High Frequency Cell Phone Survey on the Socio-Economic Impacts of Ebola in Sierra Leone

Basic Information Document

ACRONYMS

EVD	Ebola Virus Disease
IPA	Innovations for Poverty Action
LFS	Labor Force Survey
SSL	Statistics Sierra Leone

Table of Contents

INTRODUCTION

The purpose of this document is to provide detailed information on the three rounds of the High Frequency Cell Phone Survey on the Socio-Economic Impacts of Ebola in Sierra Leone conducted between November 2014 and May 2015. The surveys were implemented as part of a wider effort to measure the socioeconomic impacts of the Ebola Virus Disease (EVD) crisis, conducted by the Government of Sierra Leone, with technical assistance from the World Bank and Innovations for Poverty Action (IPA). The survey was designed to provide rapid indicators from a large sample of households across the country at a time when traditional face-to-face surveys were not possible, and followed a sample of households for whom cell phone numbers were recorded during the nationally-representative Labor Force Survey (LFS) conducted in July-August 2014. The main focus of the data collection is to capture the key socio-economic effects of EVD, including impacts on labor market indicators, agricultural production, food security, migration, and utilization of non-Ebola essential health services.

Three reports were produced using the three rounds of the High Frequency Cell Phone Survey on the Socio-Economic Impacts of Ebola in Sierra Leone:

- The socio-economic impacts of Ebola in Sierra Leone : results from a high frequency cell phone survey (round one) – released in January 2015
  http://www.worldbank.org/content/dam/Worldbank/document/Poverty%20documents/Socio-Economic%20Impacts%20of%20Ebola%20in%20Sierra%20Leone,%20Jan%2012%20(final).pdf
- The socio-economic impacts of Ebola in Sierra Leone : results from a high frequency cell phone survey (round two) – released in April 2015
  http://www.worldbank.org/content/dam/Worldbank/document/Poverty%20documents/Socio-Economic%20Impacts%20of%20Ebola%20in%20Sierra%20Leone,%20April%202015%20(final).pdf
- The socio-economic impacts of Ebola in Sierra Leone : results from a high frequency cell phone survey (round three) – released in June 2015
  http://documents.worldbank.org/curated/en/2015/06/24646532/socio-economic-impacts-ebola-sierra-leone-results-high-frequency-cell-phone-survey-round-three

A similar project was conducted Liberia consisting of five rounds of data collection.

SURVEY INSTRUMENTS

As the survey was administered by telephone, the length of the questionnaire was targeted as 20 to 25 minutes. In round 1, the questionnaire focused on employment and labor market conditions, non-agricultural business operations, agricultural activity, food security, health responses (covering only fever and pregnancy), remittances, travel, trust and knowledge about Ebola. In round 2, questions were added on social assistance and education on the radio, and there were small changes to the existing questions based on the results from round 1. Questions on earnings were revised to match the LFS questions more closely, in particular to account for earnings that were expressed in time unit other than months, and questions on the incidence and treatment of child diarrhea were adding using identical wording to the DHS. The most substantial changes were to the migration section as the round 1 analysis found inconsistencies in the migration reporting. Details of these changes can be found in the round 2 report. In round 3, the agriculture, social assistance, and education sections were expanded while the trust section was dropped due to limited variation between rounds 1 and 2.

The only questions on EVD specifically were in round 1 and focused on whether the respondent had heard of Ebola and what were their main sources of information were. This section was placed at the end of the questionnaire in order to elicit unbiased responses in other sections, since people may be distrustful of the government especially regarding Ebola, at a time of such emergency. Questions related directly to incidence of EVD within the household were excluded for two reasons. First EVD is a relatively rare event and the sample was unlikely to yield sufficient observations for meaningful analysis, and secondly, the respondents will be called repeatedly as part of the high frequency survey therefore it was necessary to avoid sensitive questions that may increase attrition in later rounds. The included questions were worded in such a way as to facilitate differences-in-differences comparisons. The vast majority of questions were identical in their wording to those asked during

the LFS or other nationally representative surveys for which detailed data were available including the Demographic and Health Survey (DHS), the National Public Services survey (NPS) and the Agricultural Households Tracking Survey (AHTS). (See appendix for links to data sources.) In a few cases, the time period over which the questions were asked was shortened to make it relevant to the last few months during which the outbreak has been growing. For example, the NPS asked about remittances in the last year whereas in November 2014, respondents were asked about remittances received in the last month.
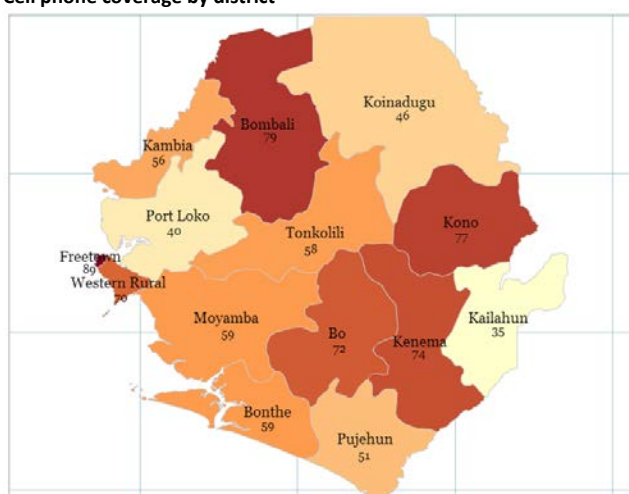
SAMPLE DESIGN

The sampling frame for the cell phone survey was the Sierra Leone Labor Force Survey (LFS). The LFS is a nationally representative stratified cluster sample survey conducted in July and August 2014, and includes the oversampling of urban areas. As part of the LFS, a total of 4199 households in 280 enumeration areas (EAs) were interviewed. Interviewers collected the phone number, if available, for the head of household, and 2,764 households interviewed in the LFS included phone numbers. All available numbers from the LFS were included in the cell

phone survey. See table in the response and attrition section for percentages by district. The phone numbers were reported for 43 percent of rural households and 82 percent of urban households. Those households reporting numbers are unevenly distributed across the sample though there is at least partial coverage in all districts, ranging from 93 percent in Freetown (Western urban) to 30 percent in Kailahun district.

Cell phone coverage by district

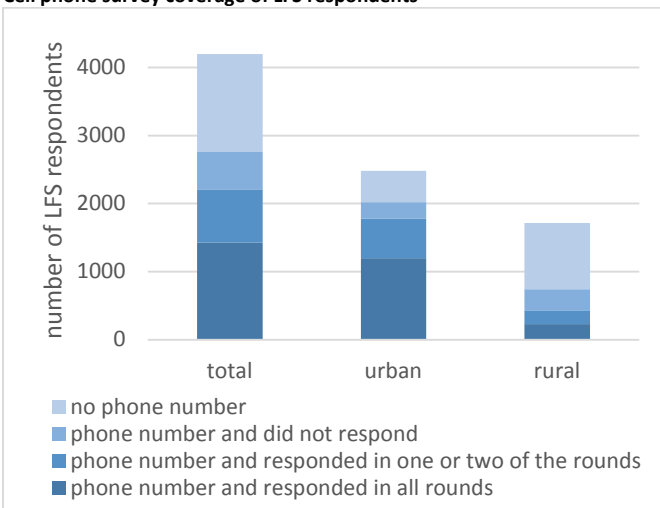*Source: LFS 2014*

IMPLEMENTATION

The survey was implemented by enumerators recruited by SSL and IPA from SSL's Freetown offices in three rounds. The first round was from November 12 – November 25, 2014; the second round was from January 22 – February 4, 2015; and the final round was from May 1 – May 12, 2015. The questionnaire was administered using computer assisted telephone interviewing from a CSPro application run on desktop computers. If respondents did not answer the phone after the initial attempts, a text message was sent to explain the purpose of the call. Respondents also received an incentive in the form of 50 phone units (valued up to 50 US cents) in cell phone credit for completed calls. A maximum of nine attempts were made to contact target respondents over the course of 14 days, with no more than three attempts being made in a single day. Interviewers called requested to speak to household heads. If a household head was not available after three tries, a spouse or another adult was interviewed. Of the households reached, 96 percent were household heads. If the respondent was not an original household member, the call was ended and an incorrect number was recorded.

Cell phone survey coverage of LFS respondents

- no phone number
- phone number and did not respond
- phone number and responded in one or two of the rounds
- phone number and responded in all rounds

Source: LFS 2014 and cell phone surveys

RESPONSE AND ATTRITION RATES

Due to differing characteristics between responding and non-responding households, the results should be considered "descriptive" rather than representative of the Sierra Leonean population. Overall the response rate was higher than expected given the nature of the survey and the difficult conditions under which it was conducted. In Sierra Leone, of the 4,199 households interviewed in the LFS, 65.8 percent (2,764 households) recorded a cell phone number for the household head, and, of those, 80.0 percent responded to at least one round of the cell phone survey. The unweighted sample was 59.1 percent urban (2,483 households) and 40.9 percent rural (1,716 households). Of urban households, 81.4 percent (2,021 households) listed a cell phone number for the household head, and, of those, 88.1 percent (1,780 households) responded in at least one of the three rounds of the cell phone survey. Of rural households, 43.1 percent (740 households) listed a cell phone number for the household head, and, of those, 58.1 percent (430 households) responded in at least one of the three rounds.

	Labor Force Survey		% of LFS found in round 1	% of LFS found in round 2	% LFS found in round 3
	Freq.	Percent			
Kailahun	210	5.0	17.6	19.1	19.0
Kenema	420	10.0	51.0	49.8	48.1
Kono	420	10.0	58.1	56.0	51.9
Bombali	330	7.9	47.6	47.3	43.3
Kambia	181	4.3	32.6	37.6	33.7
Koinadugu	180	4.3	31.1	29.4	30.6
Port Loko	179	4.3	27.4	28.5	24.6
Tonkolili	180	4.3	25.6	25.6	22.2
Bo	421	10.0	43.9	44.7	39.0
Bonthe	269	6.4	42.0	37.9	37.2
Moyamba	180	4.3	34.4	40.0	32.8
Pujehun	180	4.3	24.4	28.3	25.0
Western Rural	288	6.9	51.7	37.9	40.6
Western Urban	761	18.1	63.2	64.3	56.1
Total	4199	100.0	45.2	44.4	40.8

The largest component of non-response was phones that rang but were not answered. Table A2 shows a breakdown of the call outcomes including unanswered calls, phone being switched off, rescheduled but never completed, refusal, bad network/call drops off, incorrect phone number, and number disconnected. Comparing the characteristics of respondents to the overall sample frame, 96 percent were household heads and all were original household members. Overall, the characteristics of the respondents were similar to those in the original sample. Comparing the average age of respondents to that of the original sample of household heads, the most logical available comparison group, the average age of a respondent was 44.9, compared to 44.7 in the original LFS sample. Thirty percent of the cell phone survey respondents were female compared with 29 percent female household heads in the original sample. The table above shows the distribution of location for cell phone survey respondents and the original LFS sample.

Call Result (round 1)	
Survey Completed	1896
Phone switched off	745
Incomplete	37
Wrong number	36
Mobile company no longer active	22
Call unanswered	13
Rescheduled but never completed	6
Refusal	5
Bad network/call drops off	4
Total	2764

The base weights for the cell phone survey were the probability weights from the LFS. Sampling weights for the LFS households were calculated by,

$$\text{Household weight} = 1/(P_{EA,strata} * P_{HH,EA})$$

where $P_{EA,strata}$ is probability of EA being selected within strata, and,
                $P_{HH,EA}$ is probability of household being selected within the EA.

To account for higher likelihood of more populated EA's being selected, $P_{EA,strata}$ is calculated as,

$$P_{EA,strata} = (n_{EA,strata} * N_{HH,EA})/N_{HH,strara}$$

where $n_{EA,strata}$ is number of EA's selected within the strata,
                $N_{HH,EA}$ is the total number of households within that EA, and,
                $N_{HH,strara}$ is total number of households across all EAs in that strata.

Household selection probability was calculated using,

$$P_{HH,EA} = n_{HH,EA}/N_{HH,EA}$$

To compensate as much as possible for non-response and low coverage rates, an attrition adjustment was applied. A propensity score adjustment, which uses the available characteristics of the household head from the LFS (age, gender, location, and employment sector) to calculate an aggregate probability of response, was calculated. These calculations need to be done separately for each combination of data sets, meaning the attrition calculations between the LFS and round 1 would be different than those between the LFS and round 2, which would also be different than those between the LFS and households that answered in both rounds 1 and 2. As an example the results of this analysis between the LFS and round 1 of the cell phone survey are presented in Table A1 in the appendix. The inverse of this probability is then applied to the probability weights, therefore increasing the weight for underrepresented groups. As a final step, a post-stratification correction was applied, adjusting the weights to match known population totals at the district and urban/rural levels.

DATASET
The dataset is compiled into three datafiles – slhfcps_r1, slhfcps_r2, and slhfcps_r3 – corresponding to each of the three survey rounds.

Linking to LFS
The identifying variable for each dataset is hhid_lfs, which can be used to merge the datsets together, and to merge into the original LFS. This variable is called "SERIAL_NUMBER" in the LFS data and would need to be renamed to merge. Variable names are text descriptions of the variable content with the question number included as part of the label. Note that question numbers differ somewhat between rounds of the survey.

TECHNICAL NOTES ON USE OF DATA IN THE REPORTS ON SOCIO-ECONOMIC IMPACTS OF EBOLA IN SIERRA LEONE
For the three reports on the Socio-Economic Impacts of Ebola in Sierra Leone, there are a number of specific decisions on how to use the data for analysis. Six key areas are outlines below related to defining eligible households and measurement of employment, wealth index, monthly wage earnings, household revenues, and pregnancy.

Eligible households
For most of the three reports, the LFS and the three rounds of the cell phone survey are used as repeated cross sections and not as a panel. To be as consistent with the round 1 report as possible, the repeated cross sections

were created as follows. The round 1, round 2, and round 3 households are the full sample of households for which cell phone survey data was collected in November 2014, January – February 2015, and May 2015, respectively. The LFS cross section includes all the households that were surveyed in any of the three cell phone rounds, even if not in all rounds. This means that for the employment section, the sample is slightly different in each of the three reports. There are some household heads in the cell phone survey for whom there is no employment data in LFS dataset as they were considered not part of the labor force during the LFS. In the reports, for the employment section, the round 1, round 2, and round 3 samples are restricted to those household heads for whom employment data was collected in the LFS. Most of the results in the employment section are based on using repeated cross sections, except for two sub-sections where the sample is restricted to being the panel sample of households across all three rounds (i.e. the sample of households that are in LFS, round 1, round 2, and round 3). The two sub-sections where the panel sample is used are those on employment transitions and earnings.

Employment

Given the high frequency nature of the four surveys used and the nature of the EVD crisis, a slightly modified definition of employment was used in the analysis. Households heads were categorized as in the labor force in any given round of the surveys if they were working, looking for work or expected to return to work. For the round 1 report, if a household head was in the labor force in either the LFS or round 1 of the cell phone survey, he was categorized as in the labor force in both rounds. This was done because both rounds of the survey were conducted within three months or less of the previous round and it is unlikely that someone who was working in the LFS suddenly decided to exit the labor force rather than become unemployed due to EVD. Such high frequency labor force surveys are contrary to most other employment surveys and thus necessitate different definitions of labor force participation. In the reports, to be consistent with the round 1 and round 2 report's approach and because of the high frequency nature of these employment surveys, a household head was categorized as in the labor force in all four rounds of surveys if they were in the labor force in any one round. As a result, none of the changes observed in employment rates are due to changes in the composition of the labor force.

NOTES ON VARIABLE CONSTRUCTION FOR REPORTS

Wealth Index

As consumption data is not available for either the LFS or the cell phone survey, a wealth index using principal components analysis is used to proxy differences in well-being. The index includes information on livestock assets (goats, pigs, chickens), educational attainment (literacy and completion of primary school), housing structure (electricity, material of walls, and toilet facilities), and dwelling characteristics (water source and lighting source). This index is then divided at the median into an indicator variable for wealthier households.

Monthly Wage Earnings

Most wage workers report earnings in monthly terms, and therefore results associated with wage earnings are reported this way. For respondents who report wage income in other time units, the analysis translates their wages into monthly terms under the assumption they work at a standard capacity, i.e., 8 hours a day, 22 days or 4.3 weeks a month, and 12 months a year. The earnings data was not collected in round 1 in a way that allowed direct comparison to the LFS, which is the reason only LFS and round 2 are compared. Since earnings data tend to be noisy and a few large outliers can have a big impact on average wages, the figures reported in the reports exclude earnings for the highest 5 percent. As a robustness check, median earnings were also analyzed and the same trends held.

Household enterprise revenues

Business revenues are noisy so the main results in the reports have the top percentile of revenues trimmed. As LFS has the highest revenues, a large fraction of the outliers are from LFS. For this reason the reports present results based on an alternative approach, i.e. the top 1 percent of revenues in each round is trimmed. As a robustness check, revenues were also calculated by excluding the top 1 and 5 percent. Same trends emerged from the analysis of these alternative measures. Also, revenues are collected at the household level in the cell phone surveys

compared to the enterprise survey in the LFS. This was done to reduce the complexity and length of the time needed for the cell phone survey.

Pregnancy

In the DHS each woman in the household was interviewed individually whereas in the cell phone surveys the respondent was the household head. It is possible that the cell phone survey underreports utilization if household heads are not always aware of clinic visits made by household members. It is also possible that the cell phone survey over reports pregnancies if the household head misremembers dates of birth and include pregnancies and child births that took place more than 2 months prior to the interview. A final difference is that while the DHS reports on current pregnancies and births in the last 2 months the cell phone survey captures anyone who was pregnant in the last two months i.e., the cell phone survey also captures visits from those who had miscarriages or abortions in the last 2 months while these are not include in the DHS utilization figures.

OBTAINING DATA

The micro data for the high frequency cell phone surveys can be downloaded from the World Bank's microdata catalogue. The original LFS data was collected as a collaboration between the World Bank's Social Protection and Labor Global Practice and Statistics Sierra Leone. Access to the LFS data can be obtained through the World Bank microdata portal or by written request to the Secretary General at Statistics Sierra.

Weight Calculations

Table A1: Propensity Score Regression Results (round 1)

	Coefficient	Std. Err.	z	P>|z|
Age	0.0234	0.0127	1.8400	0.0650
Age Squared	-0.0002	0.0001	-1.7400	0.0820
Gender	-0.0395	0.0797	-0.5000	0.6200
Wage Sector	0.2169	0.1157	1.8800	0.0610
Agriculture Sector	-0.4008	0.1068	-3.7500	0.0000
Non-Agriculture Self Employed Sector	0.2907	0.1025	2.8400	0.0050
Unpaid Workers	0.0310	0.2149	0.1400	0.8850
Household Head Can Read and Write	0.2598	0.1255	2.0700	0.0380
Household Head Has More Than Primary Education	0.3905	0.1308	2.9900	0.0030
Household Owns Livestock	0.1312	0.1988	0.6600	0.5090
Household Owns Goats	0.2068	0.1211	1.7100	0.0880
Household Owns Pigs	-0.5166	0.4247	-1.2200	0.2240
Household Owns Chicken	-0.0824	0.1881	-0.4400	0.6610
Stratum: Kailahun, Urban	1.3802	0.4089	3.3800	0.0010
Stratum: Kenema, Rural	0.9552	0.4124	2.3200	0.0210
Stratum: Kenema, Urban	2.7604	0.3564	7.7500	0.0000
Stratum: Kono, Rural	1.2543	0.3971	3.1600	0.0020
Stratum: Kono, Urban	3.0357	0.3591	8.4500	0.0000
Stratum: Bombali, Rural	1.2652	0.4028	3.1400	0.0020
Stratum: Bombali, Urban	2.6045	0.3669	7.1000	0.0000
Stratum: Kambia, Rural	1.3583	0.4004	3.3900	0.0010
Stratum: Kambia, Urban	1.9302	0.4307	4.4800	0.0000
Stratum: Koinadugu, Rural	0.7195	0.4236	1.7000	0.0890
Stratum: Koinadugu, Urban	2.6063	0.4349	5.9900	0.0000
Stratum: Port Loko, Rural	0.9414	0.4129	2.2800	0.0230
Stratum: Port Loko, Urban	1.6943	0.4305	3.9400	0.0000
Stratum: Tonkolili, Rural	1.0921	0.4095	2.6700	0.0080
Stratum: Tonkolili, Urban	1.8056	0.4305	4.1900	0.0000
Stratum: Bo, Rural	0.9186	0.4109	2.2400	0.0250
Stratum: Bo, Urban	2.2318	0.3555	6.2800	0.0000
Stratum: Bonthe, Rural	1.2118	0.4014	3.0200	0.0030
Stratum: Bonthe, Urban	2.2747	0.3777	6.0200	0.0000
Stratum: Moyamba, Rural	1.3079	0.3984	3.2800	0.0010
Stratum: Moyamba, Urban	2.5419	0.4300	5.9100	0.0000
Stratum: Pujehun, Rural	0.5161	0.4322	1.1900	0.2320
Stratum: Pujehun, Urban	1.9748	0.4314	4.5800	0.0000
Stratum: WA Rural, Rural	1.5570	0.3766	4.1300	0.0000
Stratum: WA Rural, Urban	1.5193	0.3783	4.0200	0.0000
Stratum: WA Urban	2.5619	0.3474	7.3700	0.0000
Constant	-2.9696	0.4575	-6.4900	0.0000
Number of obs		4,199		
Log likelihood		-2428.3048		

The three High Frequency Cell Phone Survey questionnaires to assess the Socio-Economic Impacts of Ebola in Sierra Leone were designed to be compared to previous data. These data sources are described here. Among them, the LFS was the basis for the sample of the High Frequency Cell Phone Surveys.

Sierra Leone Labor Force Survey (2014)

The Sierra Leone Labor Force Survey is a nationally representative household survey and was conducted by Statistics Sierra Leone and the World Bank in July and August of 2014. The total sample size was 4,199 households, 59 percent in urban areas and 41 percent in rural areas. As urban areas were oversampled, probability weights are used to obtain unbiased national estimates. Of the total sample, 66 percent reported cell phone numbers. The main topics covered by the LFS were household listing and demographic information, education, training, and migration, unemployment and inactivity, current main and secondary economic activities, usual economic activity, industrial relations and occupational injuries, time use, family/household non-farm enterprises, and farming activities. The first cases of Ebola were detected in Sierra Leone in May.[1] Therefore while the LFS therefore does not represent a clean pre-Ebola baseline, the outbreak was much more geographically contained. At the end of June, there were only four cases outside of Kailahun and Kenema, the first two districts to be infected. The first restrictions placed on economic activity were imposed at the start of August when a cordon was imposed on Kailahun and Kenema, bars were asked to close throughout the country, and schools were closed. Thus, 50 percent of the LFS was collected in a month when the economic impacts of Ebola were likely to be minimal and 50 percent was collected when the impacts were likely to be restricted to specific regions and sectors. To the extent that Ebola may have been already negatively impacting economic activity in July and August of 2014, the estimate of the economic impact of Ebola may be underestimated.

To analyze the effects of the EVD outbreak on employment, the reports using the high frequency cell phone surveys replieon measures of employment that ensure comparability across surveys while capturing short term changes. The reports used a measure of labor force participation that includes household heads in the labor force if they were in the labor force either at the time of the LFS or in the November round of the phone survey. It also no longer considers someone employed in November if they are temporarily absent from work for three months or less and the reason cited is Ebola. Since it is unlikely that household heads dropped out of the labor force over such a short period of time for reasons unrelated to the outbreak, this allows short term changes in work induced by the outbreak to be captured. When defined this way, the labor force participation rate among working-age heads is 98 percent.

Sierra Leone Demographic and Health Survey (2013)

The 2013 Sierra Leone Demographic and Health Survey (DHS) is a nationally representative survey focused on topics related to family and child health issues, including fertility, family planning, maternal and childhood mortality, maternal and child health, nutrition, and HIV/AIDS. The DHS was implemented by the Ministry of Health and Sanitation and SSL. Three types of questionnaires were administered, a household questionnaire, a women's questionnaire for all women aged 15-49, and a men's questionnaire for men aged 15-49 in every second household. The survey was administered from June to October 2013, and includes data on 12,629 households. Further information on the DHS methodology is available at http://www.dhsprogram.com.

National Public Services Survey (2008)

The National Public Services Survey (NPS) was conducted in 2008 by the Institutional Reform and Capacity Building Project, a joint initiative of the Government of Sierra Leone and the World Bank. It involved two questionnaires: a household questionnaire that was administered to ten households in each EA; and a community questionnaire that, in rural areas, was completed during an informal village meeting, and, in urban areas, was completed based on the survey enumerator's own observations. It is a nationally representative survey that focuses on the state of public services, political attitudes, and community organization in Sierra Leone. It covers a number of topics including the quality of, cost of, and satisfaction with public services;

---

[1] http://wwwnc.cdc.gov/travel/notices/warning/ebola-sierra-leone

participation in and the accountability of various levels of government; and social capital and political attitudes. The survey contains data on 6,343 households in 634 enumeration areas (EA). These are the same EAs and, to a large extent, the same households covered when the survey was previously conducted in 2005 and 2007. In order to provide results representative at a district level, the sample over samples EAs in small districts. The results here are reweighted to ensure that the results are representative at the national level. Further information is available at
http://thedata.harvard.edu/dvn/dv/RepublicSierraLeone/faces/study/StudyPage.xhtml?globalId=hdl:1902.1/16786.

Agricultural Household Tracking Survey (2010)
The Agricultural Household Tracking Survey (AHTS) was commissioned by the Office of the President of Sierra Leone, and implemented collaboratively by the Ministry of Agriculture, Forestry and Food Security, Statistics Sierra Leone, and the Innovations for Poverty Action. It is a nationally representative survey of farming households in Sierra Leone. The questionnaire was designed to capture information on the agricultural activities of smallholder farmers, and covered topics such as: farmers' decisions; yields and production levels; access to services and technology; and food security. The survey was conducted between March and May of 2010 and contains data on 8,803 households in 917 EAs. The sampling of EAs was stratified by district, and the questionnaire was administered to ten households in each EA. The outcomes have been reweighted to make the results representative of agricultural households in the country as a whole. There are insufficient nonagricultural households included in the sample to allow for reweighting to give nationally representative outcomes.
Further information is available at:

http://thedata.harvard.edu/dvn/dv/ahts/faces/study/StudyPage.xhtml?studyId=85626&tab=catalog.

Questionnaires