Ethiopia Socioeconomic Survey (ESS)
Wave Two (2013/2014)

Basic Information Document

Central Statistical Agency
&
Living Standards Measurement Study (LSMS), World Bank
March 2015

ACRONYMS

AgSS	Annual Agricultural Sample Survey
BMGF	Bill and Melinda Gates Foundation
CSA	Central Statistical Agency (Ethiopia)
DHS	Demographic and Health Survey
EA	Enumeration Area
EIAR	Ethiopia Institute of Agricultural Research
ERSS	Ethiopia Rural Socioeconomic Survey
ESS	Ethiopia Socioeconomic Survey
HCES	Household Consumption and Expenditure 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
WB	World Bank
WFP	World Food Program

Table of Contents

1. Introduction

The Ethiopian Socioeconomic Survey (ESS) 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) project. [1] The objective of the LSMS-ISA is to collect multi-topic panel household level data with a special focus on improving agricultural statistics and the link between agriculture and other household income activities. The idea of the ESS 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 ESS, 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. The LSMS-ISA, 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 basis of analyses looking for insights into how innovation and efficiency can be fostered in the agriculture sector.

ESS began as ERSS (Ethiopia Rural Socioeconomic Survey) in 2011/12. The first wave of data collection in 2011/12 included only rural and small town areas. The survey name dropped the word "Rural" in the second wave of data collection when the sample was expanded to include all urban areas. The urban supplement was done in such a way to ensure that the ESS wave 2 data can provide nationally representative estimates. Accordingly, the number of enumeration areas (EAs) covered by the survey increased from 333 (or 3,776 households) to 433 (or 5,262 households). For the rest of this document, ESS will refer generally to the survey, ESS1 will refer to the first wave of the ESS carried out in 2011/12,[2] and ESS2 will refer to the recent second wave of the ESS carried out in 2013/14. ESS1 and ESS2 together create a panel data set of households from rural and small town areas (ie. the same households that were interviewed in ESS1 were tracked and re-interviewed in ESS2). ESS2 also serves as a nationally representative sample because the sample of rural and small town households was expanded to include all urban areas as well.

---

[1] For more information on the LSMS and LSMS-ISA go to www.worldbank.org/lsms
[2] In other words, ESS1 is the new term now used to refer to the ERSS.

Several innovative approaches were incorporated in the ESS:

♦ 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;

♦ Implementing concurrent data entry using a decentralized data entry arrangements;

♦ 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;

♦ Tracking of households who have left their original location for other places in the country;

♦ Active dissemination of agriculture statistics.

ESS2 covered all regional states including the capital, Addis Ababa. The majority of the sample comprises rural areas as it was carried over from ESS1. The ESS2 was implemented in 433 enumeration areas (EAs) out of which 290 were rural, 43 were small town EAs from ESS1, and 100 were new EAs from major urban areas.

ESS is a long-term project to collect panel data. The project responds to the data needs of the country, given the dependence of a high percentage of households in agriculture activities in the country. The ESS collects 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 ESS 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 ESS is the first panel survey to be carried out by the CSA that links a multi-topic household questionnaire with detailed data on agriculture.

The purpose of this document is to provide detailed information on the second wave of the survey fielded by the CSA in 2013-14. The survey consists of three rounds of visits to the household. The first round of the ESS2 was carried out in September and October 2013 and collected information on post-planting agriculture activities. The second round was conducted in November-December 2013 and fielded the livestock questionnaire to collect information on ownership, production and utilization of livestock, and livestock by products. The third round took place from February-April 2014 to collect information for the post-harvest agriculture, household, and community questionnaires.

2. The Survey Instruments

The survey consisted of five questionnaires. These questionnaires are the same as those used in the ESS1 with revisions based both on the results of the ESS1 and also on identified areas of need for new data (see Section 7). The ***household questionnaire*** was administered to all households in the sample. The ***community questionnaire*** was administered to a group of community members 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 ESS data to other available geographic data sets (See Appendix 1 for discussion of the geo-data provided with the ESS).

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 questionnaires 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: Household Questionnaire

Section	Topic	Respondent	Description	New questions added in the second wave
Cover	Cover	Field staff	Household location identification; household size and head's name; field staff identification; and observation notes by enumerator regarding the interview.	panel or new household (12a), CSA branch code (34)
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.	Age/sex verification questions (4b, 4c, 4d, 4e, 4f, 4g, 4h), migration questions from 22-37
2	Education	Individuals 5 years and above	Educational attainment, enrollment, attendance, school characteristics,	-

---

[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.

Section	Topic	Respondent	Description	New questions added in the second wave
			and expenditures for the 2013-14 academic year. [4]	
3	Health	Individuals	Health problems, types of injury/illness, medical assistance/consultation, disabilities, and anthropometrics (children under seven years). [5]	Child filter for disability (12a), Additional child health questions (20a, 20b, 20c, 20d, 20e, 20f)
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.	Kocho and Bulla- asked separately (17a & 17b)
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.	Meal away from home (6 &7: meal types A-H)
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.	Reference period for Q6 changed to two years
9	Housing	Household head or eligible adult	Dwelling ownership, and characteristics of the dwelling and utilities.	Additional questions on utilities: water (13a, 14a, 15a) & electricity (20a, 20b)
10	Household assets	Household head or eligible adult	Household ownership of assets.	Jewels split into gold and silver (item #24a,b)
11	Non-farm Enterprises	Owner or manager of enterprise	Characteristics of enterprises owned by the household.	Panel filter (8a), verification (1a-1c), respondent and decision maker (3b-3d), credit/loan for the NFE (4b-4h)
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. (Confidential-not included in the public data)	-

---

[4] The 2013-14 school year started in September 2013 and ended in July 2014.

[5] First wave measured children under five.

Table 2.2: Community Questionnaire

Section	Topic	Respondent	Description	New questions added in the second wave
Cover (1.1&1.2)	Cover	Field staff	Community location identification; field staff identification; date and time of interviews.	Branch office code (26)
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.	Questions on who undertakes various agriculture activities in community (7b, 7c, 9b); reference period changed from five to two years for 16, 17, and 18
7	Changes	Informants	Important events in the community in the last five years.	Reference period for section changed from five to two 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.	Q4 has been modified: Past 2 years (instead of 5 years)
9	Productive Safety nets Program	Informants	Participation in the productive safety nets program. Management and performance of the program in the community.	-
10A	Market Prices	Sellers in the nearby market-Market center 1	Market prices in the first closest market center.	Section reorganized to clearly capture unit & unit code
10B	Market Prices	Sellers in the nearby market-Market center 2	Market prices in the other closest market center.	Section reorganized to clearly capture unit & unit code

Table 2.3: Post-planting Agriculture Questionnaire[6]

Section	Topic	Respondent	Description	New questions added in the second wave
Cover	Cover	Field Staff	Holder[7] location identification; household head name, holder name, household size, agriculture holding type (farming, livestock, or both); field staff identification.	Holder's education level (13a), CSA branch office code (35)
1	Household Roster	Household head or eligible adult	Name, age, and gender of each household member and holding type (farming, livestock or both).	-
2	Parcel Roster	Owner or manager of the parcel	Information on all parcels owned and/or managed by the holder:	respondent verification (1c &1d), rights over the parcel (3b, 3c & 8a), soil quality (14 & 15)
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, labor inputs, and other details on the fields.	Question 3 reorganized now 3 and 3b, Fallowing practices (3c&3d), Reason for not having GPS/ Rope and Compass record (8c), Decision maker on planting and crop choice (10a, 10b, 10c); additional fertilizer acquisition questions (16b-d for Urea and 19b-d for DAP and 20a-d for inorganic fertilizer).
4	Crop Field Roster	Manager of field	Crop planting/management information for each crop on each field.	Quantity of seed (11b&11c), type of crop sowing technique (12b).
5	Seeds Roster	Manager of field	Seed related information for each crop planted on each field.	Section reorganized. Presented by seed type. 0a, 0b and 1b new questions on seed type and crop code.
7	Miscell-aneous	Manager of field	Information on holder characteristics including chemical fertilizer use, and access and use of credit, extension/advisory services.	Fertilizer questions revised and new questions added: (16a-c), watershed management (17-20), gender focused questions on participation in agriculture activities (20-23).
NR	Network Roster	Manager of field	Roster of individuals, places or businesses from/where the holder purchases agricultural inputs.	

---

[6] There is no Section 6 in the post-planting questionnaire.

[7] 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.

Table 2.4: Post-harvest Agriculture Questionnaire

Section	Topic	Respondent	Description	New questions added in the second wave
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/	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). This section excludes permanent, tree and root crops. Crop cut only for 23 crop types.	Note: Crop cut size is 4mX4m (not 2mX2m)
9	Crop Harvest by Field	Holder/ Manager of field	All crops- harvest information. Information on crop use, area harvested, amount harvested and damage to crops.	Section reorganized during the second wave. All the questions in this section are either new or modified. In the first wave, some of the questions (e.g. crop damage) were asked together with the crop cut.
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.	Decision maker for disposition (1b&1c), quantity consumed (10a)
12	Crop Disposition / Sales (Tree / Permanent and Root Crops)	Holder/ Manager of field	Harvest information on permanent, tree, and root crops.	Additional question on tree crops harvest (3b), decision on earnings from sale (8a), quantity consumed (10a)
NR	Network Roster	Manager of field	Roster of individuals, places or businesses to/where the holder sells crops.	-

Table 2.5: Livestock Questionnaire

Section	Topic	Respondent	Description	New questions added in the second wave
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.	Holder's education level (13a), branch office code (35)
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.	Herding responsibilities (12b), livestock selling decisions and earnings (60b, 60c&60d)
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.	Byproduct selling decisions and earnings (7b,7c &7d)

3. Sample Design

3.1 Wave 1 coverage, Rural and small towns (ESS1)

The ESS sample coverage and design changed over the first two waves of data collection. In wave 1, the ESS1 sample was designed to be representative of rural and small town areas in Ethiopia.[8] The ESS sample is drawn from a population frame that includes all areas of Ethiopia except for three zones of Afar and six zones of Somalie region.[9] The frame for rural areas is the 2011/2012 Agricultural Sample Survey (AgSS), so the ESS rural sample is a complete subset of the AgSS sample. The small town and urban area samples come from the universe of small town and urban EAs, excluding the same three zones of Afar and six zones of Somalie. The wave 1 sample design provides representative estimates for all rural-area households and for the combination of rural-area and small-town households (excepting the 9 zones excluded from the frame).

The wave 1 sample design is a stratified, two-stage design where the regions of Ethiopia serve as the strata. Quotas were set for the number of EAs in each region to ensure a minimum number of EAs are drawn from each EA. The data is representative at the regional level for the most populous regions: Amhara, Oromiya, SNNP, and Tigray. The sample size is insufficient to support region-

---

[8] 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.

[9] Zones excluded from the sampling frame include Zones 2, 4, and 5 in the Afar Region and Zones 3, 4, 5, 6, 7, and 8 in the Somali Region.

specific estimates 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. For the purpose of sub-national anlaysis, there are 5 representative domains of analysis: Amhara, Oromiya, SNNP, Tigray and all other regions combined as one category.

3.2 Wave 2 coverage, Rural and all urban areas (ESS2)

In part due to the success of the first wave of data collection, the Government of Ethiopia decided to invest in an urban sample supplement to the ESS1 sample. The key objective of the urban expansion is to ensure that the ESS2 will be able to provide inferences to all of Ethiopia, and not just rural areas and small towns. To reach this objective, the ESS1 sample was supplemented with an urban sample that was drawn independently from a stratified urban frame of households.

A secondary objective of the expansion was to ensure that one can also make inferences to the population of urban households. This requires both that the urban sample is a separate stratum and also that it is of sufficient size to provide adequate precision for the urban estimates. In order to be more specific, adequate precision was set to mean that the estimated relative standard error (RSE) of the mean of consumption in urban areas is approximately the same size as the RSE for the mean of consumption as estimated in ESS1.[10] This rule helped guide the size of the urban supplement.

Prior to expanding the sample, CSA and the World Bank team examined potential options for expansion based on the criteria of providing approximately the same level of precision with the urban estimates of consumption as is obtained with the rural estimates of consumption. To assess the needed sample size, the team used urban estimates of the overall variance and within cluster variance of consumption from the national Household Consumption and Expenditure Survey (HCES). The level of within cluster variance in urban areas as estimated by the HCES is significantly less than estimates of within cluster variance in rural areas. An implication of this is that a relatively small urban supplement would be adequate to provide sufficient power to meet the objectives.[11] After careful consideration of the power analysis and budget constraints, CSA decided on supplementing the wave 1 sample with 1,500 urban households.

The population frame for the urban expansion consists of all households in towns with population greater than 10,000 people. This population cut off of 10,000 people is the same threshold that is used to define small towns.[12] In addition to the existing sample of rural and small towns, this expansion means now that all households in Ethiopia have some positive probability of selection into the expanded ESS2 sample.[13]

---

[10] The RSE for consumption is 5 percent in wave 1, based on rural and small towns combined and treating outliers with both a 1 percent and 2 percent symmetric trim.

[11] Under some assumptions, the required sample size was as low as 720 households.

[12] 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.

[13] There is an exception to this statement. The ESS and the AgSS surveys both exclude 9 zones from the population frame -- zones 2, 4, and 5 in the Afar Region and zones 3, 4, 5, 6, 7, and 8 in the Somali Region.

In order to correspond with the existing, wave 1 design while ensuring that all urban areas are included, the population frame was stratified to ensure that the ESS will be able to provide population inferences for the same five domains as in ESS1 plus an additional domain for the city state of Addis Ababa.[14] Specifically, the 6 strata are: Addis Ababa, Amhara, Oromiya, SNNP, Tigray, and "other regions" (including Dire Dawa). In those regions with the greatest urban populations, Addis Ababa and Oromiya, 20 EAs were selected; while in all other strata, 15 EAs were selected.

In an effort to improve efficiency, the frame was further stratified based on town size. While we did not have empirical evidence of important types of heterogeneity across big and medium cities, the assumption held was that big cities look quite different from medium cities and therefore stratification by size ensures coverage across these two types of urban areas. To this end, all strata except for the city state of Addis Ababa were stratified in medium-sized (population between 10,000 and 100,000) and big-sized (greater than 100,000) towns. Table 3.1 below provides the allocation of EAs in accord with the sample design for the wave 2 urban expansion.

Table 3.1: Urban expansion design, stratification, and allocation of EAs (n=1500)

Strata = Domains of Analysis	Sub-stratification By city size	EAs (prob. proportional to population)	Households (Systematic, Fixed-interval random)
Addis Ababa		20	300
Amhara	Big || Medium	8 || 7	225
Oromiya	Big || Medium	10 || 10	300
SNNP	Big || Medium	8 || 7	225
Tigray	Big || Medium	8 || 7	225
All other regions	Big || Medium	8 || 7	225
Total		*100*	*1500*

Note: By design, 15 households per EA are selected. Medium is defined as a town with population between 10,000 and 100,000. Big is defined as a town with population greater than 100,000.

3.3 ESS Design Details, ESS1 and ESS2

The ESS sample is a two-stage probability sample. The first stage of sampling entailed selecting enumeration areas (ie. the primary sampling units) using simple random sampling (SRS) from the sample of the AgSS enumeration areas (EAs). The AgSS EAs were selected based on probability proportional to size of population (PPS). For the rural sample, 290 EAs were selected from the AgSS EAs. For small town EAs, a total of 43 EAs were selected by PPS. Similarly, for the newly added urban areas, a total of 100 EAs were selected (PPS). Table 3.2 shows how the EAs of the ESS1 (rural and small town EAs) and ESS2 (rural, small town, mid and large towns) are spread across regions of Ethiopia.

---

[14] The city state of Dire Dawa already appears in the existing sample design.

Table 3.2: First Stage Sample Design

	Total EAs	Rural EAs	Urban	
			Small town EAs	Mid and Large Town EAs
National	433	290	43	100
Regions				
Tigray	49	30	4	15
Afar	13	10	2	1
Amhara	86	61	10	15
Oromiya	85	55	10	20
Somali	26	20	3	3
Benishangul-Gumuz	11	10	1	0
SNNP	99	74	10	15
Gambela	12	10	1	1
Harari	14	10	1	3
Dire Dawa	18	10	1	7
Addis Ababa	20	NA	NA	20

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 (SRS) 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 (SRS) from the listing of each EA, with no stratification as to whether the household is engaged in agriculture/livestock. The final number of household interviewed in wave 1 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 percent. Of these 3,969 households in ESS1, ESS2 successfully re-interviewed 3,776 households. This implies a panel attrition rate of 5 percent, or successful follow-up rate of 95 percent.

As with the rural sample, the urban sample was selected following a multi-stage, clustered design. In the second stage of selection, 15 households were selected (SRS) in the mid- and large-size town EAs. Due to nonresponse and fieldwork issues, the final number of households interviewed was slightly less than the 1,500 urban households as planned in the design. A total of 1,486 households were interviewed with a response rate of 99 percent. The distribution of the final ESS1 sample of households in rural and small towns, and the ESS2 sample of households in rural and all urban areas is shown in Table 3.3.

Table 3.3: Distribution of Final Sample for ESS
by Region and Zone in Rural and Urban areas

Region Code	Region Name	Zone Code	Zone Name	Rural		Small towns (urban)		Mid- and large-sized towns (urban)	
				EA	HH	EA	HH	EA	HH
1	Tigray	1	North Western	5	58	1	12	1	15
		2	Central	7	79	1	8	2	30
		3	East	6	71	1	11	2	30
		4	South	6	68	1	12	1	15
		5	West	6	69	0	0	1	15
		6	Mekelle Town	0	0	0	0	8	120
2	Afar	1	Zone-1	7	70	1	12	1	13
		3	Zone-3	3	33	1	8	0	0
3	Amhara	1	North Gonder	7	77	2	23	4	60
		2	South Gonder	6	67	1	10	1	15
		3	North Wollo	8	94	1	12	1	15
		4	South Wollo	6	67	1	9	2	30
		5	North Shewa	7	79	1	11	1	15
		6	East Gojam	6	70	1	10	1	15
		7	West Gojam	7	80	1	12	1	15
		8	Wag Himra	3	31	1	10	0	0
		9	Awi	5	58	1	9	1	15
		10	Ormiya	3	34	1	11	0	0
		11	Bahirdar	0	0	0	0	3	45
		12	Argoba Special Wereda	3	35	0	0	0	0
4	Oromiya	1	West Wellega	4	48	1	12	1	15
		2	East Wellega	3	34	1	12	0	0
		3	Ilu Aba Bora	3	35	1	11	1	15
		4	Jimma	4	47	1	9	0	0
		5	West Shewa	3	35	1	10	1	15
		6	North Shewa	3	35	1	8	0	0
		7	East Shewa	4	44	0	0	2	29
		8	Arsi	3	34	1	11	1	15
		9	West Hararge	4	48	1	11	1	15
		10	East Hararge	3	36	1	12	0	0
		11	Bale	4	45	1	12	1	15
		12	Borena	3	36	0	0	0	0
		13	South West Shewa	3	35	1	8	1	15
		14	Guji	3	34	0	0	0	0
		15	Adama town	0	0	0	0	6	90
		16	Jimma	0	0	0	0	2	30
		17	West Arsi	3	36	0	0	3	45
		18	QellemeWollega	2	24	0	0	0	0

Region Code	Region Name	Zone Code	Zone Name	Rural		Small towns (urban)		Mid- and large-sized towns (urban)	
				EA	HH	EA	HH	EA	HH
		19	HoroGudruWellega	3	35	0	0	0	0
5	Somali	1	Shinille	5	53	1	7	0	0
		2	Jijiga	8	88	1	10	3	45
		9	Liben	7	78	1	9	0	0
6	BenshangulGumuz	2	Metekel	3	33	0	0	0	0
		3	Assosa	3	36	1	12	0	0
		4	Kamashi	2	22	0	0	0	0
		6	Mao Komo	2	22	0	0	0	0
7	SNNP	1	Gurage	5	58	1	10	1	15
		2	Hadiya	4	48	1	12	1	15
		3	KembataTimbaro	3	36	0	0	0	0
		4	Sidama	6	72	1	11	1	14
		5	Gedeo	5	57	0	0	0	0
		6	Wolayita	4	46	1	11	2	30
		7	South Omo	3	36	1	9	0	0
		8	Sheka	3	33	0	0	0	0
		9	Kefa	4	48	1	11	0	0
		10	GamuGofa	5	57	1	11	1	15
		11	Bench Maji	5	56	1	11	1	15
		12	Yem	3	36	0	0	0	0
		13	Amaro Special	2	24	0	0	0	0
		14	Burji Special	2	24	0	0	0	0
		15	Konso Special	3	36	0	0	0	0
		16	Derashe	2	23	0	0	0	0
		17	Dauro	4	47	1	12	0	0
		18	Basketo	3	36	0	0	0	0
		19	Konta	2	23	0	0	0	0
		20	Silti	4	48	1	10	0	0
		21	Alaba	2	24	0	0	0	0
		22	Hawassa city	0	0	0	0	8	114
12	Gambella	1	Anuak	3	32	1	11	1	15
		2	Nuer	2	21	0	0	0	0
		3	Mejenger	4	45	0	0	0	0
		4	Etang Special	1	6	0	0	0	0
13	Harari	1	Zone 1	10	120	0	0	3	45
15	Dire Dawa	1	Zone 1	10	118	0	0	7	104
14	Addis Ababa	1	Akaki Kaliti Sub City	0	0	0	0	2	30
		2	Nefas Silk-Lafto Sub City	0	0	0	0	2	30
		3	Kolfe Keraniyo Sub City	0	0	0	0	3	43

Region Code	Region Name	Zone Code	Zone Name	Rural		Small towns (urban)		Mid- and large-sized towns (urban)	
				EA	HH	EA	HH	EA	HH
		4	Gulele Sub City	0	0	0	0	2	30
		5	Lideta Sub City	0	0	0	0	1	15
		6	Kirkos Sub City	0	0	0	0	2	30
		7	Arada Sub City	0	0	0	0	1	14
		8	Addis Ketema Sub City	0	0	0	0	2	30
		9	Yeka Sub City	0	0	0	0	3	45
		10	Bole Sub City	0	0	0	0	2	30
				290	**3,323**	**43**	**453**	**100**	**1,486**

4. Training, Data Collection, Tracking & Monitoring

4.1 Training, ESS2

Six training sessions were held. These included two training of trainers (TOT) (July 2013 and January 2014), two field staff (enumerator and supervisor) training sessions in August 2013 and January-February 2014, and two data entry training in October 2013 and March 2014. The TOT and the field staff training focused on the content of the questionnaires as well as practical applications in data collection and supervision. Data entry clerks were trained in Addis Ababa. All of the trainees had survey experience and most of them had participated in the first wave of the ESS.

The TOT sessions were held in Addis Ababa and were led by the LSMS-ISA team. The first TOT (in July 2013) lasted approximately two weeks, and the second TOT (in January) took about 10 days.

The first field staff training was led by CSA experts together with the LSMS team. The first field staff training was held in four locations: Adama, Bahirdar, Hawassa and Jimma, and lasted approximately for three weeks in August 2013.[15] the second field staff training took place for about three weeks from January 16- February 7th, 2014. It was held in three different locations: Bahirdar, Hawassa and Jimma.[16] The training discussed household, community, and post-harvest questionnaires for the entire sample. The training sessions discussed post-planting and livestock questionnaires as well as new fieldwork management methods introduced in the second wave, such as concurrent data entry and household tracking. The January 2014 training also included field staff for the new urban areas. The urban team was trained separately on all modules in one location in Hawassa.

4.2 Field Work Organization & Data Collection

The survey was designed to be implemented in three visits following the AgSS field schedule. For panel households (the rural and small town households in both ESS1 and ESS2), the first visit took place between September and October 2013. In this round, the post-planting agriculture questionnaire was administered. The second visit took place between November and December 2013 when the livestock questionnaire was administered. The third visit took place from February through April 2014 when the household, community and post-harvest agriculture questionnaires were administered.

For newly added urban households, all questionnaires were administered in one visit that took place between February and April 2014.

---

[15] The distribution of branches for the field training was as follows: (1) Adama, Asayita, Asebe Teferi, Debreberhan, Dire Dawa, Harar and Jijiga in Adama town, (2) Bahirdar, Dessie Gondar, Mekelle and Shire branches in Bahirdar town; (3) Arba Minch, Goba, Hawassa, Hosaena, Negelle and Sodo branches in Hawassa town, (4) Assossa, Gambella, Jimma and Nekemte branches in Nekemte town.

[16] Those who had taken the training in Adama in August went to Hawassa for the January training.

Table 4.1: Implementation Timeline

ESS Questionnaire	2013				2014		
	Sep	Oct	Nov	Dec	Feb	Mar	Apr
Post-planting agriculture questionnaire	■	■					
Livestock questionnaire			■	■			
Household, community, and post-harvest agriculture questionnaire					■	■	■
All modules in major urban EAs					■	■	■

Note: Agriculture interviews for urban areas didn't include crop-cut and land measurement by the enumerator. Agriculture information collected from those EAs in major urban areas was based on self-response.

The post-planting and post-harvest questionnaires collected information on the same fields and crops in two different visits.[17] While the post-planting questionnaire collected field and crop level information such as 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 asked of 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 the field was small, by compass and rope[18]. 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 individual-level sections, the information was collected from each individual, except for children, in which case the information was collected from a parent or knowledgeable adult in the household (see Table 2.1). The enumerator conducted the anthropometric measurements for children 6-83 months.

---

[17] Since post-planting and post-harvest information was collected in two different visits for panel households, there are approximately 200 plots with PP data but no PH data. These cases include households that moved away or passed away between visits, those plots that contained permanent fruits or tree crops not harvested during the year, and plots that were used for other purposes excluding cultivation.

[18] GPS estimates become less accurate for smaller fields; in these cases, enumerators were instructed to use rope and compass to measure the area.

The community questionnaire was collected from focus groups of community informants and through direct observation (Table 2.2). This questionnaire was administered by the field supervisors, rather than the enumerators. 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 were collected with the help of sellers in the market centers.

Resident enumerators were used to administer the household, agriculture (post-planting and post-harvest), and livestock questionnaires in rural and small-town EAs. The CSA assigned one resident enumerator for each of these EAs. The enumerator lived in the EA for the entire survey period from September 2013 to April 2014. 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. Temporary mobile enumerator teams were used for the mid- and large-size town EAs.

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 Tracking of First Wave Households

During the second wave, some of the households interviewed during the first wave were not found in their original location. Most of those who were in the country were tracked and interviewed.

The tracking protocol was as follows. If any member of the original household was found in the original dwelling, they were interviewed. If no household member from wave 1 was found in the dwelling, the household was tracked if it moved to a different location within the country. The household was not tracked if it left Ethiopia. If a household split, the part of the household that contained the wave 1 head of household was tracked. Table 4.2 presents the tracking information for households included in wave 1.

Table 4.2: Status of panel households: rural and small town areas

Location Status	Households	Interviewed	Not interviewed
Same Dwelling	3,512	3,448	64
Moved	457	328	129
Within same Kebele	*160*	*158*	*2*
Nearby Kebele	*75*	*51*	*24*
Not nearby Kebele, but within same Woreda	*63*	*43*	*20*
Outside Woreda, within same Region	*70*	*48*	*22*
Outside of Region	*38*	*19*	*19*
Other	*51*	*9*	*42*
Total	**3,969**	**3,776**	**193**

4.4 Fieldwork Monitoring and Evaluation

Routine supervision by CSA's field supervisors entailed the field-level coordination by all CSA branch offices. Branch level statisticians and supervisors who were assigned to this project conducted the routine supervision. The branch supervisors made three extended visits to the EAs between September, 2013 and April, 2014. 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 2013 when interviews with the Post-planting questionnaire were being conducted. The second visit was in February-April 2014 when the household, community, and post-harvest agriculture data were being collected.

5. Data Management and Description of Datasets

5.1 Data Management

5.1.1 Concurrent Data Entry

The completed paper questionnaires were keyed in the branch offices. The CSPro software was used to design the data entry programs. Near-concurrent data entry arrangement was employed. In this arrangement, completed questionnaires were sent to the branch offices for entry. Then the data files were checked by a Stata program for errors. The questionnaires were then sent back to the field for correction. This process of parallel data collection and entry continued throughout until the fieldwork was completed.

5.1.2 Second Data Entry

The questionnaires were entered for a second time after the concurrent data entry process was completed in the field.[19] The first and the second entry were compared to check if there were data entry errors. The comparison flagged a number of data entry errors. Most of the errors were incorrect values and misalignment of rows in sections with a roster structure. The errors were corrected by referring to the hardcopy of the questionnaires.

5.1.3 Final Data Cleaning

Final data cleaning was carried out on all data files. Only errors that could be clearly and confidently fixed by the team were corrected; errors that had no clear fix were left in the datasets. Cleaning methods for these errors are left up to the data user.

5.1.4 Weighting of Data

The ESS2 data needs to be weighted to represent the national-level population of rural, small and large town households. A sample weight with post-stratification adjustments was calculated for the households and this weight variable is included in all the datasets. It reflects the adjusted probability of selecting the household into the sample. The inverse of this weight can be considered an expansion factor that sums to the total population of households in the nation. When this weight is used in a household-level file, it sums to the population of households. When this weight is used in an individual-level file, it sums to the population of individuals. If the data user wishes to produce an estimate for the population of individuals in a household-level file, an approximate expansion factor is the sample weight times the household size of each household.

The ESS2 sample weights were calculated in two stages. In the first stage, weights were separately calculated or adjusted for the three different sampling frames (rural, small town, and large town[20]).

---

[19] The second data entry was done for all modules, but the community questionnaire.
[20] Large town in this case refers to the urban supplement of mid- and large-sized towns. All towns greater than a population of 10,000.

For the rural sample, the wave 1 weights were adjusted to account for relisting, non-response, and attrition of households in the sample frame (the AgSS) between the two waves. In each of the waves, the rural EAs were re-listed which reflects EA-specific population growth patterns. The post-stratification adjustment accounts for this change. For the small town sample, the EAs were not re-listed so the wave 1 weights were adjusted to reflect national population growth between the two waves.

For the mid- and large-town sample, new weights were calculated in accordance with the sampling design and then adjusted to reflect household listing and nonresponse encountered during field work. The sample weights were estimated as the product of the first-stage of sampling (which was PPS selection of EA) and the second stage of sampling (which was SRS sampling of 15 households within each EA). The first stage selection was based on population estimates. The EAs were listed after selection, and the first-stage probability was adjusted to account for the difference between the listing counts and the EA population estimates used for the first stage selection. The product of the adjusted first stage weight and the second stage weight were then adjusted to account for non-response.[21] Finally, these sampling weights were further post-stratified to ensure that the sum of the weights within each region matched the 2014 population estimates provided by CSA. [22,23]

5.2 Description of Public Datasets

The electronic datasets are organized by questionnaire with the following labels on file names in parentheses: household (*hh*), community (*com*), post-planting agriculture (*pp*), post-harvest agriculture (*ph*), and livestock (*ls*). The data within each questionnaire do not contain any constructed variables. For example, the ESS data provide most all variables needed to construct an estimate of total household consumption, but the data set does not contain an estimated value of total consumption. The only compiled data that are included with the ESS files are the geo-spatial variables described below.

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_w2*' for household wave 2 data, and '*com_w2*' for community wave 2 data. Similarly, the suffixes for post planting and post-harvest as well as livestock wave 2 data are '*pp_w2*', '*ph_w2*', and '*ls_w2*' respectively.

For example, the data set that corresponds with the section 1 of the household questionnaire is in the data file '*sect1_hh_w2*'. 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 to the reference period collected in the section. In this case, the name of the corresponding files will be '*sect5a_hh_w2*', '*sect5b_hh_w2*', *sect5c1_hh_w2*' and *sect5c2_hh_w2*'.

---

[21] Given the low level of nonresponse, this is a very minor adjustment.
[22] The population estimates are from the *Population Projection of Ethiopia for All Regions At Wereda Level from 2014 – 2017* produced by CSA and is based upon the National Population and Housing Census of Ethiopia conducted in May 2007.
[23] For Afar and Somalie, only the cumulative population of the zones within the sample frame was calculated and compared (i.e. not the population of the entire region).

Each dataset has identification variables, a rural area/small town/large town indicator variable (*rural*), and sampling weight (*pw2*).

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, as well as names of field and data entry staff, 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 24 data files (Table 5.1).

Table 5.1: Household Data Files

Section	Section Name/ Content	Dataset Filename	Unique identification variables
Cover	Cover	*sect_cover_hh_w2*	*household_id2*
1	Roster	*sect1_hh_w2*	*individual_id2*
2	Education	*sect2_hh_w2*	*individual_id2*
3	Health	*sect3_hh_w2*	*individual_id2*
4	Labor and Time Use	*sect4_hh_w2*	*individual_id2*
5A	Food Consumption Last 7 days	*sect5a_hh_w2*	*household_id2 hh_s5aq00*
5B	Food Consumption Last 7 days	*sect5b_hh_w2*	*household_id2 hh_s5bq00*
5C1	Meal Sharing Last 7 days	*sect5c1_hh_w2*	*household_id2*
5C2	Number of Days Meal shared Last 7 days	*sect5c2_hh_w2*	*household_id2 hh_s5cq00*
5D	Food away from home in Last 7 days	*sect5d_hh_w2*	*household_id2 hh_s5cq0b*
6A	Non-Food Expenditure (one month)	*sect6a_hh_w2*	*household_id2 hh_s6aq00*
6B	Non-Food Expenditure (one year)	*sect6b_hh_w2*	*household_id2 hh_s6bq00*
7	Food Security	*sect7_hh_w2*	*household_id2*
8	Shocks	*sect8_hh_w2*	*household_id2 hh_s8q00*
9	Housing	*sect9_hh_w2*	*household_id2*
10	Household assets	*sect10_hh_w2*	*household_id2 hh_s10q00*
11A	Non-Farm Enterprises Filter	*sect11a_hh_w2*	*household_id2*

11B	Non-Farm Enterprises (at the household-enterprise level)	*sect11b_hh_w2*	*household_id2 hh_s11bq00*
11C	Non-Farm Enterprises (questions 18-19 at the household level)	*sect11c_hh_w2*	*household_id2*
12	Other Income	*sect12_hh_w2*	*household_id2 hh_s12q00*
13	Assistance	*sect13_hh_w2*	*household_id2 hh_s13q00*
14A	Credit (Household Level)	*sect14a_hh_w2*	*household_id2*
14 B	Credit (Loan Level Information)	*sect14b_hh_w2*	*household_id2 hh_s14q00*
14C	Credit (Household Level)	*sect14c_hh_w2*	*household_id2*

5.2.2 Community Data

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

Table 5.2: Community Data Files

Section	Topic	Dataset Filename	Unique identification variables
1A	Cover/ Identification	*sect1a_com_w2*	*ea_id2*
1B	Cover/ Community Overview/ Observation	*sect1b_com_w2*	*ea_id2*
2	Roster of Informants	*sect2_com_w2*	*ea_id2 cs2q01*
3	Community Basic Information/ Demographics	*sect3_com_w2*	*ea_id2*
4	Access to Basic Services/ Infrastructure	*sect4_com_w2*	*ea_id2*
5	Economic Activities/ Employment	*sect5_com_w2*	*ea_id2*
6	Agriculture	*sect6_com_w2*	*ea_id2*
7	Changes/Events	*sect7_com_w2*	*ea_id2 cs7q00 cs7q01_b*
8	Community Needs	*sect8_com_w2*	*ea_id2 cs8q00*
9	Productive Safety nets Program	*sect9_com_w2*	*ea_id2*
10A1	Market Prices: Market 1 Location	*sect10a1_com_w2*	*ea_id2*
10A2	Market Prices in Market 1	*sect10a2_com_w2*	*ea_id2 cs10a2q02*
10B1	Market Prices: Market 2 Location	*sect10b1_com_w2*	*ea_id2*
10B2	Market Prices in Market 2	*sect10b2_com_w2*	*ea_id2 cs10b2q02*

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: Post-planting Agriculture Data Files[24]

Section	Section Name	Dataset Filename	Unique Identification Variables
Cover	Cover	sect_cover_pp_w2	holder_id
1	Household Roster	sect1_pp_w2	holder_id pp_s1q00
2	Parcel Roster	sect2_pp_w2	holder_id parcel_id
3	Field Roster	sect3_pp_w2	holder_id parcel_id field_id
3RCA	Field Roster/Area Measurement	sect3_rca_pp_w2	holder_id parcel_id field_id
3RCB	Field Roster/Area Measurement	sect3_rcb_pp_w2	holder_id parcel_id field_id pp_rcq0d
4	Crop Field Roster	sect4_pp_w2	holder_id parcel_id field_id crop_code
5	Seed acquisition	sect5_pp_w2	holder_id crop_code pp_s5q01*
7	Holder questions	sect7_pp_w2	holder_id
NR	Network Roster	sect_nr_pp_w2	holder_id

* There are exceptions for this data set, see Section 7.2 for more detail

Table 5.4: Post-harvest Agriculture Data Files

Section	Section Name	Dataset Filename	Unique Identification Variables
Cover	Cover	sect_cover_ph_w2	holder_id
1	Household Roster	sect1_ph_w2	holder_id ph_s1q00
9a	Crop Cut by field (for selected fields and crops only)	Sect9a_cc_w2	holder_id parcel_id field_id crop_code
9	Harvest by field	sect9_ph_w2	holder_id parcel_id field_id crop_code
10	Harvest Labor	sect10_ph_w2	holder_id parcel_id field_id crop_code
11	Crop Utilization (cereals, pulses, oil seeds)	sect11_ph_w2	holder_id crop_code*
12	Crop Utilization (permanent tree and root crops)	sect12_ph_w2	holder_id crop_code*
NR	Network Roster	sect_nr_ph_w2	holder_id

* There are exceptions for this data set, see Section 7.2 for more detail

---

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

Table 5.5: Livestock Data Files

Section	Section Name	Dataset Filename	Unique Identification Variables
Cover	Cover	*sect_cover_ls_w2*	*holder_id*
1	Household Roster	*sect1_ls_w2*	*holder_id ls_s1q00*
8A	Livestock Inventory	*sect8a_ls_w2*	*holder_id ls_s8aq00*
8B	Livestock Utilization/Sales	*sect8b_ls_w2*	*holder_id*
8C	Livestock byproduct	*sect11_ls_w2*	*holder_id ls_s8cq00*

5.2.4 Geospatial Data

The ESS data files also include additional geospatial data computed for data users. The geovariables are stored in two data files: field-level data (Pub_ETH_PlotGeovariables_Y2), and household-level data (Pub_ETH_HouseholdGeovariables_Y2). Information on the ESS 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 ESS data to facilitate converting non-standard units into square meters. The ESS data provide both standard unit measures of land (from GPS and/or rope and compass measures) and self-reported values which may be reported in local units. The included conversion factors have not been used to convert self-reported land values into standard unit measures. 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 ESS program. The conversion factors were sourced primarily from a report completed by the CSA[25] and a study on land certification in Ethiopia[26], and supplemented by other internal CSA sources.

---

[25] 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.

[26] 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. Using the ESS Public Data

6.1 File Structure

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

6.2 Merging Datasets

6.2.1 Within wave 2

All households are uniquely identified by the variable *household_id2* 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 datasets, where there is more than one observation per household, 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_id2*. So in order to merge any two individual type files, the variable *individual_id2* 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*.

The community questionnaire is administered at the EA level. A unique EA identifier, *ea_id2,* is in every data file. This variable is the concatenation of the variables *region, zone, wereda, town, subcity, kebele* and *ea*, and is used as the unique key variable in the merging of all community-level data files. In some of the community datasets, additional key variables may be needed.

Due to the sequence in which the ESS visits occur (see Table 4.1), there were 30 households which were administered the post-planting and/or livestock questionnaires but were not administered the household and post-harvest questionnaires. These 30 households could not be located during the subsequent household visit and were dropped from the entire ESS2 sample.

6.2.2 Between wave 1 and wave 2

In order to merge datasets between wave 1 and wave 2 (note: this is only possible for rural and small town households), you will need to use different variables from those mentioned above. To merge household-level data between the two waves, for example, you should use the variable *household_id.* The variable *household_id* is carried over from wave 1 and thus is empty for new urban households as they were not included in wave 1. The variable *ea_id* is also carried over from wave 1 and is empty for all urban EAs new to wave 2. *ea_id* can be used to merge wave 1 community data with wave 2 community data (for EAs included in both waves).

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 N01, N02 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 N01, N02, N03 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 Second Wave of the Survey and Recommendations for the Third Wave of the Survey

Designing and implementing a complex survey such as the ESS 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

During the preparation of the second wave survey the team gathered feedback from data users in Ethiopia and beyond. Most of the feedback was on additional questions. There were also suggestions for modifying some questions. Most of the inputs were incorporated and reflected in the questionnaires. This however, required revising the entire format of the modules as well as the data entry programs.

7.2 Fieldwork

ESS enumerators in rural areas were also responsible for conducting AgSS interviews in their respective EAs. This was an added burden as the enumerators had to measure fields, conduct crop cuts and carryout interviews. This contributed to the delay of the fieldwork in these areas.

Due to the large-scale nature of this survey, there are sometimes mistakes made during the interviewing process. The following points highlight such errors in the data:

-HH Section 1: There were approximately 430 panel individuals where the sex listed in wave 1 was different from the sex listed in wave 2. Due to the gender ambiguity of some names in Ethiopia, the team was only able to correct 77 such cases with 100% certainty. Therefore, the data user will find approximately 350 individuals where the sex between waves does not match.

- PP Section 5, PH Sections 11 and 12: The information for these sections is collected at the level of the holder/crop/seed type. However, in the case of approximately 200 holders, the enumerator mistakenly collected the information at the level of the holder/**parcel**/**field**/crop/seed type for PP section 5 and holder/**parcel/field/**crop for PH sections 11 and 12. Therefore there are approximately 800 cases in PP section 5, 170 cases in PH Section 11, and 300 cases in PH Section 12 that are not uniquely identified by holder/crop(/seed type). The data for these observations can be aggregated by the data users as they see fit.

- Due to confusion in the data collection process, there are two households that have post-harvest data but do not have post-planting data. Thus there are three plots that can be found in PH sections 9 and 10, but not in PP section 4.

7.3 Tracking

Coordinating with different branch offices to trace those households who had moved to other regions required additional resources and field organization. Some households could not be tracked

due to lack of information about their new location. The contact information collected during the first wave was either inaccurate or not enough to trace the locations. Neighbors and community leaders in the original EAs did not have the information.

7.4 Decentralized Data Entry

The second wave employed decentralized data entry arrangement; the questionnaires were keyed in the field to facilitate the concurrent data entry arrangement. While this facilitated error correction in the field it also created a challenge on troubleshooting problems in terms of lower-quality technical infrastructure (eg. computers) in the field offices and also more logistical challenges in terms of fixing errors in the data entry programs in each office rather than centrally. The LSMS team made frequent travels to the branch offices to work with the data entry clerks in the field. Another major challenge in the field-based data entry was frequent power interruption. In some places the power outages lasted for several days.

Appendix 1: Geospatial Data with the ESS

The ESS 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 ESS public use data.

To increase the use of the ESS data, a set of geospatial variables has been provided by using the geo-referenced 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.

Pub_ETH_PlotGeovariables_Y2

The field-level geo file Pub_ETH_PlotGeovariables_Y2 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_Y2

The household-level geo file Pub_ETH_HouseholdGeovariables_Y2 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 ESS household id.

To partially satisfy user interest in geo-referenced location, while preserving the confidentiality of sample household and communities, modified EA-level coordinates are provided as part of the household geovariable table. Modified coordinates are generated by applying a random offset within a specified range to the average EA value (following the MeasureDHS approach). For households that have moved and are more than 10 km from their baseline location, the offset is with respect to the new household location.

More specifically, the coordinate modification strategy relies on random offset of EA center-point coordinates (or average of household GPS locations by EA in ESS) within a specified range determined by the urban and rural classification. For small towns and urban areas, an offset range of 0-2 km is used. In rural areas, where communities are more dispersed and risk of disclosure may be higher, a range of 0-5 km offset 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 A1.1 Field-level geo-referenced data linked to the ESS 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#

Table A1.2 Household-level geo-referenced data linked to the ESS 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/
	WorldPop	Africa 2010 Demography (v ap10_180313)	pop_density	string	2010	0.00833 dd	2010 Population Density Range (people per km2), with national totals adjusted to match UN population division estimates, 2012 revision	http://www.worldpop.org.uk/
	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	2001-2013	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	2001-2013	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	2001-2013	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)	h2013_tot	Continuous	2013	0.1 dd	12-month total rainfall (mm) in Jan-Dec, starting January 2013	ftp://ftp.cpc.ncep.noaa.gov/fews/newalgo_est_dekad/
	NOAA CPC	Rainfall Estimates (RFE)	h2013_wetQ	Continuous	2013	0.1 dd	Total rainfall in wettest quarter (mm) within 12-month periods starting January 2013	ftp://ftp.cpc.ncep.noaa.gov/fews/newalgo_est_dekad/
	NOAA CPC	Rainfall Estimates (RFE)	h2013_wetQstart	Continuous	2013	0.1 dd	Start of wettest quarter in dekads 1-36, where first dekad of January 2013 =1	ftp://ftp.cpc.ncep.noaa.gov/fews/newalgo_est_dekad/
	BU	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	eviarea_avg	Continuous	2001-2013	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	2001-2013	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	2001-2013	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	2001-2013	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)	h2013_eviarea	Continuous	2013	0.004176 dd	Total change in greenness (integral of daily EVI values) within main (Meher) growing season of 2013, averaged by zone	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
		MOD12Q2 Land Cover Dynamics (PHENOLOGY)	h2013_evimax	Continuous	2013	0.004176 dd	EVI value at peak of greenness within main (Meher) growing season of 2013, averaged by zone	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
	BU	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	h2013_grn	Continuous	2013	0.004176 dd	Onset of greenness increase in day of year 1-356, within growing season of 2013, averaged by zone	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
	BU	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	h2013_sen	Continuous	2013	0.004176 dd	Onset of greenness decrease in day of year 1-356, within main (Meher) growing season of 2013, averaged by zone	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005

Appendix 2: 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 website:

http://www.worldbank.org/lsms

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	Development Research Group
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