is scaled up to project production of crops at the EA level. However, it does not allow for holder specific production analysis. The original design intent for the ERSS was to ask holders to report the total amount of crops harvested from each field on which the crop was grown. This is how the post-harvest questionnaire is designed. The self-reported harvest would also be accompanied by the crop-cut data in the subset of cases where the crop/field was randomly selected for the crop cut.

However, in practice, the ERSS manual *incorrectly* instructed field staff to only collect self-reported harvest for the crops/ fields which were selected for crop-cutting. In practice, then, despite the intention of the survey, most crop/ fields do not have self-report harvest information.

7.7 Utilization of Permanent Tree and Root Crops

In the Post-harvest agriculture module Question number 19 of Section 12 was not included in the fielded version of the paper questionnaire. Therefore, information on the utilization of permanent tree and root crops is available only for those enumeration areas where CAPI was implemented.

7.8 Child Height Outliers in the Public Data

After the data were launched in May 2013, additional outliers in child height were checked. The outliers were due to data collection and entry errors. Therefore, the height variable in Section 3 of the household data needs to be updated using the correct values listed in Appendix 3.

Appendix 1: Geospatial Data with the ERSS

The ERSS collects confidential information on respondents. The confidential variables pertain to (i) names of the respondents to the household and community questionnaires, (ii) village and constituency names, (iii) descriptions of household dwelling and agricultural field locations, (iv) phone numbers of household members and their reference contacts, (v) GPS-based dwelling and agricultural field locations, (vi) names of the children of the head/spouse living elsewhere, (vii) names of the deceased household members, (viii) names of individuals listed in the network roster, and (ix) names of field staff. To maintain confidentiality, this information is not included in the ERSS public use data.

To increase the use of the ERSS data, a set of geospatial variables has been provided by using the georeferenced field and dwelling locations in conjunction with various spatial databases that were available to the survey team. These include measures of distance, climatology, soil and terrain, and other environmental factors. Time-series on rainfall and vegetation have also been used to describe the survey agricultural season relative to normal conditions. These variables are intended to provide some understanding of how geophysical characteristics vary at the landscape level. The two tables below provide the name, type, source, reference period, resolution, description, and source of each variable.

All geospatial variables have been produced using the unmodified GPS data. Efforts have been made to correct obvious data entry errors, but this is not always possible. A variable of *qa_type* and *plot_qa_type* is included in the household and field-level geovariable files as an indicator of coordinate reliability, with a value of "5" indicating that the location falls outside expected range of values (more than 10 km from the EA centerpoint for households and more than 10 km from household for fields).

Pub_ETH_PlotGeovariables_Y1

The field-level geo file Pub_ETH_PlotGeovariables_Y1 contains four geospatial variables measuring field distance to household, slope, elevation and potential wetness index for field locations. The observations are uniquely identified by the combination of holder_id, household_id, parcel_id and field_id. The observations included in this file are fields that are owned and/or cultivated by the household and that have been visited for GPS-based land-area measurement.

Pub_ETH_HouseholdGeovariables_Y1

The household-level geo file Pub_ETH_HouseholdGeovariables_Y1 contains a range of variables measuring (on the basis of the household dwelling) distance to main points, climatology, landscape typology, soil and terrain, and crop season parameters. The observations are uniquely identified by the ERSS household id.

Additionally, modified coordinates are provided to satisfy the user interest in geo-referenced location while preserving the confidentiality of sample household and communities. It provides the average of household GPS coordinates within each EA, applied a random offset within a specified range to the average EA value (following the MeasureDHS approach).

More specifically, the coordinate modification strategy relies on random offset of EA center-point coordinates (or average of household GPS locations by EA in ERSS) within a specified range determined by the small town and rural classification. For small towns, an offset range of 0-2 km is used. Additionally, an offset range of 0-10 km is applied to 1% of EAs, effectively increasing the known range for all points to 10 km while introducing only a small amount of noise. Offset points are constrained at the zone level, so that they still fall within the correct zone for spatial joins, or point-in-polygon overlays. The result is a set of coordinates, representative at the EA level, that fall within known limits of accuracy. Users should take into account the offset range when considering different types of spatial analysis or queries with the data. Analysis of the spatial relationships between locations in close proximity would not be reliable. However, spatial queries using medium or low resolution datasets should be minimally affected by the offsets.

Table A.1 Field-level geo-referenced data linked to the ERSS survey data

Theme	Source	Dataset Title	Variable Name in Stata file	Variable Type	Reference Period	Resolution	Description	Web
Distance	LSMS-ISA	Field Distance to Household	dist_household	Continuous	N/A	N/A	Field distance to household	
Soil & Terrain	NASA	SRTM 90m	plot_srtm	Continuous	N/A	0.000833 dd	Elevation (m)	ftp://xftp.jrc.it/pub/srtmV4/arcasci/
	USGS	Slope (percent)	plot_srtmslp	Continuous	N/A	0.000833 dd	Derived from unprojected 90m SRTM using DEM Surface Tools	
	AfSIS	Topographic Wetness Index	plot_twi	Continuous	N/A	0.000833 dd	Downloaded from AfSIS website. Derived from modified 90m SRTM. Local upslope contributing area and slope are combined to determine the potential wetness index: WI = ln (A s / tan(b)) where A s is flow accumulation or effective drainage area and b is slope gradient.	http://www.ciesin.columbia.edu/afsis/bafsis_fullmap.htm#

TableA. 2 Household-level geo-referenced data linked to the ERSS survey data

Theme	Source	Dataset Title	Variable Name in Stata	Variable Type	Reference Period	Resolution	Description	Web
	Ethiopian Roads Agency	Household Distance to Main Road	dist_road	Continuous	2005	N/A	Household distance to nearest major road (ERA functional classes 'Trunk', 'Main Access', 'Link')	
	Central Statistics Agency / CityPopulation	Household Distance to Towns	dist_popcenter	Continuous	2007	N/A	Household distance to nearest town of >20,000 based on 2007 Census	http://www.citypopulation.de
	USAID FEWSNET	Household Distance to Key Market Centers	dist_market	Continuous	N/A	N/A	Household distance to nearest major market (FEWSNET key market centers)	
	Tracks4Africa	Household Distance to Border Posts	dist_borderpost	Continuous	N/A	N/A	Border control posts from PADKOS database, plus other border crossings on major road	
	Central Statistics Agency / CityPopulation	Household Distance to Regional Capital	dist_admctr	Continuous	N/A	N/A	Household distance to to the capital of the region of residence	http://www.citypopulation.de
Climatology	UC Berkeley	WorldClim Bioclimatic Variables	af_bio_1	Continuous	1960-1990	0.008333 dd	Average annual temperature calculated from monthly climatology, multiplied by 10 (°C)	http://www.worldclim.org/bioclim
	UC Berkeley	WorldClim Bioclimatic Variables	af_bio_8	Continuous	1960-1990	0.008333 dd	Average temperature of the wettest quarter, from monthly climatology, multiplied by 10. (°C)	http://www.worldclim.org/bioclim
	UC Berkeley	WorldClim Bioclimatic Variables	af_bio_12	Continuous	1960-1990	0.008333 dd	Total annual precipitation, from monthly climatology (mm)	http://www.worldclim.org/bioclim
	UC Berkeley	WorldClim Bioclimatic Variables	af_bio_13	Continuous	1960-1990	0.008333 dd	Precipitation of wettest month, from monthly climatology (mm)	http://www.worldclim.org/bioclim
	UC Berkeley	WorldClim Bioclimatic Variables	af_bio_16	Continuous	1960-1990	0.008333 dd	Precipitation of wettest quarter, from monthly climatology (mm)	http://www.worldclim.org/bioclim

Landscape Typology	ESA and UC Louvain	GlobCover v 2.3	fsrad3_lcmaj	Categorical	2009	0.002778 dd	Majority landcover class within approximately 1km buffer	http://ionia1.esrin.esa.int/
	ESA and UC Louvain	GlobCover v 2.3	fsrad3_agpct	Continuous	2009	0.002778 dd	Percent under agriculture within approx 1 km buffer	http://ionia1.esrin.esa.int/
	IFPRI	IFPRI standardized AEZ based on elevation, climatology	ssa_aez09	Categorical		0.008333 dd	Agro-ecological zones created using WorldClim climate data and 0.0833dd resolution LGP data from IIASA.	http://harvestchoice.org/production/biophysical/agroecology
Soil & Terrain	NASA	SRTM 90m	Srtm	Continuous		0.000833 dd	Elevation (m)	ftp://xftp.jrc.it/pub/srtmV4/arcasci/
	AfSIS	Topographic Wetness Index	Twi	Continuous		0.000833 dd	Downloaded from AfSIS website. Derived from modified 90m SRTM. Local upslope contributing area and slope are combined to determine the potential wetness index: WI = ln (A s / tan(b)) where A s is flow accumulation or effective drainage area and b is slope gradient.	http://www.ciesin.columbia.edu/afsis/bafsis_fullmap.htm#
	LSMS-ISA	Terrain Roughness	srtm_5_15	Categorical		0.000833 dd	Derived from 90m SRTM using 15 Meybeck relief classes and 5x5 pixel neighborhood	
	FAO	Harmonized World Soil Database	SQ1	Categorical		0.083333 dd	Nutrient availability	http://www.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/
	FAO	Harmonized World Soil Database	SQ2	Categorical		0.083333 dd	Nutrient retention capacity	http://www.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/
	FAO	Harmonized World Soil Database	SQ3	Categorical		0.083333 dd	Rooting conditions	http://www.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/
	FAO	Harmonized World Soil Database	SQ4	Categorical		0.083333 dd	Oxygen availability to roots	http://www.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/
	FAO	Harmonized World Soil Database	SQ5	Categorical		0.083333 dd	Excess salts	http://www.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/

	FAO	Harmonized World Soil Database	SQ6	Categorical		0.083333 dd	Toxicity	http://www.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/
	FAO	Harmonized World Soil Database	SQ7	Categorical		0.083333 dd	Workability (constraining field management)	http://www.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/
Crop Season Parameters	NOAA CPC	Rainfall Estimates (RFE)	anntot_avg	Continuous	2001-2011	0.1 dd	Average 12-month total rainfall (mm) for Jan-Dec	ftp://ftp.cpc.ncep.noaa.gov/fews/newalgo_est_dekad/
	NOAA CPC	Rainfall Estimates (RFE)	wetQ_avg	Continuous	2001-2011	0.1 dd	Average total rainfall in wettest quarter (mm) within 12-month periods from Jan-Dec	ftp://ftp.cpc.ncep.noaa.gov/fews/newalgo_est_dekad/
	NOAA CPC	Rainfall Estimates (RFE)	wetQ_avgstart	Continuous	2001-2011	0.1 dd	Average start of wettest quarter in dekads 1-36, where first dekad of Jan =1	ftp://ftp.cpc.ncep.noaa.gov/fews/newalgo_est_dekad/
	NOAA CPC	Rainfall Estimates (RFE)	h2011_tot	Continuous	2011	0.1 dd	12-month total rainfall (mm) in Jan-Dec, starting January 2011	ftp://ftp.cpc.ncep.noaa.gov/fews/newalgo_est_dekad/
	NOAA CPC	Rainfall Estimates (RFE)	h2011_wetQ	Continuous	2011	0.1 dd	Total rainfall in wettest quarter (mm) within 12-month periods starting January 2011	ftp://ftp.cpc.ncep.noaa.gov/fews/newalgo_est_dekad/
	NOAA CPC	Rainfall Estimates (RFE)	h2011_wetQstart	Continuous	2011	0.1 dd	Start of wettest quarter in dekads 1-36, where first dekad of January 2011 =1	ftp://ftp.cpc.ncep.noaa.gov/fews/newalgo_est_dekad/
	BU	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	eviarea_avg	Continuous	2001-2011	0.004176 dd	Average total change in greenness (integral of daily EVI values) within main (Meher) growing season, averaged by zone	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
	BU	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	evimax_avg	Continuous	2001-2011	0.004176 dd	Average EVI value at peak of greenness, averaged by zone	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005

	BU	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	grn_avg	Continuous	2001-2011	0.004176 dd	Average timing of onset of greenness increase in day of year 1-356, within main (Meher) growing season, averaged by zone	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
	BU	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	sen_avg	Continuous	2001-2011	0.004176 dd	Average timing of onset of greenness decrease in day of year 1-356, within main (Meher) growing season, averaged by zone	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
	BU	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	h2011_eviarea	Continuous	2011	0.004176 dd	Total change in greenness (integral of daily EVI values) within main (Meher) growing season of 2011, averaged by zone	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
		MOD12Q2 Land Cover Dynamics (PHENOLOGY)	h2011_evimax	Continuous	2011	0.004176 dd	EVI value at peak of greenness within main (Meher) growing season of 2011, averaged by zone	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
	BU	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	h2011_grn	Continuous	2011	0.004176 dd	Onset of greenness increase in day of year 1-356, within growing season of 2011, averaged by zone	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
	BU	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	h2011_sen	Continuous	2011	0.004176 dd	Onset of greenness decrease in day of year 1-356, within main (Meher) growing season of 2011, averaged by zone	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005

Appendix 2: Updates to the ERSS Household Data:

In February 2017, updates were made to the 2011-2012 ERSS data. As of February 2017, all data downloaded from the web site included the following updates. If the data that you are using does not include these updates, you should download a newer version of the data by going to the LSMS web site (see Appendix 4). The changes made to the data are:

1. Revision of survey weights
2. Inclusion of nonstandard unit conversion factors
3. Desk edits to household consumption data
4. Updated consumption aggregate file

A2.1 Revision of survey weights

An error was discovered in the original treatment of small town areas in the calculation of the survey weights. This led to underweighting of small town areas. The weights included in the revised data have been corrected. This change only affects the weights for small town enumeration areas; the weights for rural enumeration areas are unchanged.

A2.2 Inclusion of food nonstandard unit conversion factors

In the household food consumption section (Section 5a), the majority of food quantities are reported in standard units (e.g. kilograms, grams, liters, etc.). However, there are some cases where the quantities are reported in nonstandard units (e.g. number, roll, cup, etc.). Previously, no conversion factors were available to convert these nonstandard units into standard units. Included in the February 2017 update are a set of conversion factors that can be applied to the ERSS (ESS1) consumption module to convert all units to kilograms: *Food_CF_Wave1.dta*. In order to make the conversions, the data user simply has to merge the consumption module with the conversion factors data set on item code (*hh_s5aq00*) and relevant unit code (for example *hh_s5aq02_b*) and then multiply the relevant quantity by the included conversion factor (*mean_cf_nat*). The conversion factors were collected as a part of a nationwide market survey conducted prior to ESS3.

A2.3 Desk edits to household consumption data

Household Section 5a, the module on household food consumption in the last 7 days, underwent a slightly more rigorous cleaning process at headquarters. This section was reviewed for unreasonable values and, utilizing our team's field experience and familiarity with common data entry errors, we identify errors where we feel confident we can reasonably assume the nature of the error and the subsequent fix. Outliers that we cannot explain using basic assumptions about data entry errors are left unchanged. These changes only affect the food consumption section (*sect5a_hh_w1.dta*).

1. One common error found in recording quantities of a given item involves incorrectly specifying the corresponding unit; instead of specifying grams, the enumerator accidentally writes down (or the data entry person incorrectly inputs) kilograms. In order to identify these cases, we convert household quantities consumed, purchased, produced, and received as gifts into daily per capita quantities. This allows us to identify quantities that are biologically impossible to consume by one person in one day's time. After identifying these cases, we make the assumption that the corresponding unit should have been grams (or cubic centimeters in the case of Liters). After looking at the distribution of daily per capita quantities, we selected a cut-off point of 7.14 kgs or Ls/person/day (or 50 kgs/person each week). For observations where the quantity exceeded this amount (and the unit was kgs or liters), the unit was changed from kilograms to grams, or from liters to cubic centimeters.

> **Example**: A household's reported quantity of maize consumption is listed as 1,200 kg in the last 7 days. The household has 5 members, so this translates to a weekly per capita consumption of 240 kgs and a daily per capita consumption of 32.3 kg of maize. We change the 1,200 kg >>> 1,200 grams.

> **Resulting changes**:
>
> | Variable | # of observations |
> |---|---|
> | hh_s5aq02_b (unit of consumption) | 157 |
> | hh_s5aq03_b (unit of purchases) | 131 |
> | hh_s5aq05_b (unit of own production) | 67 |
> | hh_s5aq06_b (unit of gifts) | 13 |

2. The scenario described above can also occur in the opposite direction, i.e., a suspiciously small quantity of consumption reported in grams, should likely have been reported in kgs. Here, our basis for assuming an error is not founded on biologically impossible amounts of consumption, but rather we assume that households would not register and/or report consumption of 5 g of maize, for example. For this scenario, we selected a lower cut-off of 1 g per person/per week. For observations where reported consumption is less than this amount (and the unit was grams), the unit was changed from grams to kilograms.

> **Example**: A household's reported quantity of teff consumption is listed as 6 g in the last 7 days. The household has 7 members, so this translates to a weekly per capita consumption of 0.86 grams of teff. We change the 6 g >>> 6 kgs.

> **Resulting changes**:
>
> | Variable | # of observations |
> |---|---|
> | hh_s5aq02_b (unit of consumption) | 61 |
> | hh_s5aq03_b (unit of purchases) | 0 |
> | hh_s5aq05_b (unit of own production) | 50 |
> | hh_s5aq06_b (unit of gifts) | 5 |

3. This correction expands on #2, asserting that the lower threshold selected above can be increased slightly for items that are not extremely light in weight. For observations where

reported consumption is less than 5 grams per person/week, AND the food item is not sugar or salt, we change the unit to kgs.

> **Example**: A household's reported quantity of horse bean consumption is listed as 10 g in the last 7 days. The household has 5 members, so this translates to a weekly per capita consumption of 2 grams of horse beans. We change the 2 g >>> 2 kgs.

> **Resulting changes**:
>
> | Variable | # of observations |
> |---|---|
> | hh_s5aq02_b (unit of consumption) | 50 |
> | hh_s5aq03_b (unit of purchases) | 15 |
> | hh_s5aq05_b (unit of own production) | 31 |
> | hh_s5aq06_b (unit of gifts) | 2 |

4. On the paper questionnaire, all quantity amounts can be entered up to the thousandths decimal place. In ESS1, the value was split into two cells: the digits to the left of the decimal and the digits to the right of the decimal. Therefore, it is possible for a data entry person to enter the value of '4' into the decimal cell, turning the value into '.004' (values are entered beginning with the right-most digit).

We identify observations where the quantity consumed is reported in kgs, but the value is less than 0.05. These values are not only implausibly small, but they are very likely errors resulting from the situation described above. For these cases, we multiply the quantity amount by 1,000.

> **Example**: A household reports consuming 20 kgs of wheat in the paper questionnaire:
>
> | | 2 | 0 | . | | | |
>
> The data entry person accidentally enters this value of '20' into the decimal entry cell:
>
> | | | | | 2 | 0 |
>
> Which is the equivalent of:
>
> | | 0 | . | 0 | 2 | 0 |
>
> We rectify this error by multiplying 0.02*1,000=20.0

> **Resulting changes**:
>
> | Variable | # of observations |
> |---|---|
> | hh_s5aq02_a (qty. consumed) | 75 |
> | hh_s5aq03_a (qty. purchased) | 0 |
> | hh_s5aq05_a (qty. produced) | 58 |
> | hh_s5aq06_a (qty. received as gifts) | 17 |

5. In wave 1, we identified a few cases where a household only reported consumption and purchase amounts (not production or gifts), the quantity purchased was exactly 1,000 times greater than the reported quantity consumed, and the corresponding units for purchase and consumption did not match this disparity (for example, hh_s5aq03_a= hh_s5aq02_a*1000, but it

was not the case that q03_b was grams and 02_b was kilograms). Furthermore, in these instances, the unit for purchase amounts did not make sense with the corresponding quantity and/or food item. Here, we assume that the value and unit entered in questions 2_a and 2_b are the correct ones.

Example: A household reports 25 kgs of salt consumption, 25,000 cups of purchases, and no production or gifts. We replace 25,000 cups of purchases with 25 kgs.

Resulting changes:

Variable	# of observations
hh_s5aq03_a (qty. consumed)	40
hh_s5aq03_b (qty. purchased)	36

A2.4 Updated consumption aggregate

The consumption aggregate file (***cons_agg_w1***) has been updated based on the desk edits made to the household consumption file as well as integration of the food conversion factors. Please see *ESS1_consumption_aggregate_documentation.pdf*[15] for the methodological details of the consumption aggregate.

---

[15] Available for download at http://go.worldbank.org/ZK2ZDZYDD0

Appendix 3: ERSS Household Data: Data Entry and Data Collection Errors in Anthropometric Section of the Health Module

individual_id	household_id	saq01 (Region)	hh_s3q23 (height in cm)	In the Questionnaire	Correct Value	Error Source
0101030180414508	01010301804145	Tigray	13	086	86	1
0102030100112807	01020301001128	Tigray	1	094	94	1
0102070010102706	01020700101027	Tigray	1.9	1.09	109	3
0102070010105703	01020700101057	Tigray	1	0.80	80	3
0102070010108704	01020700101087	Tigray	1	78	78	1
0102070010111704	01020700101117	Tigray	0.92	0.92	92	2
0102070010135704	01020700101357	Tigray	1	99	99	1
0104010120411803	01040101204118	Tigray	1	89	89	1
0104030040402706	01040300404027	Tigray	0.74	0.74	74	2
0203060060109810	02030600601098	Afar	1.3	1.35	135	2
0203060060109812	02030600601098	Afar	1	1.00	100	2
0301010020406102	03010100204061	Amhara	1	1.14	114	3
0301140081613203	03011400816132	Amhara	6	65	65	1
0304090120313405	03040901203134	Amhara	1	0.79	79	3
0304120010504206	03041200105042	Amhara	1.5	1.05	105	2
0304140210100604	03041402101006	Amhara	8	0.77	77	3
0307010140602204	03070101406022	Amhara	14	73	73	1
0309010120200107	03090101202001	Amhara	1.2	1.02	102	3
0309010120201105	03090101202011	Amhara	1	0.76	76	3
0310040040810005	03100400408100	Amhara	11.1	11.1		2
0310040040810006	03100400408100	Amhara	7.1	7.1		2
0403020180313305	04030201803133	Oromia	9.5	9.5	95	2
0404120010402905	04041200104029	Oromia	5	0.85	85	3
0405130240500205	04051302405002	Oromia	0.93	0.93	93	2
0405130240502105	04051302405021	Oromia	1.11	1.11	111	2
0405130240504103	04051302405041	Oromia	1	1.00	100	2
0405130240506205	04051302405062	Oromia	0.98	0.98	98	2
0405130240514607	04051302405146	Oromia	0.9	0.9	90	2
0405130240516504	04051302405165	Oromia	1.15	1.15	115	2
0405130240516505	04051302405165	Oromia	0.84	0.84	84	2
0406080190201704	04060801902017	Oromia	1.05	1.05	105	2
0408120240100403	04081202401004	Oromia	0.8	0.89	89	3
0408120240102207	04081202401022	Oromia	5	1.00	100	3
0408120240102208	04081202401022	Oromia	0.9	0.90	90	3
0408120240104003	04081202401040	Oromia	0.9	0.96	96	3
0410090360513007	04100903605130	Oromia	1	1	100	2
0411040130402204	04110401304022	Oromia	1.8	1.08	108	3
0413090100301906	04130901003019	Oromia	1	107	107	1
0413090100303409	04130901003034	Oromia	1	86	86	1
0413090100304909	04130901003049	Oromia	1	104	104	1
0413090100307908	04130901003079	Oromia	1	100	100	1
0413090100309405	04130901003094	Oromia	1	83	83	1
0419040010108007	04190400101080	Oromia	1	103.7	104	1
0419040010108008	04190400101080	Oromia	1	68.3	68	1
0501010090206107	05010100902061	Somalie	1.17	1.17		2

individual_id	household_id	saq01 (Region)	hh_s3q23 (height in cm)	In the Questionnaire	Correct Value	Error Source
0501020010100405	05010200101004	Somalie	1	0.89	89	3
0501020010100406	05010200101004	Somalie	1	0.65	65	3
0501030110100203	05010301101002	Somalie	1	1.80	.	2
0501030110105206	05010301101052	Somalie	1	0.73	73	3
0502020010203208	05020200102032	Somalie	10	10.7		3
0502060380300303	05020603803003	Somalie	9.8	9.8		2
0502060380300304	05020603803003	Somalie	7.1	7.1		2
0502060380300703	05020603803007	Somalie	8.1	8.1		2
0502060380301503	05020603803015	Somalie	9.1	9.1		2
0502060380302104	05020603803021	Somalie	9.7	9.7		2
0502060380302105	05020603803021	Somalie	8.5	8.5		2
0502060380302707	05020603803027	Somalie	7.3	7.3		2
0502060380303303	05020603803033	Somalie	6.3	6.3		2
0502060380304504	05020603803045	Somalie	7.3	7.3		2
0502070480100503	05020704801005	Somalie	1.2	1.2		2
0502070480100504	05020704801005	Somalie	1.1	1.19	119	3
0502070480102303	05020704801023	Somalie	0.7	0.77	77	3
0502070480107705	05020704801077	Somalie	1.7	1.7	107	3
0502070480111503	05020704801115	Somalie	0.8	0.89	89	3
0502070480113303	05020704801133	Somalie	1.8	1.08	108	3
0502070480113304	05020704801133	Somalie	0.9	0.96	96	3
0502070480115104	05020704801151	Somalie	0.9	0.98	98	3
0502070480116903	05020704801169	Somalie	1.2	1.02	102	3
0502070480116904	05020704801169	Somalie	0.8	0.89	89	3
0502070480116905	05020704801169	Somalie	0.7	0.72	72	3
0509020120100105	05090201201001	Somalie	10.8	10.8		2
0701110040100806	07011100401008	SNNP	1.1	111	111	1
0701110040100807	07011100401008	SNNP	0.9	"091"	91	1
0707020290415003	07070202904150	SNNP	14	.	.	1
0708030020616507	07080300206165	SNNP	1.1	1.01	101	3
0710090170314606	07100901703146	SNNP	0.72	0.72	72	2
0710090170316406	07100901703164	SNNP	1.3	1.03	103	3
0710090170316407	07100901703164	SNNP	0.75	0.75	75	2
0720010120303406	07200101203034	SNNP	1	100	100	1
0721010620219506	07210106202195	SNNP	1.2	1.20	120	2
1202030040102707	12020300401027	Gambelia	10.2	10.2		2
1202030040102708	12020300401027	Gambelia	9	9		2
1301010060104007	13010100601040	Harari	1.5	1.05	105	3

Note:

1. Source of Error: 1=Data Entry Error; 2= Data Collection Error; 3=Both (Data Collection and Entry Error)
2. The outliers were first flagged from the entered data. The errors were then identified by reviewing the questionnaires. The outliers were due to data entry error, data collection error or both. The data entry errors were those errors that the data entry clerk missed to enter what was in the questionnaire. The data collection errors are those errors that the data collector used meters instead of centimeters. Data collection and data entry errors

(both) are those wrongly recorded errors by the data collector were wrongly keyed by the data entry clerk.

Appendix 4: How to Obtain Copies of the Data

The data are available through the CSA web site:

http://www.csa.gov.et/

or through the LSMS-ISA website:

http://www.worldbank.org/lsms-isa

Users do not need to obtain the permission of the CSA to receive a copy of the data, but will be asked to fill in a data access agreement. In this agreement, users agree to: (a) cite the Central Statistical Agency of Ethiopia as the collector of the data in all reports, publications and presentations; (b) provide copies of all reports publications and presentation to the Central Statistical Agency (see address below) and the Poverty and Inequality Division of the World Bank (see address below); and (c) not pass the data to any third parties for any reasons.

The Director General	LSMS Database Manager
Central Statistical Agency	Poverty and Inequality Division
Addis Ababa, Ethiopia	The World Bank
www.csa.gov.et	1818 H Street, NW
Phone: +251-111-553011	MSN MC3-306
	Washington, DC 20433
	www.worldbank.org/lsms-isa
	Email: lsms@worldbank.org