

The Uganda National Panel Survey (UNPS) 2009/10

Basic Information Document

**UGANDA BUREAU OF STATISTICS
P. O. BOX 7186
KAMPALA UGANDA**

Acronyms

BHPS	British Household Panel Survey
EA	Enumeration Area
GoU	Government of Uganda
GPS	Global Positioning System
ISCO	International Standard Classification of Occupations
ISIC	International Standard Industrial Classification
LC1	Local Council 1
LSMS-ISA	Living Standards Measurement Study – Integrated Surveys on Agriculture
NAADS	National Agricultural Advisory Services
NDP	National Development Plan
NDS	National Development Strategy
NSDS	National Service Delivery Surveys
PSID	Panel Study of Income Dynamics
UBOS	Uganda Bureau of Statistics
UDHS	Uganda Demographic and Health Survey
UNHS	Uganda National Household Survey
UNPS	Uganda National Panel Survey

TABLE OF CONTENTS

LIST OF TABLES	4
1 OVERVIEW	5
1.1 SURVEY OBJECTIVES	5
1.2 SURVEY DESIGN	6
2 SURVEY QUESTIONNAIRES – REVIEW OF SECTIONS	7
2.1: EXPLANATORY NOTES BY SECTION – HOUSEHOLD QUESTIONNAIRE	7
<i>Section 1A: Household Identification Particulars.....</i>	<i>7</i>
<i>Section 1B: Staff Details and Survey Time.....</i>	<i>7</i>
<i>Section 2: Household Roster.....</i>	<i>7</i>
<i>Section 3: General Information on Household Members</i>	<i>8</i>
<i>Section 4: Education.....</i>	<i>8</i>
<i>Section 5: Health</i>	<i>9</i>
<i>Section 6: Child Nutrition and Health</i>	<i>9</i>
<i>Section 7: Disability.....</i>	<i>9</i>
<i>Section 8: Labour Force Status.....</i>	<i>9</i>
<i>Section 9: Household and Housing Conditions</i>	<i>10</i>
<i>Section 10: Energy Use.....</i>	<i>10</i>
<i>Section 11: Household Incomes</i>	<i>10</i>
<i>Section 12: Non-agricultural Enterprises/Activities.....</i>	<i>10</i>
<i>Section 13: Financial Services</i>	<i>11</i>
<i>Section 14: Household Assets</i>	<i>11</i>
<i>Section 15: Household Consumption Expenditure.....</i>	<i>11</i>
<i>Section 16: Shocks and Coping Strategies</i>	<i>11</i>
<i>Section 17: Welfare Indicators and Food Security.....</i>	<i>11</i>
<i>Section 18: Transport Services.....</i>	<i>12</i>
2.2: EXPLANATORY NOTES BY SECTION – AGRICULTURE QUESTIONNAIRE	14
<i>Section 1A: Household Identification Particulars.....</i>	<i>14</i>
<i>Section 1B: Staff details and survey time</i>	<i>14</i>
<i>Section 2: Current land Holdings and land that the household has access through use rights</i>	<i>14</i>
<i>Section 3A & 3B: Agricultural and labour inputs.....</i>	<i>15</i>
<i>Section 4A & 4B: Crops grown and type of seeds used.....</i>	<i>15</i>
<i>Section 5A & 5B: Quantification of Agricultural Production.....</i>	<i>15</i>
<i>Sections 6A, 6B & 6C: Livestock ownership</i>	<i>15</i>
<i>Section 7: Livestock expenditure and income.....</i>	<i>15</i>
<i>Section 8: Livestock Products</i>	<i>16</i>
<i>Section 9: Fishing.....</i>	<i>16</i>
<i>Section 10: Extension Services.....</i>	<i>16</i>
2.3: EXPLANATORY NOTES BY SECTION – COMMUNITY QUESTIONNAIRE	18
<i>Section 1: Community Identification Particulars</i>	<i>18</i>
<i>Section 2: Availability of services within the community.....</i>	<i>18</i>
<i>Section 3: Education (Primary).....</i>	<i>18</i>
<i>Section 4: Health services.....</i>	<i>18</i>
<i>Section 5: Works and Transport.....</i>	<i>18</i>
2.4: EXPLANATORY NOTES BY SECTION – WOMAN QUESTIONNAIRE	21
<i>Section 2A: Contraception</i>	<i>21</i>
<i>Section 2B: Birth History</i>	<i>21</i>
3 OTHER RELATED INSTRUCTIONS/CODES	22
3.1: AREA MEASUREMENT USING GLOBAL POSITIONING SYSTEM (GPS).....	22

3.2:	OTHER CODES	22
4	FIELD WORK ORGANIZATION	23
4.0:	TRACKING OF HOUSEHOLDS	24
4.1:	TRACKING OF HOUSEHOLDS	24
4.2:	TRACKING OF SPLIT-OFFS	24
5	LINKING UNHS 2005/06 & UNPS 2009/10	25
6	CALCULATION OF UNPS 2009/10 PANEL WEIGHTS	26
6.1:	BASE WEIGHTS FROM 2005 SAMPLE	28
6.2:	PROBABILITY OF SELECTION INTO UNPS	28
6.3:	PROBABILITY OF SELECTION INTO TRACKING	28
6.4:	FAIR SHARE CORRECTION	30
6.5:	POOLING	31
6.6:	ATTRITION CORRECTION FACTOR	31
6.7:	TRIMMING	33
6.8:	POST-STRATIFICATION	34
7	REFERENCES	35
	APPENDIX: CONFIDENTIAL INFORMATION, GEOSPATIAL VARIABLES.....	36

LIST OF TABLES

Table 1 :	Organization of the UNPS 2009/10 Household Questionnaire	13
Table 2 :	Organization of the UNPS 2009/10 Agriculture Questionnaire.....	17
Table 3 :	Organization of the UNPS 2009/10 Community Questionnaire.....	19
Table 4 :	Organization of the UNPS 2009/10 Woman's Questionnaire	21
Table 5:	EA Probability of Selection.....	28
Table 6:	Summary Statistics on Heads of Missing Households vs. Missing Individuals	32

1 Overview

Uganda has experienced strong economic growth over the past two decades, and has made great strides towards improving the quality of life and access to services. In order to continue to promote pro-poor economic growth, the Government of Uganda (GoU) developed the National Development Plan (NDP) and a Joint Budget Support strategy as part of the implementation of the National Development Strategy (NDS).

The GoU recognizes the need for adequate data collection to effectively monitor outcomes of the National Development Strategy (NDS). For this purpose, the Uganda Bureau of Statistics (UBOS) is implementing the Uganda National Panel Survey (UNPS) program, with financial and technical support from the Government of Netherlands, and the World Bank Living Standards Measurement Study – Integrated Surveys on Agriculture (LSMS-ISA) project.

The UNPS is a multi-topic panel household survey that commenced in 2009/10. One of the primary uses of the UNPS is to inform policymaking in advance of the Budget, through descriptive reports that are made ready in time for the initial work on sector budget framework papers.

1.1 Survey Objectives

The UNPS aims at producing annual estimates in key policy areas and at providing a platform for experimenting with and assessing of national policies and programs. Explicitly, the objectives of the UNPS include:

1. To provide information required for monitoring the National Development Strategy, of major programs such as National Agricultural Advisory Services (NAADS) and General Budget Support, and also to provide information to the compilation of the National Accounts (e.g. agricultural production);
2. To provide high quality nationally representative information on income dynamics at the household level and provide annual information on service delivery and consumption expenditure estimates to monitor poverty and service outcomes in interim years of other national survey efforts, such as the Uganda National Household Survey (UNHS), Uganda Demographic and Health Survey (UDHS) and National Service Delivery Surveys (NSDS);
3. To provide a framework for low-cost experimentation with different policy interventions to e.g. reduce teacher absenteeism, improve ante- and post-natal care, or assessing the effect of agricultural input subsidies;

4. To provide a framework for policy oriented analysis and capacity building substantiated with the UGDR and support to other research which will feed into the Annual Policy Implementation Review; and
5. To facilitate randomized impact evaluations of interventions whose effects cannot currently be readily assessed through the existing system of national household surveys.

1.2 Survey Design

The UNPS is carried out annually, over a twelve-month period on a nationally representative sample of households, for the purpose of accommodating the seasonality associated with the composition of and expenditures on consumption. The survey is conducted in two visits in order to better capture agricultural outcomes associated with the two cropping seasons of the country. The UNPS will therefore interview each household twice each year, in visits six months apart.

Starting in 2009/10, the UNPS has been set out to track and reinterview 3,123 households that were distributed over 322 enumeration areas (EAs), selected out of the 783 EAs that had been visited by the Uganda National Household Survey (UNHS) in 2005/06. The UNPS EAs covered all 34 EAs visited by the UNHS 2005/06 in Kampala District, and 72 EAs (58 rural and 14 urban) in each of the (i) Central Region with the exception of Kampala District, (ii) Eastern Region, (iii) Western Region, and (iv) Northern Region.

Within each stratum, the UNPS EAs were selected from the UNHS 2005/06 EAs with equal probability, and with implicit stratification by urban/rural and district (in this order), except for the rural portions of the ten districts that were oversampled by the UNHS 2005/06. In these districts, the probabilities were deflated, to bring them back to the levels originally intended. Since IDP camps are now mostly unoccupied, the extra EAs in IDP camps are not a part of the UNPS subsample. This allocation strives for reasonably reliable estimates for the rural portion of each region, and for the set of urban areas out of Kampala as a whole, as well as the best possible estimates for Kampala that can be expected from a subsample of the UNHS 2005/06. Therefore, the UNPS strata of representativeness include (i) Kampala City, (ii) Other Urban Areas, (iii) Central Rural, (iv) Eastern Rural, (v) Western Rural, and (vi) Northern Rural.

Prior to the start of the 2009/10 field work, 2 UNPS households were also randomly selected in each EA for the purposes of tracking baseline individuals that moved away from original locations since the UNHS 2005/06. The initial UNPS sample will be subject to three consecutive waves of data collection after which, parts of the sample will start to be replaced by new households extracted from the updated sample frames developed by the UBOS as part of the 2012 Uganda Population and Housing Census. In addition,

the UNPS will fit within the Long Term Census and Household Survey Program and therefore both the questionnaires and the timing of data collection will be coordinated with the current surveys and census implemented by UBOS. To suit its multiple objectives, the UNPS comprises of a set of survey instruments, namely:

- Household Questionnaire
- Woman Questionnaire,
- Agriculture Questionnaire, (administered to the subset of UNPS households engaged in agricultural activities)
- Community Questionnaire, and
- Market Questionnaire.

2 Survey Questionnaires – Review of Sections

As mentioned earlier, the UNPS had five questionnaires namely: Household Questionnaire; Woman Questionnaire; Agriculture Questionnaire; Community Questionnaire and Market Questionnaire. Each of these questionnaires is divided into a number of sections and the level of observation for each section varies accordingly. The tables 1-4 below provides an overview of the sections of the Household, Agriculture, Community and Woman questionnaires, associated data files, and key identifiers.

2.1: Explanatory notes by section – Household Questionnaire

Section 1A: Household Identification Particulars

Information in this section was distributed by the Headquarters staff to the field teams before starting data collection. Names and codes pertaining to the selected Enumeration Areas (EAs) were provided by UBOS to the team leaders prior to fieldwork. An EA generally does not have its own name but is known by the name of the Local Council 1 (LC1) that is associated with it.

Section 1B: Staff Details and Survey Time

The Supervisors, interviewers and data entry operators were all required to record their particulars in this section. Time taken to conduct interviews was recorded. The data also include the date on which the household questionnaire was administered in full.

Section 2: Household Roster

The purpose of this section is to:

- (i) Identify all persons who are members of the household;

- (ii) Provide basic demographic information such as age, sex and marital status of each household member; and
- (iii) Identify any changes to household members since the first visit

The respondent for this section was mainly the household head. In the absence of the household head the next person who is acting as household head would be interviewed. ***It was a requirement that respondent must be a usual member of the household and should be capable of providing all the necessary information about other members of the household.*** Other household members also helped in providing information or details on particular questions concerning them.

In UNPS 2009/10, a **household** was defined as a group of people who have ***normally*** been living and eating their meals together for at least 6 of the 12 months preceding the interview. Therefore, the member of the household is defined on the basis of their usual place of residence.

Section 3: General Information on Household Members

This section captured general information on all members of the household specifically on:

- (i) Parents of household members who sometimes do not live in the same dwelling as the household members.
- (ii) The salient moves (migration status) made by members of the household.
- (iii) Malaria indicators: use and treatment of mosquito nets.

The respondents for questions in the first half of this section (columns (1) – (7)) were all members of the household below 18 years while the questions in the second half applied to all members of the household. To the extent possible each person was asked directly. If someone was not available or too young to answer then the household head, spouse, or another well-informed member of the household would answer these questions.

Section 4: Education

The objective of this section was to measure the level of education or formal schooling of all household members aged 5 years and above, and to collect educational expenditures associated with each. Information was mainly collected on (i) the literacy status of household members – i.e. member of the household who could read and write; (ii) the educational attainment of each respondent and the type of school attended; and (iii) amount spent on education of household member's during the past 12 months, among others.

Section 5: Health

This section collected information on illness and injuries among household members during the past 30 days, use of health facilities and medical expenses for treating the illnesses or injuries. The respondents for the section were all members of the household, but parents or a knowledgeable adult (preferably female) could answer for young children.

Section 6: Child Nutrition and Health

These questions were asked with a view of obtaining a better picture of the diversity of the child's diet. Only children aged from 0 to 59 months and living with a mother or caretaker in the sampled households are eligible for the questions. The questions were answered by the mothers /caretakers of the children because they are considered more knowledgeable about the children. Height and weight measures were obtained for all children aged 6 to 59 months using anthropometric equipment.

Section 7: Disability

This section gathered information on:

- (i) Self-reported limitations on usual activities due to illness and caring for sick member of the household;
- (ii) Disability as a difficulty to be measured (both adults and children)

The questions applied to all members of the household aged 5 years and above. In some instances, parents or knowledgeable adults (preferably female) would answer for young children.

Section 8: Labour Force Status

This section acted as a screen to determine which respondents should be asked about employment and which should be asked the questions that address labor force participation, unemployment, and job search. It also determined the reason for absence for those people who had a job or business but were not at work the previous week. All household members aged 5 years and older were classified into three broad groupings i.e. employed, unemployed, and not in the labor force.

Employed persons were defined as those who were working at a paid job or business or who were working unpaid at a household business or farm **for at least one hour during the reference week**, or who did not work during the reference week but held a job or had a business from which they were temporarily absent.

Unemployed persons were classified as those individuals who did not work at all during the reference week and who were not absent from a job, but who actively looked for work during the past four weeks and were available to work in the reference week. Persons who were on layoff from a job to which they expected to return and were available to work during the reference week are also classified as unemployed, even if they did not actively look for work. The sum of the employed and the unemployed constituted the **labor force**. (*Persons **not in the labor force** were those who were neither employed nor unemployed. They did not work, they were not absent from work and they did not actively look for work in the past four weeks*).

Section 9: Household and Housing Conditions

Data from this section was aimed at measuring the quality of housing occupied by the household currently. Information was collected on the type of dwelling, occupancy status, the physical characteristics of the dwelling, and access to basic services (including water, electricity and sanitation). A dwelling was defined as a building or a group of buildings in which the household lived. It could be a hut, a group of huts, a single house, a group of houses, an apartment, several one-room apartments, etc.

Section 10: Energy Use

Information obtained in this section aimed at measuring the access and utilization of energy fuels especially for lighting and cooking.

Section 11: Household Incomes

This section gathered information on income transfers i.e. all incomes of household members other than that from paid and/or self employment during the past 12 months.

Section 12: Non-agricultural Enterprises/Activities

This section collected information the presence of non-agricultural household enterprises. It includes information income and employment derived from non-agricultural household enterprises. It identified which household member was responsible for each enterprise in terms of decision making and the allocation of income generated. It also covered the involvement of household enterprises in the credit market. The respondent for each enterprise was a member of the household best informed about the activities of the enterprise.

An enterprise was defined as any undertaking which is engaged in the production and/or distribution of some goods and/or services meant mainly for the purpose of sale whether fully or partly.

Section 13: Financial Services

Information was collected on households' access and use of financial services. These included credit unions, micro financial institutions (MFI) and Savings and Credit Cooperatives (SACCO).

The main purpose of this section was to collect data to estimate the value of household, farm and non-farm enterprise assets. Information on ownership of assets was also collected.

Section 14: Household Assets

This section aimed at collecting data to estimate the value of household, farm and non-farm enterprise assets. It also collected information on ownership of assets.

Section 15: Household Consumption Expenditure

This section covered expenditures of the household with different reference periods depending on the frequency of purchases. It is separated into four parts which include: (a) food, beverages and tobacco; (b) non-durable goods and frequently purchased services; (c) semi-durable and durable goods and services; and (d) non-consumption expenditure.

The major emphasis of the section was mainly on consumption and not monetary expenditures. Although the two are very close, they are not the same. Household consumption expenditures in cash, kind or through barter were recorded for the household only. For bartered items the value of the item paid for (not the value one got in exchange) was recorded. Food, beverages or tobacco served to other members and guests in the household during the reference period were however included. The respondent for this section was the person (household member) who managed the household budget and was the best informed about the household's consumption expenditure.

Section 16: Shocks and Coping Strategies

Shocks were defined as events that happen suddenly. Usually they have a marked beginning and end. While they last for a short time, a few days or weeks, usually their effects are felt for a longer time. It was noted that a shock can be household specific or community wide. Examples of shocks include floods, rebel raids, livestock disease, fire, etc. For example, petty theft of household property was not considered as a shock.

Section 17: Welfare Indicators and Food Security

The Purpose of this section was to collect information on vital needs and living conditions of households during the last 12 months. It provided additional information to assess household welfare. **Food security**

was defined as the availability of food and one's access to it. A household was considered food secure when its occupants did not live in or fear of starvation.

Section 18: Transport Services

Information was collected on access to and use of transport services. These included access to road infrastructure. A **road** was defined as an open way for the passage of vehicles, persons or animals. **Trunk roads** are main roads maintained by the central government and they are normally connecting a district to other districts. These can either be tarmac or murram roads and they are normally six (6) metres and above in width. **District/Feeder roads** are major roads joining Trunk roads and are maintained by the district authorities. **Community roads** on the other hand are roads (excluding footpaths) connecting villages and are normally maintained by the communities themselves.

Table 1 : Organization of the UNPS 2009/10 Household Questionnaire

Section	Level of Observation	Data File	Key Identifiers
Household Identification Particulars	Household	GSEC1.dta	HHID
Household Roster	Individual	GSEC2.dta	PID
General Information on Household Members	Individual	GSEC3.dta	PID
Education	Individual	GSEC4.dta	PID
Health	Individual	GSEC5.dta	PID
Child Nutrition and Health	Individual	GSEC6.dta	PID
Disability	Individual	GSEC7.dta	PID
Labour Force Status	Individual	GSEC8.dta	PID
Housing Conditions, Water and Sanitation	Household	GSEC9.dta	HHID
Energy Use	Household	GSEC10A.dta	HHID
Energy Use cont'd	Fuel Type	GSEC10.dta	HHID h10q13
Other Household Income in Past 12 months	Income Type	GSEC11.dta	HHID h11aq03
Non-Agricultural Household Enterprises/Activities	Enterprise	GSEC12.dta	HHID h12q01a
Financial Services Use	Household	GSEC13.dta	HHID
Household Assets	Asset Type	GSEC14.dta	HHID h14q2
Household Consumption Expenditures – No. of Household Members present	Household	GSEC15A.dta	HHID
Household Consumption Expenditures – Food, Beverages and Tobacco (Last 7 days)	Consumption Item	GSEC15B.dta	HHID h15bq2
Food Fortification	Consumption Item	GSEC15BB.dta	HHID h15bqid
Household Consumption Expenditures – Non-Durable Goods and Frequently Purchased Services (Last 30 days)	Consumption Item	GSEC15C.dta	HHID h15cq2
Household Consumption Expenditures – Semi-durable and Durable Goods and Services (Last 365 days)	Consumption Item	GSEC15D.dta	HHID h15dq2
Non-Consumption Expenditures (Last 365 Days)	Consumption Item	GSEC15E.dta	HHID h15eq2
Shocks and Coping strategies	Shock Type	GSEC16.dta	HHID h16q00
Welfare and Food Security	Household	GSEC17.dta	HHID
Transport Services and Road Infrastructure	Road Type	GSEC18.dta	HHID h18q1
Transport Services and Road Infrastructure cont'd	Household	GSEC18A.dta	HHID
Transport Services and Road Infrastructure cont'd	Activity	GSEC18B.dta	HHID h18q9

2.2: Explanatory notes by section – Agriculture Questionnaire

The purpose of the agricultural module in the household survey was to give a better descriptive picture of Uganda's farm economy, and deeper insight into factors affecting farm incomes. These would include a better understanding of the influence of farmers' resources and marketing opportunities on farm-household income, and some sense of how farmers' situation has changed in the past few years.

The agriculture module was administered in two visits to the selected households. During the first visit, agricultural production data was collected on the first cropping season of 2009 (January – June 2009). The second visit collected agricultural production data on the second cropping season of 2009 (July – December 2009).

The main or first agricultural season normally refers to the growing cycle of temporary crops that are planted and harvested in the first half of the year, occasionally extending up to the end of June. It thus covers the period between January and June. The second agricultural season is generally the period between July and December. It should be noted that seasons are directly related to rains and only indirectly related to the growing cycle of crops. The first rains are generally longer than the second rains. However, it is also noted that some areas in Uganda have only one significant agricultural season.

Section 1A: Household Identification Particulars

Information in this section was distributed by the Headquarters staff to the field teams before starting data collection. Names and codes pertaining to the selected Enumeration Areas (EAs) were provided by UBOS to the team leaders prior to fieldwork. An EA generally does not have its own name but is known by the name of the LC1 that is associated with it.

Section 1B: Staff details and survey time

The Supervisors, interviewers and data entry operators were all required to record their particulars in this section. Time taken to conduct interviews was also recorded.

Section 2: Current land Holdings and land that the household has access through use rights

The purpose of this section was to have a complete list of all the parcels owned and/or operated as well as rented by the household during the first season of 2009 and the second season of 2009. This section captures information in two parts; Section 2 part A captures information concerning current land holdings and section 2 part B captures information pertaining to land that a household has access to through use rights. The questions were administered to households who had been involved in crop farming during the last completed and the current cropping seasons. Information was collected on agricultural land that

these households had access during the reference period. Issues of land tenure status and land user rights were also investigated.

Section 3A & 3B: Agricultural and labour inputs

This section collected information on non-labor and labor input applications at the parcel-plot-level during the first cropping season (January-June 2009) and second cropping season (July –December 2009) in part A and B, respectively.

Section 4A & 4B: Crops grown and type of seeds used

The purpose of this section was to collect information on crop cover of parcels farmed by the household. Data was collected on crops planted by the household during the first cropping season (January-June 2009) and second cropping season (July –December 2009) on each plot on each parcel accessed by the household through ownership or user rights, in part A and B, respectively.

Section 5A & 5B: Quantification of Agricultural Production

Information on agricultural production is collected at the parcel-plot-crop-level separately for the first cropping season (January-June 2009) and second cropping season (July –December 2009) in part A and B, respectively. This section also collects data on how the household used the harvested produce.

Sections 6A, 6B & 6C: Livestock ownership

The data on the ownership of (i) cattle and pack animals, (ii) small animals, and (iii) poultry and other animals are solicited in sections 6A, 6B, and 6C, respectively. Each section collects information on dynamics of household livestock ownership at animal-type level over a given reference period, earnings from animal sales, and expenditures on animal purchases. If the household cared for animals that belonged to others, interviewers were instructed to record only ownership, sales and purchases of animals the household was entitled to keep, for instance the baby goats or sheep that the household keeps in return for caring the flock.

Section 7: Livestock expenditure and income

The purpose of this section was to estimate cash earnings from livestock products i.e. the difference between the revenues a household earns from selling animal by-products and the expenditures necessary to raise the animals. Earnings from the sale of animal products produced from other purchased animal products were not included, for example, revenues from the sale of butter/cheese produced from milk bought in the market.

Section 8: Livestock Products

This section collected information on the production and sales of livestock by-products. The reference period was generally last 12 months unless otherwise.

Section 9: Fishing

This section collected information on fishing activities conducted by the household. The questions sought to establish the type fishing practice used and quantity of fish caught.

Section 10: Extension Services

The section collected information on agricultural technology and extension services. It covered access to extension services and access to and demand for agricultural technology. Extension workers were defined as individuals employed by the government or non-governmental organizations who work as an agricultural development agents for contacting and demonstrating improved farming methods to farmers. They are responsible for organizing, disseminating, guiding and introducing technical methods in agricultural production directly to farmers, and for facilitating farmers coming into contact with cultivation methods to promote agricultural production.

Table 2 : Organization of the UNPS 2009/10 Agriculture Questionnaire

Section	Level of Observation	Data File	Key Identifiers
Household Identification Particulars	Household	AGSEC1.dta	HHID
Current Land Holdings - 1 st /2 nd Visit	Parcel	AGSEC2A.dta	HHID a2aq2
Land That the Household Has Access Through Use Rights - 1 st /2 nd Visit	Parcel	AGSEC2B.dta	HHID a2bq2
Agriculture and Labour Inputs – 1 st Visit	Parcel-Plot	AGSEC3A.dta	HHID a3aq1 a3aq3
Crops Grown and Types of Seeds Used – 1 st Visit	Parcel-Plot-Crop	AGSEC4A.dta	HHID a4aq2 a4aq4 a4aq6
Quantification of Production – 1 st Visit	Parcel-Plot-Crop	AGSEC5A.dta	HHID a5aq1 a5aq3 a5aq5
Agriculture and Labour Inputs – 2 nd Visit	Parcel-Plot	AGSEC3B.dta	HHID a3bq1 a3bq3
Crops Grown and Types of Seed Used – 2 nd Visit	Parcel-Plot-Crop	AGSEC4B.dta	HHID a4bq2 a4bq4 a4bq6
Quantification of Production – 2 nd Visit	Parcel-Plot-Crop	AGSEC5B.dta	HHID a5bq1 a5bq3 a5bq5
Livestock Ownership – Cattle and Pack Animals	Livestock Type	AGSEC6A.dta	HHID a6aq3
Livestock Ownership – Small Animals	Livestock Type	AGSEC6B.dta	HHID a6bq3
Livestock Ownership – Poultry and Others	Livestock Type	AGSEC6C.dta	HHID a6cq3
Livestock Expenditure	Expenditure Type	AGSEC7.dta	HHID a7q2
Livestock Products and Income	Livestock Product	AGSEC8.dta	HHID a8q2
Fishing	Household	AGSEC9A.dta	HHID
Fishing – Methods of Utilization	Purpose	AGSEC9B.dta	HHID a9q6purp
Fishing – Ownership of Fishing Equipments	Fishing Equipment	AGSEC9C.dta	HHID a9q10a
Fishing – Operational Cost	Fishing Expenditure Item	AGSEC9D.dta	HHID a9q11a
Fishing – Form of Sale	Household	AGSEC9E.dta	HHID
Extension Services	Extension Source	AGSEC10.dta	HHID a10q2

2.3: Explanatory notes by section – Community Questionnaire

The community survey aimed at collecting information relating to communities residing in the sampled EAs. The administrative unit for collection of community data was mainly the LC1, although there were specific questions for the Sub-county Chief. The community survey information was collected by interviewing key informants within the institutions of interest. These included community members and heads of selected facilities.

Section 1: Community Identification Particulars

Most of the information in this section was obtained from headquarters by field teams before starting data collection. A provision was made to record details for each of the subsequent 4 sectors on which data was collected. These included names of respondents and responses status for each sector.

Section 2: Availability of services within the community

The purpose of this section was to obtain general information on the social infrastructure nearest to the community. Information was collected from community leaders. The social facilities on which data was collected included schools/other education facilities, banks, markets, agricultural and fisheries services, police and army facilities, various types of health facilities, water and sanitation facilities as well as works and transport services .

Section 3: Education (Primary)

Information for this section was provided by a knowledgeable school official preferably the headmaster or someone nominated by him/her. Data was collected on both the most popular and the nearest primary schools. These schools on which data was collected were not necessarily located within the LC1 covered.

Section 4: Health services

In this section, information was collected on the most commonly used public and private health facilities. The respondent for this section was an authorized or knowledgeable health official at the facility preferably the head of the facility. The health facility targeted would be a place that had qualified doctors/nurses/medical attendants for treating patients including dressing and emergency attention facilities and would in addition be selling medicines to patients. Individual doctors, practitioners, etc, doing only consultation, with very limited supply of medicines were excluded. However, Doctors with moderate treatment and medical attention facilities were included.

Section 5: Works and Transport

The respondent for this section was the sub-country chief. Information was mainly collected on the availability, use and maintenance of works and transport infrastructure.

Table 3 : Organization of the UNPS 2009/10 Community Questionnaire

Section	Level of Observation	Data File	Key Identifiers
Identification Particulars	EA	CSECTION1.dta	comcod
Service Availability in LC1	Service type	CSECTION2A.dta	comcod c2asn
Client satisfaction with health facilities	EA	CSECTION2B.dta	comcod
Water and Sanitation	EA	CSECTION2C.dta	comcod
Primary school identification and management	EA	CSECTION3.dta	comcod
Availability of Facilities at School	Facility type	CSECTION3A.dta	comcod c3asn
Condition of toilets	EA	CSECTION3B.dta	comcod
Water facilities at the School	Water facility type	CSECTION3C.dta	comcod c3csn
Payment for Services by Parents/Guardians	Item	CSECTION3D.dta	comcod c3dsn
Academic Performance of pupils in PLE	Year	CSECTION3E.dta	comcod c3e
Incidence of leaving school prematurely	Year	CSECTION3F.dta	comcod c3f
School meetings	Type of meeting	CSECTION3G.dta	comcod c3gsn
Staffing at the School	Staffing position	CSECTION3H.dta	comcod c3hsn
Supervision/Monitoring of School during last 12 months	Supervisor/monitor	CSECTION3I.dta	comcod c3isn
Problems/constraints faced by School	Problem type	CSECTION3J.dta	comcod c3jsn
Learner attendance, Teacher presence and qualifications and other classroom elements	Class	CSECTION3K.dta	comcod c3kq51 visit
Health facility identification and management	EA	CSECTION4.dta	comcod
Work at Night	EA	CSECTION4A.dta	comcod
Availability of equipments/ services at the facility	EA	CSECTION4B.dta	comcod
Services offered by Health facility	Service type	CSECTION4C.dta	comcod c4csn
Common diseases reported at Health facility	EA	CSECTION4D.dta	comcod
Common stock-outs reported by Health Facility	Drug supplies	CSECTION4E.dta	comcod c4esn

Table 3 (Cont'd)

Items bought by patients visiting the Health facility	EA	CSECTION4F.dta	comcod
Deliveries at the facility	EA	CSECTION4G.dta	comcod
Health facility Identification for HMIS	EA	CSECTION4H1.dta	comcod
Validation of HMIS	Data element, period	CSECTION4H2.dta	comcod
Epidemic reporting	EA	CSECTION4I.dta	comcod
General operations	EA	CSECTION4J.dta	comcod
Sanitary Facilities Available at the Health Facility	EA	CSECTION4K1.dta	comcod
Access to Water at the Health facility	Water facility type	CSECTION4K2.dta	comcod c3csn
Factors Limiting provision of Health Services	Limiting factor	CSECTION4L.dta	comcod c4lsn
Supervision/Monitoring of Health Facility	Supervisor/monitor	CSECTION4M.dta	comcod c4msn
Village Health Teams	EA	CSECTION4N.dta	comcod
Staffing at the Health Facility	Positions	CSECTION4O1.dta	comcod c4osn
List of Medical Staff working at the Facility	Medical staff visit	CSECTION4O2.dta	comcod c40sn2 visit
Works and Transport	EA	CSECTION5.dta	comcod
Infrastructure availability and condition	Item type	CSECTION5A.dta	comcod c5asn
Maintenance and Repair of Infrastructure	Item type	CSECTION5B.dta	comcod c5bsn
Funding for Maintenance of Roads/Bridges/Culverts	Item type	CSECTION5C.dta	comcod c5csn
Constraints faced in the maintenance and repair of roads	Item type	CSECTION5D.dta	comcod c5dsn

2.4: Explanatory notes by section – Woman Questionnaire

The intention of the Woman module in the household survey was to gather information relating to knowledge and use of contraceptives among women as well as their birth history. This questionnaire was administered to all women aged 15-49 years in the households.

Section 2A: Contraception

Information on contraceptives in the survey was collected by asking females in the households within the reproductive age of 15-49 years to name the different ways or methods that one would use to avoid or delay getting pregnant. The interviewers would then describe the methods mentioned to the respondent in case she failed to mention any spontaneously. They would go further to ask the respondent if they have ever used each of the methods and the ones that they are currently using with their partner.

Section 2B: Birth History

The purpose of this section was to obtain information on the birth history of eligible women in the household (15-49 years). Information was mainly collected on the number of children ever born whether dead or alive as well as birth information on last child born in the last five years (whether living or dead).

Table 4 : Organization of the UNPS 2009/10 Woman's Questionnaire

Section	Level of Observation	Data File	Key Identifiers	Remarks
Contraception and Birth Related Issues: Contraception	Contraception Type	GSEC2A	PID ws2q2	Information obtained for women aged 15-49 years on their knowledge and use of contraceptives as well as birth history
Contraception and Birth Related Issues: Birth History	Individual	GSEC2B	PID	

3 Other related instructions/codes

3.1: Area Measurement using Global Positioning System (GPS)

The GPS was mainly used in measuring parcels owned and/or operated by the selected households located within the EA and crop plot area for the respective cropping season of 2009. The GARMIN 12 hand-held Global Positioning System (GPS) equipment was used. The GPS equipment is in principle a high precision digital watch combined with a signal receiver. The field supervisors were responsible for ensuring availability of fully charged batteries for the GPS equipment and also ensuring that they were handled with great care and stored in a safe place when not in use. Details on GPS equipment were well documented and rigorous training about use of GPS was given to the interviewers before actual data collection.

3.2: Other Codes

There were a number of sections for which the respective codes could not fit within the cell/page where the question was located. For these questions, a separate code sheet was provided in the instructions manual. These code lists included:

- Reason for staying in the household for less than 12 months
- Highest level of education attained
- Current schooling status
- Ethnicity
- International Standard Classification of Occupations (ISCO)
- International Standard Industrial Classification (ISIC)
- Units of Quantity
- Crop Codes
- Condition and state of crops harvested

4 Field Work Organization

Prior to starting fieldwork enumerators and supervisors were trained for a period of approximately three weeks with many practical sessions to ensure competency and accuracy during household identification and data collection. The planned structure of the implementation of the UNPS was to have 9 mobile field teams, each of which was comprised of a driver, a supervisor, three enumerators, and a data entry operator. Each mobile team required a vehicle, a data entry laptop, GPS units for the enumerators, and anthropometric equipment (height and weight scales). The data entry was done in the field, which meant that the questionnaires were keyed in and checked for errors prior to the departure from each EA. Given internet access, the supervisors sent the data electronically from the field at the conclusion of data entry for each EA. The computer-assisted field-based data entry was an innovation with respect to the UNHS design, where the data entry is typically conducted in a centralized location.

The teams went on a two to three week-long trip each month. At the end of each trip, the teams reported back to the Headquarters. The main field work, which lasted from September 2009 to August 2010, was comprised of two six-month phases. All households were visited once in each phase with a portion of split-off individuals identified in phase 1 being visited only once across the 12-month period with the visit taking place in phase 2. The latter was mostly due to long-distance tracking cases where the survey teams simply did not have adequate time to track the households as part of phase 1 operations. At the end of 12 months, the UNPS 2009/10 field operations were also extended for two months, specifically for finalizing split-off tracking that was not accommodated as part of the main field work due to time and logistical constraints. The two additional months of field work took place in parallel with the UNPS 2010/11 field operation, and was largely implemented by an extra team specifically devoted to tracking,

The two-visit field work structure was designed to accommodate the difficulties associated with solicitation of information on agriculture in the presence of multiple agricultural seasons. As is well-known, Uganda has two agricultural seasons, the first running from February to July/August and the second from August/September to December. To collect accurate information for each of the two agricultural seasons and minimize recall associated with agricultural decisions that the survey seeks information on, the survey households with the exception of a portion of split-off cases noted above, were visited twice in the course of 12 months.

The dual visits also enabled splitting the questionnaire material and worked towards reducing respondent fatigue. In each cluster, approximately half of the households were randomly selected for the entire household questionnaire to be administered in visit 1. As far as the household questionnaire is concerned, these households only received a roster update in visit 2. The rest of the sample received only the household roster in visit 1 (along with the half of the agriculture questionnaire, assuming that they

were agricultural households), and the rest of the household questionnaire along with the household roster update in visit 2. This arrangement attempted to ensure an even distribution of households that reported information on household consumption in each month of the main field work. The information solicited from each household in visit 1 was fed forward for visit 2. EAs in each region were randomly assigned to each team and randomly split into six subsets of six EAs each, one for each of the first six months of fieldwork, after which households were re-visited after six months.

4.0: Tracking of Households

Tracking considers the mobility of the target population, the success with which those who move are found and interviewed, and the number of refusals. In wave 1 of the Uganda National Panel Survey 2009/10 tracking was done both at household- and individual-level. It aimed at tracking all the 3123 panel households and among these approximately 20% (2 households from each EA) was considered for individual tracking also known as split-offs tracking.

4.1: Tracking of Households

The UNPS tracked all original households including those that shifted from their original location in 2005/06 to another location either within the same EA or outside it. These are referred to as shifted households. An original household is the household interviewed in the 2005/06 baseline sample. As noted above, from the UNHS 2005/06, a total of 3,123 households were sampled as UNPS households..

Once the location of the original household was found, then the household would be interviewed. Failure to interview a household could be due to shifting to an unknown location, refusal or disintegration among others. Disintegrating means that each of the original household members had gone separate ways and none of them remained at the original location. If a household disintegrated, its members would only be tracked if it fell under the 20 percent sample for split-offs tracking otherwise none of the members of a disintegrated household would be tracked. If the tracked original household had new members as of the 2009/10 interview, the data on these individuals were solicited as well.

4.2: Tracking of Split-offs

As part of the management of individual/split-off tracking the UNPS chose to track a 20% sample households found in each of the 322 Enumeration Areas. The intention is to calibrate the size and composition of the sample of traceable split-offs (*currently referred to as tracking targets*) that will be actually tracked, so that it roughly compensates the losses due to attrition.

A random sub-sample of two households from each EA was drawn from the already sampled panel households. These two households were referred to as split-offs tracking targets. It was then identified if any of the household members in 2005/06 of these two households had left the household. These movers were referred to as split offs (tracking targets).

Once a split-off was identified, then it was tracked fully by first gathering all the contact information about this split-off/mover as well information on their new location from the original household members and any other knowledgeable person. This information was filled in a questionnaire called the individual tracking form. Based on the details filled in this questionnaire, the mover was contacted if contacts were available, traced based on the location details given by the original household or the contacted mover and then interviewed. The interviewed split-offs as well as the members of the new household that they had formed or had joined in by the time of the UNPS 2009/10 interview then became part of the UNPS sample and will be interviewed in every wave of the UNPS, even if they shift to alternative locations in subsequent waves.

It should be noted that only individuals that were related to the household head such as spouse, biological children, parents of the head or spouse, etc (codes 1-7 of Section 2 Question 4 in the household questionnaire) were tracked. Servants, other relatives and non relatives (codes 8-96) were not tracked.

The first 10 digits of the household identifier (HHID) attached to the new household in which the split-off was located in 2009/10 was the same as the household identifier for the household that the split-off was a part of in 2005/06 (i.e. parent household).

The last 2 digits of the 2009/10 split-off household identifier correspond to the 2005/06 roster line number for the split-off individual. In the event that multiple split-offs from the same parent household were found to be co-residing in 2009/10, the last 2-digits of the 2009/10 split-off household identifier correspond to the lowest UNHS 2005/06 roster line number among the split-offs.

5 Linking UNHS 2005/06 & UNPS 2009/10

As part of the dissemination package, the data from the UNHS 2005/06 sample covering 3,123 households and 322 EAs that were selected for the purposes of the UNPS 2009/10 are provided.

The UNHS 2005/06 portion of the dissemination package includes the (i) Household, (ii) Agriculture, and (iii) Community data as well as the descriptive reports, questionnaires, and manuals. At the household-level the variable **tracking sample** as part of GSEC1.dta of the UNHS 2005/06 package identifies the **643** (out of 3,123) UNHS 2005/06 households were selected for split-off tracking prior to the start of the

UNPS 2009/10 field work. The UNHS 2005/06 data that are provided could be linked with the UNPS 2009/10 data at the household-, individual- and community-levels through the unique household identifier (**HHID**), the unique individual identifier (**PID**), and the unique community identifier (**comm**), respectively.

Given the attrition at the household- and individual-level, and the addition of new households and individuals to the UNPS sample in accordance with the protocols described above, the household- and individual-level matches across the UNHS 2005/06 subsample and the UNPS 2009/10 will not be perfect. The variable **hh_status** as part of GSEC1.dta of the UNPS 2009/10 package identifies (i) original households that were interviewed at the location of the 2005/06 interview, (ii) original households that were interviewed at an alternative location with respect to the location of the 2005/06 interview, and (iii) split-off households interviewed in 2009/10.

As an ancillary data file, the UNPS 2009/10 data package also includes a complete record of all individuals interviewed in 2005/06 and/or 2009/10, with a variable, **ind_status**, that maps individuals into the following categories:

- (i) Not interviewed in 2009/10 and household selected for split-off tracking,
- (ii) Household missing in its entirety in 2009/10 and household selected for split-off tracking,
- (iii) Not interviewed in 2009/10 and household not selected for split-off tracking,
- (iv) Household missing in its entirety in 2009/10 and household not selected for split-off tracking,
- (v) New sample member interviewed only in 2009/10,
- (vi) Interviewed in 2009/10 and household selected for split-off tracking,
- (vii) Interviewed in 2009/10 and household not selected for split-off tracking,
- (viii) Died between 2005/06 and 2009/10, and
- (ix) Not tracked in 2009/10 since the individual was a servant or non-relative.

Lastly, for split-off households specifically, **parent_HHID** as part of GSEC1.dta of the UNPS 2009/10 package is the unique identifier for the original household that the split-off household is associated with.

6 Calculation of UNPS 2009/10 Panel Weights

This section presents a general description of the steps involved in the construction of panel weights for the UNPS 2009/10. The UNPS has two broad analytic goals. The first is to track the same sample of people from 2005 to 2009 to see how their lives have changed. The second is to provide a cross-sectional snapshot of the Ugandan household population. The first goal is met by revisiting the same cases in each wave. The second goal, however, necessitates the incorporation new households and individuals in each wave to account for the changing population.

The UNPS addresses this conflict by allowing for the inclusion of a sub-sample of split-off households in the second wave. Household members who were interviewed in 2005 survey and were selected for the 2009 round of data collection, but who are not living in the household at the time of the 2009 survey visit are tracked to their new location. If found, they are interviewed, along with the members of their new household. This methodology allows new units to enter into the sample. Split off households then remain part of the core sample during subsequent rounds. However, the approach does not completely offset the loss of representativeness, as those new population members who do not live in households with those eligible for selection in 2005 have no chance of selection. A new round of sample selection such as a refreshment sample would be required to include the members of such households.

The methodology described here builds upon published documentation from established panel surveys, such as the Panel Study of Income Dynamics [PSID], conducted since 1968 by the Institute for Social Research at the University of Michigan; and the British Household Panel Survey [BHPS], whose first 13 waves were conducted between 1991 and 2003 by Institute for Social and Economic Research at the University of Essex. Both the PSID and the BHPS are nationally-representative panel surveys in the USA and the UK respectively.

The weights are developed in eight steps:

- 1) Begin with the “base weights” or those calculated during the previous round of the survey;
- 2) incorporate the probability of selection from the UNHS into the UNPS;
- 3) incorporate the probability of selection into tracking;
- 4) derive fair-share weights for composition changes;
- 5) pool the weights in (i), (ii) and (iii) together;
- 6) derive attrition adjusted weights for all individuals, including split-off¹ households, then aggregate these weights to the household level;
- 7) trim these weights;
- 8) post-stratify the pooled weights to known population totals.

Each of these steps is discussed in detail below.

¹ For the purposes of this note, ‘parent’ refers to the household found at the same location as the previous round of data collection, and ‘split-off’ refers to new households entering the sample through an individual originally resident in a parent household during a previous round.

6.1: Base Weights from 2005 Sample

The panel weight calculations are based on the 2005 household weights from a subsample of the UNHS. The 2005 weights are based on the inverse probability of selection and an EA level non-response correction. These probability weights form the first component of the 2009 calculations.

$$W_1 = W_{2005}$$

6.2: Probability of Selection into UNPS

The 2009 UNPS is a subsample of the 2005 UNHS. To select the sample for the UNPS, the UNHS sample was divided into five strata (Kampala, Central, Eastern, Northern and Western). Within each stratum, EAs were selected using simple random sampling, but the probabilities of

Table 5: EA Probability of Selection	
Strata	Probability of selection (a_1)
Kampala	1.000
Central	0.393
Eastern	0.364
Northern	0.433
Western	0.400

selection varied between strata. In Kampala, all UNHS EAs were selected to ensure sufficient sample size in that stratum. In the other four strata, the probability of selection ranged between 36 and 43 percent. If the probability of selection is:

$$a_1 = \frac{n_h}{N_h}$$

Then the adjusted weights would be:

$$W_2 = W_1 * (a_1)^{-1}$$

6.3: Probability of Selection into Tracking

All households from the randomly selected subsample of the 2005 household survey are included in the 2009 UNPS with certainty. Additionally, two households per enumeration area [EA] are selected as eligible to have split off members tracked and interviewed. In general there are ten households per EA, though in some cases there were refusals during the 2005 UNHS round, and therefore there are fewer than ten in the initial pool. The selection of eligible tracking households is done prior to fieldwork at UBOS headquarters, and therefore some selected households are not found.

The sub-sampling at this stage somewhat complicates the weighting calculations. Ideally, the 'parent' household would receive a weight of 1, because it is selected into the sample with certainty (it would be interviewed regardless if it is selected into tracking). The 'split-off' households would include the probability of selection into tracking, which would, for example in the case of a 10 household EA, be $\left(\frac{2}{10}\right)^{-1}$ or 5.

The uncertainty arises, however, from the fact that the difference between a parent and a split-off is an arbitrary distinction. In general, the field manual for enumerators states that the parent household is the dwelling in the same location as the 2005 household, provided that at least one of the original members is still present in that dwelling. This would mean if one child remained at the original location and the other five family members moved to the other side of the village, the child would be the parent household and the other five members the split-off. If only the child re-located, the five original members would be the parent and the child the split-off. The designation becomes even more arbitrary if there are no members present in the original dwelling. Field supervisors could choose to designation the household with the largest number of original members, or where the household head resides, or any other criteria they choose.

Since the designation is arbitrary, there can be no mathematical difference in the probability of selection between the parent and the split off household. The probabilities of selection are therefore pooled and averaged over all households originating from a single original household. Therefore the weight to compensate for selection into tracking is as follows:

- Not selected into tracking, parent household is followed with certainty. Probability of selection is 1.
- Selected into tracking but did not split, parent household is followed with certainty. Probability of selection is 1.
- Selected into tracking and split, probabilities of selection are pooled and averaged. Examples:
 - Household from an EA with 10 UNHS households is selected into tracking and splits into two households. The probability of selection for the two households (both the 'parent and 'split-off') is $^{(5+1)}_2=3$.
 - Household from an EA with 10 UNHS households is selected into tracking and splits into three households. The probability of selection for the two households (both the 'parent and two 'split-off') is $^{(5+5+1)}_3=3.67$.

Therefore at this stage, we would add another component to our weight calculations:

$$a_2 = \begin{cases} 1 & \text{if not selected for tracking, or selected but no split offs} \\ \frac{(m+1)}{1+m(\frac{2}{q})} & \text{otherwise, where } m \text{ is the number of splits and } q \text{ is the 2005 EA size} \end{cases}$$

Then the adjusted weights would be:

$$W_3 = W_2 * (a_2)^{-1}$$

The sum of the weights at this point is equal to the size of the 2005 population eligible for selection in the initial round of data collection in 2005.

6.4: *Fair Share Correction*

As discuss above, the follow-up rules for the UNPS allow for the incorporation of people who now live with original sample members. For example a young adult living with his parents in 2005, may be 2009 have formed a new household, having gotten married and had a child. The wife and infant will be incorporated into the survey and thus require a probability of selection. Such corrections are routinely used to distribute weight to new sample members in panel surveys. See Rendtel and Harms (2009) for a discussion of several different methods of weight correction.

Because split-off individuals are tracked and interviewed in their new households, there are multiple ways that a household can become part of the survey.

- Either by being selected initially for the UNHS, and during the subsequent rounds of sub-selection.
- By receiving a member that came from a household that was selected for the UNHS and during the subsequent rounds of sub-selection.

In an ideal world, we would be able to know the probability of selection that each new member brought into the household, and adjust the household weight accordingly. This is necessary since households receiving members have higher probabilities of selection (and therefore lower weights) because the household could have been selected in multiple ways. Since we cannot know the probabilities of every member, we must make simplifying assumptions. The first simplifying assumption is that the arriving members arrived together from one other household. This would be the case if a man and woman get married and set up a new household, or in the case of an older relative moving in with adult children. In certain cases, however, arriving members come from more than one household. Assuming only two source households underestimates slightly the probability of selection (and therefore over-estimates the weights). Incidence of these cases is believed to be relatively rare, and any resulting bias should be

negligible. The second simplifying assumption we make is that the arriving members have the same probability of selection, on average, as those members that are already there. This would not be true on a case-by-case basis but would be true in the aggregate. With these simplifying assumptions, we add a factor of $\frac{1}{2}$ for all households, 'split' or 'parent' that have new members arriving from other households. This takes into account the fact that they could have been selected in two ways, and assumes the probability of selection is equal.²

$$a_3 = \begin{cases} 1 & \text{otherwise} \\ \frac{1}{2} & \text{if new members} \end{cases}$$

Then the adjusted weights would be:

$$W_4 = W_3 * a_3$$

A limitation of the panel methodology is that the represented population is not identical to the 2009 Uganda household population, as it does not include immigrants in new households. Inclusion of these groups would necessitate refreshing the sample with new households. However, the represented population is close enough to the 2009 Ugandan population to permit the desired cross-sectional estimates.

6.5: Pooling

At this point, the first four steps would be the complete calculations to calculate the panel weights in the absence of attrition.

6.6: Attrition Correction Factor

All household panel surveys must tackle the problem of attrition, sample members selected for follow-up interviews which cannot be located and/or interviewed. The methodology used to adjust weights for attrition in the UNPS follows Rosenbaum & Rubin (1984). We use predicted response probabilities from a logistic regression model based on the covariates to form the weighting classes or cells. This approach has also been adopted in the PSID; see for example, Gouskova (2008).

² New births and arriving children under age 4 do not count as 'new members' in this case because they could not have been selected in 2005.

The attrition correction in the case of the UNPS needs to take into account two distinct sources of attrition: entire households that are not found and split-off individuals that are selected for tracking but not found. The two potential options for the calculations are (1) to treat the split-off households as household heads and do the calculations at the level of the household, or (2) to treat the households that are not found as individuals and perform the calculations at the individual level. The first option is problematic as the characteristics of household heads are dissimilar to the characteristics of split-offs (see table 6). Therefore in the UNPS, the second methodology was employed.

Table 6: Summary Statistics on Heads of Missing Households vs. Missing Individuals				
Characteristic	Total Attrition Sample	Household Attrition Sample	Individual Tracking Attrition Sample	Frequency in Overall Sample
Male (%)	58.2	70.2	48.7	50.5
Age (Years)	28.6	37.9	21.2	20.9

In the UNPS, 489 out of an initial 3,123 households were not found between the 2005 UNHS and the 2009 UNPS, for a household attrition rate of 15.7 percent. Of the 18,410 individuals living a household selected for tracking, 16,956 were found to be living in their original location, and 1,454 has moved to a new household. In addition to the 1,454 individuals tracked to split off locations, an additional 375 were tracked but not found. The individual tracking attrition rate is 20.5 percent.

To obtain the attrition adjustment factor the probability that a sample household was successfully re-interviewed in the second round of surveys is modeled with the linear logistic model at the level of the individual. We create a binary response variable by coding the response disposition for eligible households that do not respond in the second round as 0, and households that do respond as 1³.

We fit a logistic response propensity model, using 2005 UNHS household and individual characteristics measured in the first wave as covariates. Included covariates are:

- gender
- age
- marital status
- presence of father in household
- presence of mother in household
- years of education

³ Note that only household members who have died are excluded from the attrition calculations. In some rare cases, there are eligible household members who were selected for tracking but for whom the field teams did not actually search. Possible reasons could include international migration or lack of time on the part of the field teams.

- current school attendance
- labor force participation
- household consumption
- household size
- residence in agricultural (crop) household
- residence in a livestock household
- residence in household owning enterprise
- residence in household receiving transfer income
- residence in dwelling owned by household member
- residence in dwelling with improved roof
- residence in household with at least one member owning mobile phone
- residence in household with savings
- rural / urban status
- district of residence
- selection into split off sample

In some cases, values of unit level variables were missing from the 2005 household dataset. These values were imputed using multivariate regression and logistic regression techniques. Imputations are done using the 'impute' command in Stata at the level of the UNPS strata. In all cases, less than one percent of the variables required imputation to replace missing values.

The estimated logistic model is used to obtain a predicted probability of response for each household member in the 2009 survey. These response probabilities were then aggregated to the household level (by calculating the mean), then using the household-level predicted response probabilities as the ranking variable, all households are ranked into 10 equal groups (deciles). An attrition adjustment factor was then defined as the reciprocal of the empirical response rate for the household-level propensity score decile.

Then the adjusted weights would be:

$$W_5 = W_4 * ac$$

6.7: *Trimming*

Complex weight calculations have the potential to produce outlier weights. These weights are abnormally high or low, and increase the standard errors of estimates. A common practice is therefore to 'trim' the weights at this stage to eliminate the outlier weights (see Little et al, 1997). Trimming introduces a small amount of bias into the estimates, but allows estimates to be much more efficient. Common values for

trimming range between one and five percent, and the UNPS weights are trimmed at the two percent level on both the top and bottom of the distribution.

6.8: *Post-stratification*

To reduce the overall standard errors, and weight the population totals up to the known population figures, a post-stratification correction is applied. Based on the projected number of households in Kampala and in the urban and rural segments of the four main regions (central, eastern, northern and western), adjustment factors are calculated. This correction also reduces overall standard errors (see Little et al, 1997).

The final weight calculations are the product of the (i) base weight, (ii) the inverse probability of selection into the UNPS, (iii) the inverse probability of selection into tracking, pooled and averaged for original households that split, (iv) a fair-share correction for new members, (v) the attrition correction, and (vi) the post-stratification adjustment. Prior to step (vi), weights are trimmed at the two percent level.

For cross-sectional estimates of population dynamics based only on the UNPS 2009/10 data, the data users must use the variable **wgt09** as part of GSEC1.dta of the UNPS 2009/10 package. This variable includes sampling weights for original as well as split-off households and is generated as a result of the procedures detailed above. As noted above, the UNPS strata of representativeness include (i) Kampala City, (ii) Other Urban Areas, (iii) Central Rural, (iv) Eastern Rural, (v) Western Rural, and (vi) Northern Rural. The variable **stratum** as part of the GSEC1.dta of the UNPS 2009/10 data package captures these.

The variable **wgt09wosplits** as part of GSEC1.dta of the UNPS 2009/10 package was computed only for the original households that were interviewed both in 2005/06 and 2009/10. **wgt09wosplits** is the multiplication of *UNHS 2005/06 sampling weight* and *the inverse of the proportion of the original households that were captured in a given EA*. **wgt09wosplits** is provided for data users interested in conducting household-level analyses of changes between 2005/06 and 2009/10.

7 References

- Gouskova, E., Heeringa, S. (2008), The 2005 PSID Transition to Adulthood Supplement (TA) Weights, PSID Technical Report. ISR, University of Michigan, Ann Arbor MI USA.
- Huang (1984) Obtaining Cross-Sectional Estimates From a Longitudinal Survey: Experiences of the Income Survey Development Program”, in Proceedings of the Section on Survey Research Methods, American Statistical Association.
- Lap-Ming Wun, et. al. (2005) Evaluation of Alternative Propensity Models for Adjusting Weights To Compensate for Dwelling Unit Nonresponse in the Medical Expenditure Panel Survey (MEPS). Journal of the American Statistical Association, 3689-3694.
- Little, R.J.A., S. Lewitzky, S. Heeringa, J. Lepkowski and R.C. Kessler. (1997) "Assessment of Weighting Methodology for the National Comorbidity Survey." American Journal of Epidemiology. 145(5).
- Lynn, Peter (Editor) (2006) Quality Profile: British Household Panel Survey: Waves 1 to 13: 1991-2003. Institute for Social and Economic Research, University of Essex.
- Rendtel, Ulrich and Harms, Torsten. (2009) “Weighting and Calibration for Household Panels.” In *Methodology of Longitudinal Surveys*, ed. P. Lynn. New York: John Wiley & Sons.
- Rosenbaum, P.R., and Rubin, D.B., (1984). “Reducing Bias in Observational Studies Using Subclassification on Propensity Score,” Journal of the American Statistical Association, 79, 516-524.

Appendix: Confidential Information, Geospatial Variables

The Uganda National Panel Survey (UNPS) collects confidential information on respondents. The confidential variables include (i) names of the respondents to the household and community questionnaires, (ii) village names, (iii) descriptions of household dwelling and agricultural parcel locations, (iv) phone numbers of household members and their reference contacts, (v) GPS-based household and agricultural parcel locations, (vi) names of field staff. To maintain the confidentiality of our respondents, certain parts of the UNPS database have not been made publicly available.

To enhance the use of UNPS data, a set of geospatial variables has been generated using the georeferenced household locations in conjunction with various geospatial databases that were available to the survey team. These include simple measures of distance, climatology, soil and terrain and other environmental factors. The variables are intended to provide some understanding of how geophysical characteristics vary across households and between communities.

All geospatial variables have been produced using the unmodified GPS data. Most of the underlying datasets are static (with exception of time-series), so the values should be largely unchanged relative to year 1, for non-mover households. Note that there may be some variation due to GPS data entry error, differences in data collection procedure, and technical limitations of the device. Geospatial variables are provided in the file *UGA_HouseholdGeovariables_Y1*.

UGA_HouseholdGeovariables_Y1

The household-level file, *UGA_HouseholdGeovariables_Y1*, contains a range of variables measuring (on the basis of the household dwelling) distance to other features, climatology, landscape typology, soil and terrain, and growing season parameters. The observations are uniquely identified by **HHID**.

This file also contains modified GPS coordinates, which enable users to generate their own spatial variables while preserving the confidentiality of sample household and communities. Following the method developed for the Measure DHS program, the coordinate modification strategy relies on random offset of cluster center-point coordinates (or average of household GPS locations by EA in the UNPS-Panel) within a specified range determined by an urban/rural classification. For urban areas a range of 0-2 km is used. In rural areas, where communities are more dispersed and risk of disclosure may be higher, a range of 0-5 km offset is used. An additional 0-10 km offset for 1% of rural clusters effectively increases the known range for all rural points to 10 km while introducing only a small amount of noise. Offset points are constrained at the state level, so that they still fall within the correct state for spatial joins, although boundary precision may be an issue for clusters located very close to the border.

In this wave of panel data collection some households are tracked to a new location. These include both local and long-distance moves, although a majority of tracked households are within 5 km of the original location. The public coordinates for new locations that are within 5 km of the original household location remain unchanged (modified coordinates of original sample EA). The public coordinates of tracked households that are more than 5 km from original location are assigned a new offset location, according to the method described above. Additionally, the distance from original location is provided for tracked households with new locations.

The result is a set of coordinates, representative at the cluster level, that fall within known limits of accuracy. Users should take into account the offset range when considering different types of spatial analysis. Analysis of the spatial relationships between locations in close proximity would not be reliable. However, spatial queries using medium or low resolution datasets should be minimally affected by the offsets. Zonal statistics (average or range of values within an area corresponding to the known range) could help minimize the effect of offsets when combining with large scale data or high resolution grids with a high degree of local variation.

Table: UGA_HouseholdGeovariables_Y1

Theme	Source	Dataset Title	Variable Name	Variable Type	Reference Period	Resolution	Description	Web
	AICD & RAFU	Household Distance to Main Road	dist_road	Continuous	N/A	N/A	Household distance to nearest international or national trunk road (functional class A, B)	
	CityPop and UBOS	Household Distance to Towns	dist_popcenter	Continuous	2011	N/A	Household distance to nearest town of >20,000 based on 2011 projections from UBOS	http://www.citypop.de/
	USAID FEWSNET	Household Distance to Key Market Centers	dist_market	Continuous	N/A	N/A	Household distance to nearest major market (FEWSNET key market centers)	http://www.fews.net/Pages/marketcenter.aspx?loc=3&gb=ug&l=en
	Tracks for Africa, PADKOS	Household Distance to Border Posts	dist_borderpost	Continuous	N/A	N/A	Household distance to nearest land border crossing on main road	http://tracks4africa.co.za/listings/
	UN COD-FOD	Household Distance to District Capital	dist_admctr	Continuous	N/A	N/A	Household distance to the headquarter of the district of residence, according to 2006 district boundaries	http://cod.humanitarianresponse.info/
Climatology	UC Berkeley	WorldClim Bioclimatic Variables	af_bio_1	Continuous	1960-1990	0.008333 dd	Average annual temperature calculated from monthly climatology, multiplied by 10 (°C)	http://www.worldclim.org/bioclim
	UC Berkeley	WorldClim Bioclimatic Variables	af_bio_8	Continuous	1960-1990	0.008333 dd	Average temperature of the wettest quarter, from monthly climatology, multiplied by 10. (°C)	http://www.worldclim.org/bioclim
	UC Berkeley	WorldClim Bioclimatic Variables	af_bio_12	Continuous	1960-1990	0.008333 dd	Total annual precipitation, from monthly climatology (mm)	http://www.worldclim.org/bioclim
	UC Berkeley	WorldClim Bioclimatic Variables	af_bio_13	Continuous	1960-1990	0.008333 dd	Precipitation of wettest month, from monthly climatology (mm)	http://www.worldclim.org/bioclim
	UC Berkeley	WorldClim Bioclimatic Variables	af_bio_16	Continuous	1960-1990	0.008333 dd	Precipitation of wettest quarter, from monthly climatology (mm)	http://www.worldclim.org/bioclim
Landscape Typology	ESA and UC Louvain	GlobCover v 2.3	fsrad3_lcmaj	Categorical	2009	0.002778 dd	Majority landcover class within approximately 1km buffer	http://ionia1.esrin.esa.int/
	ESA and UC Louvain	GlobCover v 2.3	fsrad3_agpct	Continuous	2009	0.002778 dd	Percent under agriculture within approx 1 km buffer	http://ionia1.esrin.esa.int/

Theme	Source	Dataset Title	Variable Name	Variable Type	Reference Period	Resolution	Description	Web
	IFPRI	IFPRI standardized AEZ based on elevation, climatology	ssa_aez09	Categorical		0.008333 dd	Agro-ecological zones created using WorldClim climate data and 0.0833dd resolution LGP data from IIASA.	http://harvestchoice.org/production/biophysical/agroecology
Soil & Terrain	NASA	SRTM 90m	srtm_uga	Continuous		0.000833 dd	Elevation (m)	ftp://xftp.jrc.it/pub/srtmV4/arcasci/
	USGS	Slope (percent)	slopepct_uga	Continuous		0.008333 dd	Derived from 90m SRTM, aggregated to 1km block	http://pubs.usgs.gov/of/2007/1188/ , data provided USGS upon request
	AfSIS	Topographic Wetness Index	twi_uga	Continuous		0.000833 dd	Downloaded from AfSIS website. Derived from modified 90m SRTM. Local upslope contributing area and slope are combined to determine the potential wetness index: $WI = \ln(A_s / \tan(b))$ where A_s is flow accumulation or effective drainage area and b is slope gradient.	http://www.ciesin.columbia.edu/afsis/bafsis_fullmap.htm#
	LSMS-ISA	Terrain Roughness	srtm_uga_5_15	Categorical		0.000833 dd	Derived from 90m SRTM using 15 Meybeck relief classes and 5x5 pixel neighborhood	
	FAO	Harmonized World Soil Database	SQ1	Categorical		0.083333 dd	Nutrient availability	http://www.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/
	FAO	Harmonized World Soil Database	SQ2	Categorical		0.083333 dd	Nutrient retention capacity	http://www.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/
	FAO	Harmonized World Soil Database	SQ3	Categorical		0.083333 dd	Rooting conditions	http://www.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/
	FAO	Harmonized World Soil Database	SQ4	Categorical		0.083333 dd	Oxygen availability to roots	http://www.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/
	FAO	Harmonized World Soil Database	SQ5	Categorical		0.083333 dd	Excess salts	http://www.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/

Theme	Source	Dataset Title	Variable Name	Variable Type	Reference Period	Resolution	Description	Web
Crop Season Parameters	FAO	Harmonized World Soil Database	SQ6	Categorical		0.083333 dd	Toxicity	http://www.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/
	FAO	Harmonized World Soil Database	SQ7	Categorical		0.083333 dd	Workability (constraining field management)	http://www.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/
	NOAA CPC	Rainfall Estimates (RFE)	anntot_avg	Continuous	2001-2010	0.1 dd	Avg 12-month total rainfall (mm) for Jan-Dec	ftp://ftp.cpc.ncep.noaa.gov/fews/newalg_o_est_dekad/
	NOAA CPC	Rainfall Estimates (RFE)	wetQ_avg	Continuous	2001-2010	0.1 dd	Avg rainfall (mm) in wettest quarter within Jan-Dec, or Jan-Jun for bimodal	ftp://ftp.cpc.ncep.noaa.gov/fews/newalg_o_est_dekad/
	NOAA CPC	Rainfall Estimates (RFE)	wetQ_avgstart	Continuous	2001-2010	0.1 dd	Avg start of wettest quarter in dekads 1-36, where first week of January = 1	ftp://ftp.cpc.ncep.noaa.gov/fews/newalg_o_est_dekad/
	NOAA CPC	Rainfall Estimates (RFE)	anntot_2009	Continuous	2009	0.1 dd	12-month total rainfall (mm) in Jan-Dec, starting January 2009	ftp://ftp.cpc.ncep.noaa.gov/fews/newalg_o_est_dekad/
	NOAA CPC	Rainfall Estimates (RFE)	wetQ_2009	Continuous	2009	0.1 dd	Rainfall (mm) in wettest quarter within Jan-Dec 2009, or Jan-Jun for bimodal	ftp://ftp.cpc.ncep.noaa.gov/fews/newalg_o_est_dekad/
	NOAA CPC	Rainfall Estimates (RFE)	wetQstart_2009	Continuous	2001-2010	0.1 dd	Start of wettest quarter in dekads 1-36, where first week of January 2009 = 1	ftp://ftp.cpc.ncep.noaa.gov/fews/newalg_o_est_dekad/
	NOAA CPC	Rainfall Estimates (RFE)	wetQ2_avg	Continuous	2001-2010	0.1 dd	Avg rainfall in wettest quarter in second growing season Jul-Dec, bimodal only	ftp://ftp.cpc.ncep.noaa.gov/fews/newalg_o_est_dekad/
	NOAA CPC	Rainfall Estimates (RFE)	wetQ2_avgstart	Continuous	2009	0.1 dd	Avg start of wettest quarter in second growing season in dekads, bimodal only	ftp://ftp.cpc.ncep.noaa.gov/fews/newalg_o_est_dekad/
	NOAA CPC	Rainfall Estimates (RFE)	wetQ2_2009	Continuous	2009	0.1 dd	Rainfall (mm) in wettest quarter in second growing season of 2009, bimodal only	ftp://ftp.cpc.ncep.noaa.gov/fews/newalg_o_est_dekad/
	NOAA CPC	Rainfall Estimates (RFE)	wetQ2start_2009	Continuous	2009	0.1 dd	Start of wettest quarter in second growing season in dekads 19-36, bimodal only	ftp://ftp.cpc.ncep.noaa.gov/fews/newalg_o_est_dekad/

Theme	Source	Dataset Title	Variable Name	Variable Type	Reference Period	Resolution	Description	Web
	NASA / Boston University	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	rf_regime	Categorical	2001-2010	0.004176 dd	District-level assignment of predominantly bi-modal or uni-modal growing season, derived from phenology data	
	NASA / Boston University	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	eviarea_avg	Continuous	2001-2010	0.004176 dd	Avg total change in greenness in main, or first, growing season, avg by district	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
	NASA / Boston University	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	evimax_avg	Continuous	2001-2010	0.004176 dd	Avg EVI value at peak in main, or first, growing season, avg by district	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
	NASA / Boston University	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	grn_avg	Continuous	2001-2010	0.004176 dd	Avg onset of greenness increase in day of year 1-356, avg by district	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
	NASA / Boston University	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	sen_avg	Continuous	2001-2010	0.004176 dd	Avg onset of greenness decrease in day of year 1-356, avg by district	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
	NASA / Boston University	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	eviarea_2009	Continuous	2009	0.004176 dd	Total change in greenness within main, or first, growing season 2009	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
	NASA / Boston University	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	evimax_2009	Continuous	2009	0.004176 dd	EVI value at peak of greenness within main, or first, growing season 2009	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
	NASA / Boston University	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	grn_2009	Continuous	2009	0.004176 dd	Onset of greenness increase in day of year in 2009, avg by district	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
	NASA / Boston University	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	sen_2009	Continuous	2009	0.004176 dd	Onset of greenness decrease in day of year in 2009, avg by district	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
	NASA / Boston University	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	eviarea2_avg	Continuous	2001-2010	0.004176 dd	Avg total change in greenness in second growing season, avg by district	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
	NASA / Boston University	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	evimax2_avg	Continuous	2001-2010	0.004176 dd	Avg EVI value at peak in second growing season, avg by district	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005

Theme	Source	Dataset Title	Variable Name	Variable Type	Reference Period	Resolution	Description	Web
	NASA / Boston University	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	grn2_avg	Continuous	2001-2010	0.004176 dd	Avg onset of greenness increase in second growing season, avg by district	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
	NASA / Boston University	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	sen2_avg	Continuous	2001-2010	0.004176 dd	Avg onset of greenness decrease in second growing season, avg by district	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
	NASA / Boston University	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	eviarea2_2009	Continuous	2009	0.004176 dd	Total change in greenness within second growing season of 2009	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
	NASA / Boston University	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	evimax2_2009	Continuous	2009	0.004176 dd	EVI value at peak of greenness within second growing season of 2009	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
	NASA / Boston University	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	grn2_2009	Continuous	2009	0.004176 dd	Onset of greenness increase in second growing season of 2009, avg by district	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005
	NASA / Boston University	MOD12Q2 Land Cover Dynamics (PHENOLOGY)	sen2_2009	Continuous	2009	0.004176 dd	Onset of greenness decrease in second growing season of 2009, avg by district	ftp://e4ftl01.cr.usgs.gov/MOTA/MCD12Q2.005