May 30, 2012

Concept Note

Household Income and Expenditure Survey: Liberia

I.   **Overview**

This Concept note describes the work program for Liberia Institute of Statistics and Geo-Information Services (LISGIS) on the Household Income and Expenditure Survey (HIES). This is a multi-year program that encompasses, among other features: the design and implementation of a household survey focusing on household income and expenditure which feeds into CPI construction, poverty analysis and update of household expenditures section within National Accounts. The project is also expected to provide a detailed agricultural productivity analysis and serve as baseline information for the "Agenda for Transformation" set by the Government of Liberia. Other components of this project include capacity building and cross-country knowledge sharing, alongside efforts to improve survey methodologies in Liberia.

II.   **Background Information**

The HIES and National Accounts

In Liberia, like most countries in Africa, the production approach is mostly used to prepare the GDP estimates. In the production approach of the national accounts, output and value added for all activities in the economy are estimated. After adjustments for taxes (import duties and VAT) and FISIM (Financial Intermediation Services Indirectly Measured), the total GDP of the nation is computed.

Several major sources of information are available for estimating different components of GDP. These can be classified by grouping activities by institutional sector, that is, the financial and non-financial corporations, the government, non-profit institutions serving households (NPISH) and the household sector. For the first three, books of accounts are available and reliable statistical information can be obtained from these sources. NPISH are also required to maintain proper accounts but in Liberia, enforcement of this rule is weak and often there is also not a central repository where the information is kept. The household

sector is very important in the economy of Liberia, but the weakest in terms of statistical data. The major information source for information for this component is typically a Household Income and Expenditure Survey (HIES). In the case of Liberia, this survey has never been conducted on a nationally representative scale, taking into account seasonality in income/expenditure patterns.

The HIES and Consumer Price Index

The National Accounts of a nation are compiled in constant prices for ease of comparison over time. However, much of the information going into the estimates is in current prices. Therefore, it is necessary to develop methods to restate these current-price values to constant prices. This process is called deflation and the indicators used for this purpose are the deflators. In many cases, the Consumer Price Index (CPI) is used as deflator by default, rather than choice.

The CPI measures the average change in prices of the consumption. Price collection is done on a regular basis for all products in the consumption basket. This basket comprises a representative selection of items consumed by the general population in the country and it is based on the pattern of consumption expenditures obtained from a household survey. In most countries, the HIES is used as the source of information for household consumption and expenditures which subsequently leads to the creation of the weights for the CPI. The weights provide information on how households value each item listed in the consumption basket.

As mentioned above, the HIES has never been conducted on a national scale and the current CPI estimates emanate from a limited 124 household survey restricted to Monrovia conducted in 1964, which is outdated.

Poverty Profile

As part of its efforts to track poverty and monitor household living standards, LISGIS regularly conducts a number of large-scale household surveys. These surveys include the Census of Population and Housing every ten years, the Demographic and Health Survey (DHS) every five years, Agricultural Annual Survey (AAS) every year and the Core Welfare

Indicator Questionnaire (CWIQ) every two years, and the Labor Force Survey (LFS) every five years.

Although the current set of surveys encompasses a wide range of topics relevant for monitoring welfare, three important conclusions can be drawn on the state of statistics obtained from these surveys

- While the surveys are nationally representative, they have not been conducted on a 12 month basis to account for seasonality. For example, household consumption patterns might differ right after the harvest period compared to the rest of the year. If the previous surveys are conducted over a short period of time, the seasonality effects cannot be eliminated.

- Some of these surveys are topic specific (for example, the AAS focuses on agricultural households, the DHS focuses on health, LFS on labor activities) and others do not have detailed enough modules to allow for poverty analysis, particularly factors that affect poverty numbers in different parts of Liberia (Census, CWIQ).

- Aside from the CWIQ, there is no information on consumption and expenditures, and the CWIQ numbers are plagued with many data problems. In fact, CWIQ has been identified as not sufficient for utilization since its sample size is very small and the data provides estimates at the regional level and not county level.

- Statistics in Liberia suffer from a number of problems common to other countries in the region. These include duplications and contradictory information, insufficient coverage, poor documentation and dissemination, and uneven quality both across sources and over time. This results in the existing data having both low credibility and limited use.

The Government of Liberia (GoL) recognizes that improving income, expenditure and poverty statistics is the backbone of sound sectoral policies.For this purpose, Liberia Institute of Statistics and Geo-Information Services (LISGIS), with technical assistance from the World Bankis working towards the design and implementation of a new multi-purpose

Household Income and Expenditure Survey (HIES) that addresses some of the key concerns and data gaps.

The HIES collects detailed information on the expenditure, income and household characteristics of a sample of residents in a country at a particular time. It provides data that critical in measuring the economic well-being of the population and provides information on the command over economic resources of individuals and households. This enables an environment to undertake serious analysis of assessment of levels of economic inequality, and the effectiveness of the social support system.The HIES therefore offers the unique opportunity to fill some of the existing gaps in terms of providing weights for a National Consumer Price Index (NCPI), providing crucial household expenditure information for the National Accounts and more broadly a framework to ensure that income/expenditure and poverty statistics in Liberia are more policy relevant and analytically useful.

III. Previous Efforts on CPI Construction and Update of Household Consumption and Expenditures within National Accounts in Liberia

The Household Income & Expenditure Survey (HIES) was first planned in 1963 at the national level with 752 sample households in order to determine how people expend their disposable incomes on goods and services for household use. Although the activity was abandoned due to financial reasons, alimited survey was carried out in Monrovia and its environs in November – December 1964. The sample consisted of124 of the sampled of 752 households, comprising of heads who were salariedemployees and/or wage earners with acombined income of less than US$250.00 in 1964. Based on the results of this limited survey, the first Base-Period for the Monrovia Consumer Price Index (MCPI) was derived, and is presently being used in calculating the inflation rate and indices for Liberia.

A few adjustments have been made since 1964. This includes a modification to the basket of goods and services in 1998 along with a change in the price of the reference base periods from December 1964 to May 1998. At the time, the base period prices were calculated using price data collected from a special survey conducted during March – April, 1998. However, the base period weights data from 1964 HIES remained unchanged.

More serious efforts were exerted to adjust the MCPI in order to provide an opportunity for government to source funding for the construction of a new national consumer basket. In January 2005, a consumer price specialist from the Economic Community of West Africa States (ECOWAS) visited Liberia to assist with the development of a harmonized consumer price index (HCPI) for the country. The index was intended to be based on the Classification of Individual Consumption by Purpose for Household Budget Survey (COICOP-HBS) and a new list of 515 items had been selected for the ECOWAS harmonized market basket.

In May 2006, the IMF contracted a consumer price consultant to work with the International Comparison Prices (ICP) team in re-adjusting market basket weights at the COICOP major group level using market basket data of four neighboring ECOWAS countries: Sierra Leone, Ivory Coast, Ghana and Guinea because of the similarity of consumption patterns of the people of Liberia.These new weights at the COICOP major group were then distributed across all of the 234 items selected for the new Harmonized Consumer Price Index (HCPI) market basket. This was done on the basis of the work done by ECOWAS mission prior to the fund's CPI expert.The harmonized consumer price index that evolved from this activity is an amalgamation of the COICOP twelve (12) functions.

Numerous advantages were foreseentothe usage of the harmonized consumer price index. Firstly, an enlarged market basket of 234 commodities was better than a market basket of 79 commodities in the MCPI. This enlarged market basket provided a more realistic picture of the expenditure patterns of households. For example, if we were to conduct a household income and expenditure survey today, itis anticipated that the results of this survey will show that households could spend on average, 45% - 50% of their disposable incomes on food and non- alcoholic beverages as compared with 35.5%currently recorded in the MCPI. Secondly, with the importance of communication in today's world, it is believed, based on the HCPI that households will expend on average 1.5% of their disposable incomes on communication, given the proliferations of mobile phones. Communication was not a part of the market basket of the MCPI. Finally, the HCPI stands to serve as a better instrument for wage and salary negotiations, as its measure of price movements will be more realistic than that of the MPCI because of its scope.

Despite the modifications to the weights derived from the HIES conducted in 1964, some major concerns still remain:

- Liberia has undergone many socio-economic changes since 1964, particularly in the past decade which is not reflected in the current system since the weights associated withthe basket of goods have been modified based on information from neighboring countries but not an internal data collection effort.

- The limited sample size of the survey conducted in 1964, the one month duration of the survey and its focus on Monrovia are problematic from the analytical perspective. The sample size of the survey from 1964 is too small to reflect on the ground reality of today. Additionally, the duration does not account for seasonal adjustment in consumption and expenditure patterns. Finally, when the MCPI was constructed, the operating assumption was that Monrovia was the major trading center in the country and therefore, changes in prices in Monrovia would have serious effects in other parts of the country. This clearly is not the case anymore and necessitates a nationally representative sample.

It is evident from the discussion above that a new Household Income and Expenditure survey is indispensable to not only update the existing CPI weights but also to update household expenditure within National Accounts and create a poverty profile for Liberia.

IV.   **Agricultural Statistics in Liberia and integration into the HIES**

Agricultural activities play an important role in the Liberian economy in terms of its contribution to household income generation, employment and food security. In order to provide insight into key components of the agricultural sector including production of food crops and livestock, the Liberia Institute of Statistics and Geo-Information Services (LISGIS), in collaboration with the Ministry of Agriculture implements an Annual Agricultural Survey. The survey is nationally representative and provides production estimates for crops and livestock at the county level.

Some of the problems with the implementation of this survey are listed below

- The current Annual Agricultural Survey has a non-rigorous and poorly documented multi-stage sampling design with Enumeration Areas as the Primary Sampling Unit, Agricultural Holders as Secondary Sampling Units, holders of rice farms as the tertiary sampling unit and finally experimental plots as the ultimate stage sample unit.

- The current agricultural survey only focuses on harvesting period of the rainy season with no targeting during the dry season. Many rural households, particularly in Lofa, Nimba and Bong counties also engage in agricultural activities during the dry season.

- Thirdly, the Agricultural Survey does not allow for a monitoring of welfare and does not link living standards of households (especially rural households) to livelihood strategies and measures of income diversification.The rural development literature has consistently shown that income diversification at the household and communitylevel is practiced across the globe, with agriculture still constituting a crucial sector of employment in rural economies. Higher incomes and lower risk exposure can be achieved by enhancing the linkages among the different income sources of the rural poor. However, adequate data to study these issues is lacking in Liberia.

In view of the importance that agriculture plays in the national economy and in the livelihoods of Liberian households, strengthening the availability, quality, and policy relevance of information on the agricultural sector is of utmost importance. Over the past years, WFP, FAO and other donors have made substantial investments in support of agricultural and rural development in Liberia. However, key donors and government agencies alike have often lacked the information base to guide their investment decisions and evaluate their impact. The HIES provides the unique opportunity to obtain reliable national level statistics on agriculture, allowing for, among other things, the estimation of land areas, both owned and cultivated, self-reported production figures for main crops and livestock, and detailed cost of production for crops at the household level. The data will not only provide numbers on agricultural productivity at the national level, butalso allow for disaggregation of the data by gender and counties. Talks are on-going on incorporating crop-cutting activities

into this framework.However, theyearlong nature of this survey poses a major challenge to the incorporation of crop cutting activities, which are currently part-taken right after the rainy season in the Annual Agricultural Survey.

V.   Linking the HIES to the Agenda for Transformation

The Agenda for Transformation is a new medium term economic growth and development strategy (2012 – 2017), that serves as a guide to development activities in Liberia. This medium term plan is linked to the long term national vision, Liberia RISING 2030, whose overarching goal is for Liberia to achieve middle income status by 2030. The Agenda for Transformation focuses on key investments in Infrastructure (roads, energy), Youth Skills Development & Employment, health improvement, education and manpower development, social safety net provision, security, private and public sector development.

The multi-topic nature of the Household Income and Expenditure Survey will be ideal for serving as a baseline for the focus sectors within the Agenda for Transformation. A follow up survey in 2017 using similar survey techniques (or ideally the same sample as a panel) would provide endline results for the Agenda for Transformation. A comparison of the households using the baseline and endline survey will provide key information on interventions that worked and did not work and the reasons for the same, particularly if an impact evaluation methodology is adopted. The design of the HIES will not only allow for household level analysis but also aggregate information at the county level and disaggregate results by gender.

VI.   Objectives of the Household Income and Expenditure Survey

The objectives for conducting the HIES in Liberia are multifold:

- To obtain a new set of weights for the basket of goods and services that allow for upgrading the Monrovia Consumer Price Index (MCPI) to the National Consumer Price Index (NCPI).
- To get information on household expenditure patterns in order to update the National Accounts.
- To understand the poverty dynamics across the country and factors influencing them.

- To undertake in-depth analysis of agricultural households, focusing on the links between living standards of households to livelihood strategies and measures of income diversification. The goal is to ensure that agricultural statistics in Liberia are more policy relevant and analytically useful.

- To provide a database that allows for baseline analysis of national level government policies embedded within the Agenda for Transformation.

- Emphasize capacity building and development of sustainable systems for the production of accurate and timely information on households in Liberia. A number of training activities are envisaged under this project such as a STATA workshop, household survey design and management, and policy evaluation workshop.

VII. Activities

To realize the objectives of the project, a number of interrelated activities need to be under taken in a timely manner. These activities have been grouped into five different categories, namely: **a) Preparation, b) Fieldwork, c) Analysis and Dissemination, d) Implementation Agreement and Management, and e) Timetable.**

a) Preparation:

The preparation phase involves the execution of the following activities

1. Sampling and Household Listing

Most household surveys, including the HIES have complex sample designs because of the multi-stage, stratified and clustered features. Adding to the complexity is also the multi-topic nature of the HIES. Sampling is often as much a financial and political issue as it is a technical one, and issues like total sample size and stratification are often decided along with the idea of conducting the survey. The time taken to reach final decisions on these issues mostly depends on the difficulty of establishing a consensus.

Many factors need to be accounted for, in order to finalize a good sample design for household surveys. The sample is usually placed in stages to identify the locations where interviews are to take place and to select households efficiently. Since the HIES is anticipated to be a nationally representative survey, the design must be stratified in a way that the selected sample is spread over the geography and population sub-groups of Liberia. While clusters of households are important to keep the costs manageable, this should not be done at the cost of damaging the reliability of the data. The sample size must optimally balance concerns over precision and survey costs. The sample frame to be used must be as complete, accurate and current as possible and the sample selection techniques that reduce unintentional biases caused by the implementers should be used. A listing activity needs to be undertaken in the selected enumeration areas (EA) so that a random sample of households can be selected within each EA. The design should finally be self-evaluating such that sampling errors can be estimated to guide the users in the reliability of the key results and the weights take into account any non-response that is generated through the course of the survey.

The steps involved in designing a good sample are as follows:

- Development of a Sampling Frame: In the case of Liberia, data from the2008 Population and Housing Census will be used as sample frame.

- Selecting Sampling Units: This consists of sorting the sample frame according to any desired implicit stratification criteria and selecting the required number of primary sampling units in each stratum with probability proportional to size.

- Planning the field assignments: The selected clusters need to be distributed among the field teams and the order in which they will be visited throughout the year needs to be decided.

- Dwelling Listing and Cartographic Updating: A new listing of households will be needed in the selected clusters, since it has been four years since the last Census was conducted in Liberia. This activity involves sending field teams to each of the

selected clusters with a GPS, recording at minimum, information on name of head of household, address of the household (including GPS measurements), household size disaggregated by gender and mobile phone numbers. Often, a sticker with a number is put on a dwelling to keep track of the households. Other times, information on key landmarks, occupation of head of the household and other relevant information is included. All of this needs to be geocoded using Geographic Information Systems (GIS) techniques.

- Select Dwellings in each cluster: A sample of the same number of household per cluster needs to be generated, along with reserves in case a household has to be replaced.

LISGIS is currently in the process of hiring a Sampling Expert to design the 2012 HIES sample. Selected EAs will be canvassed and all households within them will also be listed, out of which a random selection of households will be interviewed. The sample not only has to be nationally representative, but also allow for disaggregated county level estimates. A two stage clustered sampling model is anticipated with one nationally representative cycle. While designing two nationally representative cycles may be possible and would allow for national level half yearly estimates, county level half yearly estimates are not possible. It is ideal to start the data collection efforts at the beginning of the dry season (November 2012) since one of the goals of this dataset is to provide national level agricultural estimates. The size of the sample is expected to be about 8000 households at minimum, if all of the criteria listed need to be accounted for.

2. *Questionnaire Design*

Those who have never had to analyze data from a questionnaire that they have not developed themselves might think designing a questionnaire is an easy task. The process of defining the context of the questionnaire must be driven by analysts and by policy needs. Formatting a questionnaire is a complex art and proper formatting is

11

critical to survey success. It must be done by the survey planners and not relegated to clerical staff. Several steps are involved during questionnaire design.

- Identification of Policy Relevant Topics: The main issues to be addressed by the survey should be made explicit as early as possible. Consultation with various stakeholders in the government as well as donor communities will be held early on. The issues to be addressed also depend on the level of financial investment in the survey. The document, "A Medium Term Economic Growth and Development Strategy – The Agenda for Transformation" (AfT) will be used to identify policy issues to be addressed by the survey.

- Prepare Draft of the Questionnaires: The challenge here lies in the intellectual translation of all relevant concepts and policy issues into concrete questions and relevant modules. The mechanical part of physical production of a lengthy document requires the survey manager to use efficient processing software and requires about 2-3 months to complete. The questionnaire will be prepared from scratch, and a recall model will be utilized, particularly on the consumption expenditure section.

- Distribute Draft of the Household Questionnaire/Organize a seminar/Finalized Questionnaire: This activity can take anywhere between 2 weeks to a month and is required so that donor partners, subject matter specialists in relevant agencies, ministries and academics can analyze the draft. Comments received from stakeholders needs to be incorporated into the questionnaire. Subsequently, a seminar needs to be organized to receive a final set of feedback from all relevant stakeholders. Seminars are usually efficient since stakeholders can be brought in together to have an in depth conversation about the content of the questionnaire and receive last set of comments and feedback. A final version of the questionnaire will be drafted by incorporating the comments from the stakeholder's seminar and no further changes will be made beyond this point, aside from crucial ones that may emanate from the pre-testing activities.

At present, the questionnaire is anticipated to contain the following modules:

A household module with the following sections: Household Member Roster, Education, Health, Non-farm Employment, food consumption outside the household, Food Security, Housing/Water/Sanitation, consumption of food in the past week, non-food expenditures (past week, past month and past twelve months), Household Assets, Credit, Finance, Other Assistance and Group memberships.

An Agricultural Questionnaire with the following sections: A plot roster, plot details, crops by plot, crop production and sales, inputs, Processed Agricultural Products and Byproducts, Livestocks, Livestock by products, Forestry and Fishing (calendar, labor, input, output and trading).

A Community Questionnaire which at minimum includes a pricing sheet with farmer market prices for all agricultural products traded. Some other broader modules for indicators at the community level may be included. This questionnaire has to be prepared for a period in the middle of the household survey time in the locality.

Of key relevance to this project is the method in which consumption expenditure data are collected, since this will be used in the creation of weights for CPI and get information on household expenditure patterns for updating the National Accounts. Consumption expenditure data collection methods vary across three dimensions: the use of diary vs. recall, the level of aggregation or detail in the commodity list, and the reference period (Beegle et al., 2010).

- Diary vs Recall: A traditional diary based collection of consumption data requires households to be provided a diary to keep track of their daily consumption and expenditure. In some developing countries such as Brazil, China, and othersin Central Europe and Central Asia (where literacy rates are high), the diary based data collection is practiced. This contrasts with the more common practice in Living Standards Measurement Study (LSMS) surveys and other multi-topic survey instruments to base data collection on recall over a certain period. Both methodologies have advantages and disadvantages as listed in below

	Diary	Recall
Advantages	1. More detailed than Recall, therebyminimizing recall error	1. Less expensive than Diary
	2. Fewer telescoping errors than Recall	2. Easier supervision than Diary
	3. Traditional method of consumption data collection for HIES type of surveys, but in countries with high literacy rates.	3. Fewer days of field work per household than Diary
		4. Lower rates of non-response
		5. Lower chances of double counting

	Diary	Recall
Disadvantages	1. Relies heavily on literacy levels of household and motivation to maintain diaries	1. Telescoping Errors - over-reporting by fitting consumption over longer period of time into the reference period
	2. Requires support from enumerators - essentially boils down to recall	2. Recall Error - some under-reporting
	3. More Expensive than Recall	
	4. Difficult to supervise from a fieldwork perspective	
	5. More days of fieldwork per household than Recall	
	6. Household diaries - inability to capture private consumption	
	7. Double counting	
	8. Higher rates of non-response	

It is recommended that a recall module is implemented for the HIES, given the numerous challenges with the implementation of diary based methods in Liberia, particularly taking into account low literacy levels and steeply increasing costs associated with its implementation.

- The level of aggregation or detail in the commodity list:Another important issue is to decide the number of items about which data are collected within a consumption module. National consumer price index baskets often have in excess of 300 items. However, shorter lists of items are sometimes created by focusing on those items that represent the greatest share of consumption and aggregating other, less important, items into categories reducing the burden on respondents

and survey activities in general. While, a greater level of detail prompts respondents to remember more completely and accurately their consumption, the costs may be high: longer interview time, greater respondent fatigue and higher non-response. In most countries, while the reported consumption values have been higher with more items in the list, consumption decreased proportionally, implying that the ranking of households remains constant.

Since the primary goal of this survey is to provide consumption expenditure information for CPI weights and National Accounts Household Expenditure upgrading, a detailed consumption and expenditure module is anticipated. The basket of goods needs to be determined in conjunction with the National Accounts consultant.

- Reference Period:Lengthening or shortening the reference period within a survey can have ambiguous effects on the source and direction of bias. Respondents may have difficulty recalling consumption expenditure with longer reference periods due to diminished capacity to remember. On the other hand, short recall periods may produce over-estimates if respondents include expenditures just outside the reference period.

  Beegle et al. (2010), upon comparing different recall reference periods conclude that, "the savings in survey time from a reduced number of consumption categories in the recall list (a collapsed list) is minimal, compared with a substantial cost in terms of loss in accuracy. On the other hand, the subset list when scaled up based on other data performs very well in comparison with the long list. In addition, the hypothetical "usual" month recall almost doubles the interview time while most likely reducing the accuracy of measured consumption. These considerations would lead to a recommendation of a long-list recall module with a reference period of 1 or 2 weeks. Nevertheless these two variants, while not significantly more costly and relatively more accurate, are also problematic and likely subject to recall and telescoping errors of varying degrees as well as presumably the inability to capture personal out-of-household consumption. Even

though the 7-day recall module comes closest to the "gold" standard in terms of an estimate of overall mean consumption, we conclude that it is likely subject to either net telescoping or deliberate misreporting that increases in magnitude with the wealth of the respondent. Net negative reporting error also appears to increase with household size and the presence of an interpreter regardless of module".

While a final consensus has not been reached on the duration of the recall period, a combination of 7 day, one month and 12 month recall is anticipated, with items less frequently purchased (such as televisions, cell phones) appearing in the longer recall periods.

3. *Time table for field teams, and data collection plan.*

Once the sample size and the EAs to be visited have been finalized, a time table needs to be created for field teams to visit the EAs. It is important that every month, the ratio of EAs per strata visited is kept constant to account for seasonality. The best way to do this is to divide the total number of EAs in each stratum by 12 and ensure that that many EAs are visited per strata per month. The plan should try to minimize travel costs, so teams should not be assigned two EAs one after the other that lie in different parts of the country. One way to get around the problem is to subdivide all the EAs by parts of the country (for example north, south, east and west) and assign teams EAs by parts.

Each team should be given a calendar with the name of the EAs, date of visit to the EA, roster of households to be interviewed in each EA along with information collected during the listing period to locate the households.

4. *Pilot/Pretest*

Once the HIES questionnaires have received approval from the various stakeholders involved in the designing of the questionnaire, logistical arrangements need to be made for piloting (also known as pretesting) the questionnaire. This includes selecting and briefing a small number of experienced interviewers who will conduct the field

test. Their transportation, lodging and other costs should be covered along with enough printed copies of the questionnaire. Roughly 200 questionnaires should be piloted, so that there is enough data to be used for Data Entry Operator Training, troubleshooting questions that were problematic, preliminary analysis of the data obtained etc. Roughly four weeks should be allowed for field testing of the questionnaire along with a week or two for analysis of those received. It is extremely important that the PIC is closely involved in this activity.

5. *Data Entry and Double Entry Software Preparation*

This activity is subdivided into several parts.

- Development of first version of Data Entry Program: The development of the first version of the Data Entry Program (DEP) should begin as soon as the survey questionnaires are ready and this usually is accomplished shortly after the pre-test. In our case, the data entry program being used is CSPro. It usually takes about four weeks to develop the first version of the Data Entry Program. It is crucial that the CSPro software expert is able to assign fields and ranges for all variables along with the corresponding intra-record checks. The survey data manager is the main person responsible for testing and debugging the program thoroughly. However the first test of the program will come into effect during training of the data entry operators and when the first set of questionnaires will be completed and entered. It is important to note that the final program should be able to provide an error report such that at the end of entering a questionnaire, an array of problems with the data can be viewed. This should be used by supervisors to verify the quality of data and to do a revisit to the household to fix any problems that exist.

- Development of Data Entry Manual: The Data Entry Manual will take about two weeks to write and should be a comprehensive document that the data entry personnel can refer to for any questions they may have while using the program. This should be written by the person that designs the CSPro software.

- Computer Installation and Data Entry Operator Training: We intend to do field based first data entry (FDE), which means that each team will be allocated a data entry operator that will enter the questionnaires into the data entry program on the field. After verification procedures are met, they will be sending the data using internet modem sticks to the LISGIS headquarters. This means that every field team should have access to a laptop with a long lasting battery (or solar powered), a printer to view the error list and printer cartridges. CSPro needs to be installed in each of these computers and the data entry operators should be trained using both theoretical and practical sessions. This will involve the data entry operators to enter completed questionnaires arriving from the pilot/pretest.

- Preparation of Second Data Entry (SDE): The Second Data Entry program (SDE) which will mimic the first data entry program will be used for re-entering of the questionnaires received from the field. There are many advantages of undertaking this activity. Firstly, with data entry operator fatigue, some of the values entered might be incorrect, even though they were correctly noted in the paper questionnaire. Those cases cannot be troubleshooted without comparison with a SDE. Another advantage with having two versions of data entered is that you can monitor data entry operator performance on the field versus SDE. Finally, the FDE and SDE will be compared observation to observation and for those values that do not match, questionnaires will be pulled manually and the correct values will be recoded.

6. *Preparation of Field Manual*

Three types of field manuals will be created as part of this project.

- Supervisor Manual: This manual outlines the objectives, methodology and organization of the survey explicitly alongside the duties and responsibilities of the supervisor and the ways in which the supervisor is expected to be connected with the core management at LISGIS headquarters.

In addition, the procedures for carrying out fieldwork in each cluster including identification of households, liaising with local authorities for their cooperation through the project, ensuring that the questionnaires are completed and verifiable, how to deal with refusals and other forms of non-response. The relationship between supervisors and enumerators also will be outlined in the document along with forms for enumerator evaluation and interview check-up.

The manual will also specify procedures for coding open ended questions and a complete code list to be used for occupations, activities and geographic locations. Finally, the manual will outline how the print outs after first data entry needs to be interpreted along with error reports and procedures for call back interviews to the households.

- Enumerator manual: The main objective of the enumerator manual is to provide a detailed explanation of all the concepts and definitions for each question embedded within the questionnaire, define field procedures and ensure that all questions that are not self-explanatory have a uniform criteria. All sections of the questionnaires will be included, along with the survey's methodology and objective and supervisor-enumerator relationship.

- Data Entry Operator Manual: This manual will discuss the role and responsibility of the first and second data entry operators along with technical details on how the program works – entering data, deleting data, modifying data, zipping and sending data. The manual will also describe the relationship between the supervisor and enumerator, including how the error reports need to be interpreted and communicated to the supervisor.

b) Fieldwork Implementation

1. *Enumerator/Supervisor/Data Entry Operator Training*

Training is critical to the success of the HIES since it will contribute greatly to the quality of the data collection effort. Several training workshops are envisioned as part of training activities as listed below

Supervisor Training: Supervisors will be involved in the project from the time of pre-testing. All aspects of the job will be presented formally, such as the HIES objectives, sampling, contents and design of the survey, structure of the interviews, structure of management team, quality control criteria and household replacement criteria. The training usually takes about one to two weeks.

Enumerator and Data Entry Training: There will be some overlap between the enumerator and data entry training because the data entry operators need to have a good understanding of the questionnaire as well so that they can troubleshoot the errors more effectively. The training period is therefore anticipated to be about four weeks for the HIES – with two weeks of in class training which will be common to both groups and two weeks of field training. The field training for enumerators typically includes going to a cluster and performing a set of interviews to familiarize them with the questionnaire tool. The Data Entry Operators on the other hand will learn how to use the CSPro software for data entry and use the questionnaires from the field to get practice with entering, troubleshooting, zipping and emailing the data.

It is important that the training not only covers the basic structure of how to understand the questionnaire, but also covers the economic concepts involved, particularly in sections such as labor, agriculture and consumption modules which will be critical information for the CPI, NA and poverty numbers.

2. *Data collection*

They survey should be fielded as soon as possible after the supervisors, enumerators and data entry operators are trained and the logistical arrangements for their travel (itinerary, order of clusters, vehicular arrangements etc.). The expected start date for the fieldwork is November 1, 2012 and will go on for a year, until October 30, 2013.

At present, 8000 households are anticipated to be covered across the country with roughly 10 households per enumeration area (EA). Assuming uniform coverage of EA's across the year involves surveying roughly 67 EA's per month. A sample calendar for one team comprising of one driver, one supervisor, five enumerators and one data entry operator is included in Appendix 1. Given that it will take about 5 days to arrive at a new EA and complete the interviews, one team can cover about 6 EAs in a month. To arrive at the target of 67 EAs every month, 11 teams will be needed.

The field teams are expected to send the data as soon as they finish a cluster via email. Physical questionnaires will be collected from field teams as a separate effort, either during site supervision visits by staff atLISGIS Headquarters or other partners or by sending vehicles from County Offices of LISGIS. All questionnaires will be sent to the LISGIS Headquarters for additional processing

Giving small gifts to households that have taken time to answer questions for this survey is recommended since household members have to take out a significant amount of time from their schedule to act as respondents to the questionnaire. Typical gifts could include bednets or radios.

Data verification is important at this stage since regular feedback needs to be given to the teams on their performance. There are two ways to accomplish this:

- Field visits: Multiple field trips through the course of the fieldwork should be organized such that the core team at LISGIS headquarters can go and check on the work of the enumerators and supervisors. These visits should include sitting

through enumerator interviews, observing supervisor interaction with the community where the interviews are being conducted, supervisor-enumerator interactions and most importantly asking for feedback on field supplies, adequate support from the central staff. Feedback from the field teams is critical to ensure an open communication channel between all parties. It also provides insight on responses to specific questions and advise on revisions on future HIES activities.

- Weekly data checks: Since the field teams will be required to send data from remotely as soon as they finish an EA, the Project Implementation Committee will be spending some time every week looking at the data received, after it is converted into STATA. Each field team will have one counterpart in the Project Implementation Committee who will analyze the data received from the field. An automated program will be provided to those analyzing the data received from the field so that they can look at ranges, non-responses and some other errors coming from the field and provide weekly feedback to the field teams on areas of improvement.

Updates to the data entry software may be needed from time to time, which will come into light through the course of the survey process. These updates should be made available to the field teams in the most efficient and easiest possible manner.

3. *Field data entry and Second Data Entry*

As already mentioned, each field team will be assigned one data entry operator for entering the data on the field team. The field based data entry operator is expected to enter all the questionnaires from one EA and send it to the Project Implementation Committee via email before leaving for the next EA. Field based data entry has served beneficial because the program is inbuilt with checks and allows for verification of data and the possibility of call backs if need be.

Second data entry will occur once the questionnaires are sent from the field to the headquarters. Another set of data entry operators will be hired to re-enter the questionnaires received from the field. This activity will be done independently from the headquarters. The Second Data Entry Operators will undergo the same training as the field based data entry operators. However, their job is only to re-enter the data from the questionnaire into the software and NOT make any corrections to the paper questionnaire.

The data entered in the field will be compared with the data entered from the Second Data Entry exercise in the cleaning phase.

4. *Data cleaning*

Once the data have been entered and second entered, it should be subjected to four kinds of quality checks: range checks, skip checks, consistency checks, and comparison between 1DE and 2DE.

- Range Checks: Every variable in the survey will be checked for ranges. For example a Yes/No question should only have the legal codes (1 and 2) and no other values entered. Chronological variables such as dates should contain valid entries. For example, February 29 can only be allowed in a leap year survey. While these will also be inbuilt into the data entry program, it is good practice to double check ranges. This will be particularly critical for the consumption categories where units and values will often not seem consistent. This also includes scanning for typographical errors.

- Skip Checks: Skip checks verify if a skip has been followed appropriately. For example, when certain questions only need to be administered to women, there is usually a filter question at the beginning asking for the gender of the person. A Skip check looks at whether subsequent questions that pertain to women were administered to groups that weren't meant to get it. Likewise it can also check if anyone that belonged to that group were not administered those questions.

- Consistency Checks: Consistency checks verify that the values from one question are consistent with values from other questions. This can be within one module or across different modules. Also some of these checks can be across the same unit of observation while other times information needs to be compared from two or more different units of observation. There is no natural limit on the number of consistency checks that can exist. In general though, more number of checks defined is directly proportional to better quality data.

- Comparison between 1DE and 2DE: Another extremely important way of cleaning data is to check for any inconsistencies between 1DE and 2DE. If the values for the same question for the same unit of observation are different, then these questionnaires need to be pulled out for the specific question and corrected. Since 1DE is done in a more constrained environment (time and space), 2DE is generally more reliable, but a difference in values between the two should be verified. 2DE can be used later for more cleaning if need be.

5. *Creating weights*

Once the data collection activities have finished and the data has been cleaned, the sampling specialist needs to assess the need for weighting the data to account for sampling errors in order to arrive at unbiased estimates. In order for the weights calculations for the data to be accurately determined, all stages of the sampling must be carefully recorded and made available to the sampling consultant. The consultant in particular will prepare precise instructions, including the formulae, for the modification of the weights in order to take into account refusal rates, households that are removed from the sample (outliers) and/or any other modifications.

c) **Analysis and Dissemination**

1. *Reconstruction of the CPI using weights from the HIES.*

Since one of the goals of the HIES is to provide weights for the basket of goods for the HIES and to rebase the CPI, this activity is of paramount importance. This work will be undertaken by contracting an external consultant.

2. *Poverty Analysis*

   A second goal of this survey is to shed light on the several dimensions of poverty in Liberia including, who is poor, where poverty is located, how do the poor earn their living, their access to and use of government services and subsidies, construction of a poverty line, and also understanding the causes of the poverty (welfare analysis). This work will also be undertaken by contracting an external consultant.

3. *Update on National Accounts.*

   The Household final consumption expenditures are the largest component of the expenditure account. For the preparation of independent estimates of the expenditures in the National Accounts, a proper assessment of household expenditures is therefore very important. The HIES provides the sole source for the estimation of household consumption expenditure and is therefore indispensable. A consultant will be contracted for work on updating the National Accounts.

4. *Other sector specific studies*

   Another importantgoal of this survey is to track the effectiveness of various interventions and policies in the fight against poverty in Liberia. It provides information on a range of socio-economic indicators, including household income and expenditure, housing conditions, economic activities, ownership of assets, and a range of social sector indicators including school attendance and the use of health services. This work will be undertaken by the poverty analysis consultant along with the consultant administering the HIES and the Project Implementation Committee.

5. *Agricultural Report*

   This document is intended to provide reliable national level statistics on agriculture, allowing for, among other things, the estimation of land areas, both owned and cultivated, self-reported production figures for main crops and livestock, and detailed

cost of production for crops at the household level. This activity will be undertaken by LISGIS.

6. *General Report*

This document is intended to be the official report emanating from the HIES data collection activity. It will include a detailed documentation of the entire process of the HIES along with the analysis that will be sourced from the CPI, Poverty, National Accounts and other sector specific studies. This document will be complied by the HIES consultant.

7. *Basic Information Document.*

The basic information document will include a synopsis of the questionnaires, a concise but complete description of the sample design, basic field quality control techniques, guidelines for using the data, documentation of constructed datasets, description of all files and references to other ancillary documents. This document will be created by the HIES consultant.

8. *Removal of all identifiers from the data*

It is crucial that before any data dissemination, all identifiers such as names, GPS coordinates of households, locational addresses, etc. are removed to protect the identity of the households that were interviewed. Two versions of the dataset will therefore be maintained, one which has all the information which is for internal use only by LISGIS. The second will be for external use and dissemination, which will exclude all the identifiers.

9. *Data dissemination – posting on website / launch event.*

The HIES is intended for use by multiple stakeholders including donors, government, researchers and non-profit organizations. It is therefore critical that an open data access policy is maintained so that anyone can use the data for legitimate purposes. It is also important to keep track of the users of this data, so an online form

must be prompted for individuals using this data to provide information about themselves and their reasons for using this data. It is crucial to maintain the confidentiality of those that submit their information on the use of their data.

The Basic Information Document, the datasets, the general reports and any other relevant documents should be released together and an agreement on where the data will be hosted needs to be established. Negotiation on data dissemination techniques between the various stakeholders needs to be well documented.

A launch event is recommended to increase awareness on this survey.

d) Institutional Arrangements and Project Management

Liberia Institute of Statistics and Geo-Information Services (LISGIS)assumes overall responsibility for the 2012 Household Income and Expenditure Survey. In doing, so they will report progress and seek technical advice and financial approvals through the Steering Committee and Project Technical Committee (described below).

1. *Establishment of Steering Committee and Project Technical Committee*

The Steering Committee comprises of the core organizations that will act as an advisory board to the 2012 Household Income and Expenditure Survey. One or two representatives from each organization within the Steering Committee will serve on the Project Technical Committee for the HIES. The Project Technical Committee is expected to meet at least every quarter to get updates on the progress of the survey and provide feedback on the ongoing activities pertaining to thethe execution on each component of the 2012 HIES.

At present the following organizations have been recommended to serve on the Steering Committee of this project by LISGIS. Individual representatives from each organization forming the Project Technical Committee will be selected on a later date. The PTC will be chaired by the Director General of LISGIS, Dr. T. Edward Liberty, while the Chief Technical Advisor to the committee will be Mr. NagrajRao, a

consultant contracted by LISGIS financed by the World Bank Trust Fund for Statistical Capacity Building (for one year) to oversee the activities falling under this project.

Serial No.	Member Organization	Role
1.	Ministry of Planning and Economic Affairs	Chair
2.	LISGIS	Vice Chair
3.	Ministry of Finance	Member
4.	Central Bank of Liberia	Member
5.	Ministry of Agriculture	Member
6.	Ministry of Gender and Development	Member
7.	Ministry of Internal Affairs	Member
8.	Ministry of Commerce and Industry	Member
9.	Ministry of Labour	Member
10.	Ministry of Health	Member
11.	Ministry of Education	Member
12.	University of Liberia	Member
13.	International Monetary Fund	Member
14.	World Bank	Member
15.	United Nations Development Programme	Member
16.	World Food Programme	Member
17.	Food and Agriculture Organization	Member
18.	United Nations Population Fund	Member
19.	USAID	Member
20.	African Development Bank	Member
21.	European Commission	Member
22.	SIDA	Member
23.	UNICEF	Member

The list is bulky at present and a decision needs to be made on restricting the number of stakeholders to those that are relevant to the survey and/or will be funding the activities of this survey.

2. *Establishment of Project Implementation Committee*

A Project Implementation Committee will be formed within LISGIS to oversee all the technical and day to day activities related to the survey. The Resident Advisor to the Project Implementation Committee will be Mr. NagrajRao, a consultant contracted by LISGIS to oversee the activities falling under this project. Mr. Kormay Adams will serve as the Project Director for the Project Implementation Committee. The members of this committee are listed below

Title	Name
Project Director	Mr. Kormay Adams
Resident Advisor	Mr. NagrajRao
Project Coordinator	Ms. Mariah Q Gilayeneh
Asst. Project Coordinator	Mr. Boima H.M. Sonii
Local Consultant	Charles Akoi
GIS Technician	Mr. Thomas Davis
Data Manager	Mr. Joseph Nyan
Asst. Data Manager	Mr. Kwia Wilson
Accountant	Mr. Youngor Amara
Secretary	James Belleh
Public Relation Officer	TBD
Office Assistant	MavilahDedeh Johnson
Project Drivers	WehyeeLawa, John Katakpah

3. *Recruitment of Field Supervisors and Enumerators*

In addition to the Technical and Steering Committees, field supervisors and enumerators need to be interviewed and selected based on their experience with administering similar surveys. It is recommended that all the hiring occurs through the main LISGIS office so that the field supervisors and enumerators are known to the Technical Committee and take direct instructions from them. It is always better to hire a larger pool of interviewers for the training phase to allow for a final set of

enumerators, based on their performance during the practical part of the training. Selection processes usually take 3 to 6 weeks and sometimes longer if interviewers with specific geographic, ethnic or linguistic backgrounds are needed.

e)   Time-table for all activities

The timetable for all activities is included in Appendix II. It is recommended that all activities are followed as closely to the time-table as possible in order to avoid any delays.

VIII.   Financing

The budget for the HIES can vary depending on local factors such as whether or not items are provided in kind, the size of the staff needed to implement the survey (which is a function of sample size and length of questionnaire) and price of items. Having a detailed budget that covers all financial aspects of implementing will help not only with fundraising but also streamline and manage the funds received. The anticipated budget for the 2012 HIES is included in Appendix III.

A strategic fundraising plan needs to be implemented because a survey of such magnitude relies on finances through several different sources. The financing agreements need to be established as soon as an initial budget is drafted and should basically make clear, the specific donors for each stage, when and how the funds will be transferred, and what administrative procedures will be used for spending the money. It is important not only to receive confirmation from donors on the amounts committed, but also streamline the process and timing of transfer of money from each source. A survey like the HIES requires financing to be finalizedas early as possible so that any pertinent equipment needed for the survey, particularly those that need to be imported can be purchased at the earliest.

REFERENCES

Beegle, De Weerdt, Friedman, Gibson (2010). "Methods of Household Consumption Measurement through   Survey: Experimental Results from Tanzania". World Bank Policy Paper.

Government of Liberia (2012)."A Medium Term Economic Growth and Development Strategy – The Agenda for Transformation 2012-2017" (Draft).

Grosh and Munoz (1996)."A Manual for Planning and Implementing the Living Standards Measurement Study Survey".Living Standards Measurement Study Working Paper.

Grosh and Glewwe (2000)."Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 years of the Living Standards Measurement Study".Volume 1, 2 & 3.

Liberia Institute of Statistics and Geo-Information Services (2010). "Production Estimates of Major Crops and Animals" (Draft).

APPENDIX 1: SAMPLE CALENDAR FOR ONE TEAM

DAY	Vehicle	TEAM 1
DAY 1	TRAVEL TO EA1	
DAY 2	STAY IN EA1	EA1
DAY 3	STAY IN EA1	EA1
DAY 4	STAY IN EA1	EA1
DAY 5	CALL BACKS	EA1
DAY 6	TRAVEL TO EA2	
DAY 7	STAY IN EA2	EA2
DAY 8	STAY IN EA2	EA2
DAY 9	STAY IN EA2	EA2
DAY 10	CALL BACKS	EA2
DAY 11	TRAVEL TO EA3	
DAY 12	STAY IN EA3	EA3
DAY 13	STAY IN EA3	EA3
DAY 14	STAY IN EA3	EA3
DAY 15	CALL BACKS	EA3
DAY 16	TRAVEL TO EA4	
DAY 17	STAY IN EA4	EA4
DAY 18	STAY IN EA4	EA4
DAY 19	STAY IN EA4	EA4
DAY 20	CALLBACKS	EA4
DAY 21	TRAVEL TO EA5	
DAY 22	STAY IN EA5	EA5
DAY 23	STAY IN EA5	EA5
DAY 24	STAY IN EA5	EA5
DAY 25	CALL BACKS	EA5
DAY 26	TRAVEL TO EA6	
DAY 27	STAY IN EA6	EA6
DAY 28	STAY IN EA6	EA6
DAY 29	STAY IN EA6	EA6
DAY 30	CALL BACKS	EA6

TOTAL NUMBER OF EAS	800
TOTAL NUMBER OF EAS TO BE COVERED IN ONE MONTH	67
TOTAL NUMBER OF EAS COVERED BY ONE TEAM	6
TOTAL NUMBER OF TEAMS NEEDED TO COVER 67 EAS IN A MONTH	11
TOTAL NUMBER OF SUPERVISORS	11
TOTAL NUMBER OF ENUMERATORS	55
TOTAL NUMBER OF ENUMERATORS PER TEAM	5
TOTAL NUMBER OF DRIVERS	11
TOTAL NUMBER OF FIELD VEHICLES	11
TOTAL NUMBER OF FIELD BASED DATA ENTRY	11

APPENDIX 2: TIMELINE FOR THE 2012 HOUSEHOLD INCOME AND EXPENDITURE SURVEY 2012, LIBERIA

	2012								2013												2014			
	May	June	July	August	Sept	Oct	Nov	Dec	Jan	Feb	March	April	May	June	July	August	Sept	Oct	Nov	Dec	Jan	Feb	March	April
PREPARATION																								
Sampling		■	■																					
Household Listing		■	■	■																				
Questionnaire Design					■																			
Time table for field teams, and data collection plan						■																		
Pilot/Pretest					■																			
Data Entry and Double Entry software Preparation				■																				
Preparation of Field Manuals					■																			
FIELDWORK																								
Enumerator/Supervisor/Data Entry Operator Training						■																		
Data Collection							■	■	■	■	■	■	■	■	■	■	■							
First Data Entry							■	■	■	■	■	■	■	■	■	■	■							
Second Data Entry										■		■			■			■						
Data Cleaning																			■	■				
Creation of Weights																					■			
ANALYSIS AND DISSEMINATION																								
Reconstruction of the CPI using weights from the HIES.																						■	■	
Poverty Analysis																						■	■	
Update on National Accounts																						■		
Other sector specific studies																						■		
Agricultural Report																						■		
General Report																						■		
Basic Information Document																						■		
Removal of all Identifiers from Data																							■	
Data dissemination – posting on website / launch event																								■
INSTITUTIONAL ARRANGEMENT AND PROJECT MANAGEMENT																								
Agreements and institutional environment	■	■																						
Establishment of Project Technical Committee and Steering Committee	■	■																						
Recruitment of field supervisors and enumerators			■	■	■																			
FINANCES																								
Budgeting and Fundraising		■	■																					
Acquiring items needed for implementation - (computers, survey materials, cars etc)				■	■																			

APPENDIX 3: COST BREAKDOWN FOR THE SURVEY

CATEOGIES OF COSTS	TOTAL	SUB-CATEGORY	BREAKDOWN
Category 1: Preparation and Fieldwork Costs	$1,285,520.42	Training	$47,084.00
		Listing	$282,796.67
		Pretesting	$9,325.00
		Questionnaire Printing	$50,600.00
		Stationaries and Supplies	$66,336.25
		Equipment and Logistics	$829,378.50
Category 2: Project Management and Consultancy Costs	$1,427,186.00	Over time for LISGIS Staff & Per-Diems for field work	$111,385.00
		Consultant Fees (Technical Assistance)	$540,416.00
		Field Supervisors and Enumerators	$775,385.00
Category 3: Data Cleaning and Processing Costs	$16,310.00	Data Processing	$16,310.00
Category 4: Data Dissemination Costs	$2,500.00	Dissemination, Workshops	$2,500.00
Category 5: Contingency	$134,080.82	5%	$136,575.82
TOTAL	**$2,865,597.24**		**$2,868,092.24**

DEDUCTIONS	TOTAL	DESCRIPTION	BREAKDOWN
Government of Liberia	$500,000.00	Commitment from Government of Liberia	$500,000.00
World Bank	$200,000.00	World Bank Support for Technical Assistance	$200,000.00
TOTAL	**$700,000.00**		**$700,000.00**
		REMAINDER BALANCE	**$2,168,092.24**