

NATIONAL BUREAU OF STATISTICS

# Basic Information Document

National Panel Survey  
UNIVERSAL PANEL QUESTIONNAIRE

2018

[UNITED REPUBLIC OF TANZANIA]

## ACRONYMS

BID	-	Basic Information Document
HBS	-	Household Budget Survey
HHID	-	Household Identification
ILFS	-	Integrated Labour Force Survey
LSMS	-	Living Standards Measurement Survey
MDG	-	Millennium Development Goal
NBS	-	National Bureau of Statistics
NPS	-	National Panel Survey
NSCA	-	National Census Sample of Agriculture
PHC	-	Population and Housing Census
UPHI	-	Universal Panel Household Identifier
UPI	-	Universal Panel Identifier
UPQ	-	Universal Panel Questionnaire

## Table of Contents

<b>TITLE PAGE .....</b>	<b>1</b>
<b>I. INTRODUCTION .....</b>	<b>4</b>
<b>II. BACKGROUND .....</b>	<b>4</b>
<b>III. OVERVIEW.....</b>	<b>5</b>
<b>IV. SURVEY INSTRUMENT .....</b>	<b>5</b>
A. QUESTIONNAIRE LAYOUT .....	5
1. <i>Variable Names</i> .....	6
2. <i>Response Codes</i> .....	6
B. CHANGES IN QUESTIONS OVER ROUNDS.....	6
C. CHANGES IN RESPONSE CODES OVER ROUNDS .....	7
<b>V. DATASETS.....</b>	<b>7</b>
A. PANEL SEQUENCE .....	8
1. <i>Household Panel Sequences</i> .....	8
B. DATA SET LAYOUT.....	11
1. <i>Naming Convention for “r_hhid”</i> .....	11
2. <i>Merging Between Data Sets</i> .....	12
3. <i>Obtaining Data</i> .....	12
<b>VII. UPQ MODULE INFORMATION.....</b>	<b>13</b>
<b>VIII. APPENDIX .....</b>	<b>26</b>
A. APPENDIX I: NPS UPQ NAMING CONVENTION .....	26
B. APPENDIX II: QUESTIONNAIRE FOOTNOTES.....	27

## I. INTRODUCTION

Panel data possess several advantages over conventional cross-sectional and time-series data, including their power to isolate the effects of specific actions, treatments, and general policies often at the core of large-scale econometric development studies. While the concept of panel data alone provides the capacity for modeling the complexities of human behavior, the notion of *universal* panel data – in which time- and situation-driven variances leading to variations in tools, and thus results, are mitigated – can further enhance exploitation of the richness of panel information.

This Basic Information Document (BID) provides a brief overview of the Tanzania National Panel Survey (NPS), but focuses primarily on the theoretical development and application of panel data, as well as key elements of the universal panel survey instrument and datasets generated by the four rounds of the NPS. As this BID does not describe in detail the background, development, or use of the NPS itself, the round-specific NPS BIDs should supplement the information provided here.

The NPS Universal Panel Questionnaire (UPQ) consists of both survey instruments and datasets, meticulously aligned and engineered with the aim of facilitating the use of and improving access to the wealth of panel data offered by the NPS. The NPS-UPQ provides a consistent and straightforward means of conducting not only user-driven analyses using convenient, standardized tools, but also for monitoring MKUKUTA, FYDP II, and other national level development indicators reported by the NPS.

The design of the NPS-UPQ combines the four completed rounds of the NPS – NPS 2008/09 (R1), NPS 2010/11 (R2), NPS 2012/13 (R3), and NPS 2014/15 (R4)<sup>1</sup> – into pooled, module-specific survey instruments and datasets. The panel survey instruments offer the ease of comparability over time, with modifications and variances easily identifiable as well as those aspects of the questionnaire which have remained identical and offer consistent information. By providing all module-specific data over time within compact, pooled datasets, panel datasets eliminate the need for user-generated merges between rounds and present data in a clear, logical format, increasing both the usability and comprehension of complex data.

## II. BACKGROUND

The NPS is a series of nationally representative household panel surveys that collect information on a wide range of topics including agricultural production, non-farm income generating activities, consumption expenditures, and a wealth of other socio-economic characteristics. The survey was first implemented in 2008/2009 (Table 1) and all subsequent rounds have been implemented by the Tanzania National Bureau of Statistics (NBS).

**Table 1. NPS Rounds and Dates**

ROUND	NPS	DATES
1	NPS 2008/2009	October 2008 – September 2009
2	NPS 2010/2011	October 2010 – November 2011
3	NPS 2012/2013	October 2013 – November 2013
4	NPS 2014/2015	October 2014 – January 2016
5	NPS 2018/2019	Spring 2018 - TBD

The main objective of the NPS is to provide high-quality household-level data to the Tanzanian government and other stakeholders for monitoring poverty dynamics, tracking the progress of national poverty reduction strategy

---

<sup>1</sup> For simplicity, NPS rounds will be hereafter be referred to as the following: R1 (NPS 2008/09), R2 (NPS 2010/11), R3 (NPS 2012/13), and R4 (NPS 2014/15).

such as the FYPD II and MKUKUTA II<sup>2</sup>, and to evaluate the impact of other major, national-level government policy initiatives. As an integrated survey covering a number of different socioeconomic factors, it compliments other more narrowly focused survey efforts, such as the Demographic and Health Survey on health, the Integrated Labour Force Survey (ILFS) on labour markets, the Household Budget Survey (HBS) on expenditure, and the National Sample Census of Agriculture (NSCA).

As a panel household survey in which the same households are revisited over time, the NPS allows for the study of poverty and welfare transitions and the determinants of living standard changes. The NPS-UPQ survey instruments and datasets were generated following the NPS 2014/15, after recognizing the value and vast quantity of the data available and the need for easily accessible and usable datasets to measure such determinants.

### III. OVERVIEW

The NPS-UPQ consists of a universal panel survey instrument and universal panel datasets, presented by module and in a format similar to previously released round-specific NPS products. This document describes fundamental elements of the NPS-UPQ, including information on **Survey Instruments** (Section IV), **Datasets** (Section V), **Panel Sequences** (Section VI), and **Module-Specific Information** (Section VII). Please see the NPS BIDs for each round<sup>3</sup> for detailed descriptions of the sample design used in each round and their respective implementation efforts. Sample design information pertinent to the use and comprehension of the NPS-UPQ data sets is summarized in Section VI of this document.

### IV. SURVEY INSTRUMENT

The NPS-UPQ survey instrument was developed to highlight the similarities and examine the comparability of both questions and response codes across rounds and allows for easy identification of information reported in all rounds of the NPS. The NPS-UPQ survey instrument includes 26 module-specific questionnaires which correlate directly to the NPS-UPQ datasets in terms of module names, variable names, variable labels, and response codes.

#### A. Questionnaire Layout

The format of the NPS-UPQ survey instrument is similar to previously disseminated NPS survey instruments. Each module has a questionnaire and clearly identifies if the module collects information at the individual or household level. Within each module-specific questionnaire of the NPS-UPQ survey instrument, there are five distinct sections, arranged vertically: (1) the UPQ - “U” on the survey instrument, (2) R4, (3), R3, (4) R2, and (5) R1 – the latter 4 sections presenting each questionnaire in its original form at time of its respective dissemination.

The uppermost section of each module’s questionnaire (“U”) represents the model *universal panel questionnaire*, with questions generated from the comprehensive listing of questions across all four rounds of the NPS and codes generated from the comprehensive collection of codes. The following sections are arranged vertically by round, considering R4 as most recent. While not all rounds will have data reported for each question in the UPQ and not each question will have reports for each of the UPQ codes listed, the NPS-UPQ survey instrument represents the visual, all-inclusive set of information collected by the NPS over time.

---

<sup>2</sup> MKUKUTA II is a Kiswahili acronym for the National Strategy for Growth and Reduction of Poverty. The current phase began in 2010 and will finish in 2015, and is the government strategy to meet the MDGs and other national development goals. The focus of this round of MKUKUTA II includes growth and the reduction of poverty, improved quality of life and social well-being, and governance and accountability.

<sup>3</sup> Data and background documentation for the NPS are available free of charge on the Tanzania National Bureau of Statistics website ([www.nbs.go.tz](http://www.nbs.go.tz)), as well as the World Bank LSMS-ISA website ([www.worldbank.org/lsms-isa](http://www.worldbank.org/lsms-isa)).

The four round-specific sections (R4, R3, R2, R1) are aligned with their UPQ-equivalent question, visually presenting their contribution to compatibility with the UPQ. Each round-specific section includes the original round-specific variable names, response codes and skip patterns (corresponding to their respective round-specific NPS data sets, and despite their variance from other rounds or from the comprehensive UPQ code listing)<sup>4</sup>.

### 1. Variable Names

Variable names for UPQ questions were assigned following the comprehensive listing of questions from all four rounds within each particular module and used the naming convention defined in Table 2. Therefore, the UPQ variable name will not always match the original round-specific variable name.

**Table 2. UPQ Variable Naming Convention**

EXAMPLE 1: No Subset	EXAMPLE 2: Subset Question
Variable Name: <b>hbq01</b>	Variable Name: <b>hiq34_1</b>
<i>where</i> <i>h</i> = household <i>b</i> = module B <i>q</i> = question <i>01</i> = question number one	<i>where</i> <i>h</i> = household <i>i</i> = module i <i>q</i> = question <i>34</i> = question number one <i>_1</i> = subset question of question 34

### 2. Response Codes

Response codes for each UPQ question were generated from the comprehensive listing of response codes from all four NPS rounds for that particular question. Similar to the variable naming structure, since response code values for the UPQ questions were re-assigned following this comprehensive listing they will not necessarily correspond to each original round-specific response code value. The UPQ code listings intend to present the all-inclusive set of information gathered at any point in time by the NPS.

#### B. Changes in Questions Over Rounds

**Identical Questions:** Whenever possible, questions that are identical or nearly identical across rounds (i.e. “What kind of toilet facility do members of your household generally use?” and “What is the main toilet facilities usually used in this household?”) are presented as a single, unified UPQ question. These will be vertically aligned in the module’s questionnaire. Minor changes in question wording that are not anticipated to affect interpretation or results are noted as footnotes, but the slight variations will not be presented as unique UPQ questions.

**Distinct Variations of Questions:** When a question is considerably different across rounds (i.e. “What measures does this household take to ensure the safety of drinking water?” and “What do you usually do to the water to make it safer to drink in the dry season?”), all variations of the question will be included in the UPQ questionnaire with the appropriate blackouts for the respective rounds that the UPQ question is not present in. For these cases, it will be up to the user’s discretion to combine.

**Prevalence of Questions:** Whether a question exists in some or all of the four NPS rounds, it will be included in the UPQ questionnaire and will either have the respective round-specific variables presented below it, or will have the appropriate blackouts for the respective rounds that the question is not present in.

<sup>4</sup> Skip patterns within the round-specific sections of questionnaires are directed to the variable name of the UPQ to facilitate ease of direction, as more complicated modules often cannot maintain the conventional numerical order of variables within each of the round (R4 is again considered most recent and therefore is best preserved as the conventional structure).

### C. Changes in Response Codes Over Rounds

**Identical Response Codes:** Whenever possible, response codes that are identical or nearly identical (i.e. “OTHER” and “OTHER, SPECIFY” or “BOIL” and “BOILING WATER”) across rounds are presented as a single, unified UPQ response code. Minor changes in wording that are not anticipated to affect interpretation or results are noted as footnotes, but the slight variations will not be presented as unique UPQ response codes.

**Distinct Variations of Response Codes:** When a response code is considerably different across rounds (i.e. “RELIGIOUS” and “RELIGIOUS – BOARDING” or “PIT LATRINE”, “PIT LATRINE WITH SLAB (WASHABLE)” and “PIT LATRINE WITH SLAB (NOT WASHABLE)”), all variations of the response code will be included in the UPQ response code listing. For these cases, it will be up to the user’s discretion to combine.

**Prevalence of Response Codes:** Whether a response code exists in some or all of the four NPS rounds, it will be included in the UPQ code listing and will be identified in the round-specific sections by a box. If additional information is relevant, a footnote will accompany the code. In these instances, the values for codes may differ from the original, round-specific codes to accommodate for more, less, or different codes. These differences will also be identified in the round-specific sections by a box.

## V. DATASETS

The NPS-UPQ datasets were developed in accordance with the NPS-UPQ survey instrument, and aim to neatly present and facilitate the use of the wealth of panel data collected by the NPS. By offering compact, standardized data sets, data are readily available and pre-standardized for optimal panel analyses.

The design of the universal panel data sets combines the four rounds of NPS data – NPS 2008/09 (R1), NPS 2010/11 (R2), NPS 2012/13 (R3), and NPS 2014/15 (R4) – into pooled, module-specific data sets. The NPS-UPQ consists of 41 data files, each pertaining to a single module or portion of the module (i.e. “UPQ4.HH.MOD.Q1” and “UPQ4.HH.MOD.Q2”). Each data file will include data for all rounds that reported information on that module. As mentioned, modules will occasionally be split into multiple datasets if the level of observation is not consistent within the module. The complete set of data files, as well as the level of observation, unique identifier for that module, and rounds of data included, are described in the tables of Section VII.

### SAMPLE DESIGN

While the same sample of respondents was maintained over the first three rounds of the NPS, longitudinal surveys tend to suffer from bias introduced by households leaving the survey over time; i.e. attrition. Although the NPS maintains a highly successful recapture rate (roughly 96% retention at the household level), minimizing the escalation of this selection bias, a refresh of longitudinal cohorts was done for the NPS 2014/15 to ensure proper representativeness of estimates while maintaining a sufficient primary sample to maintain cohesion within panel analysis. A newly completed Population and Housing Census (PHC) in 2012, providing updated population figures along with changes in administrative boundaries, emboldened the opportunity to realign the NPS sample and abate collective bias potentially introduced through attrition.

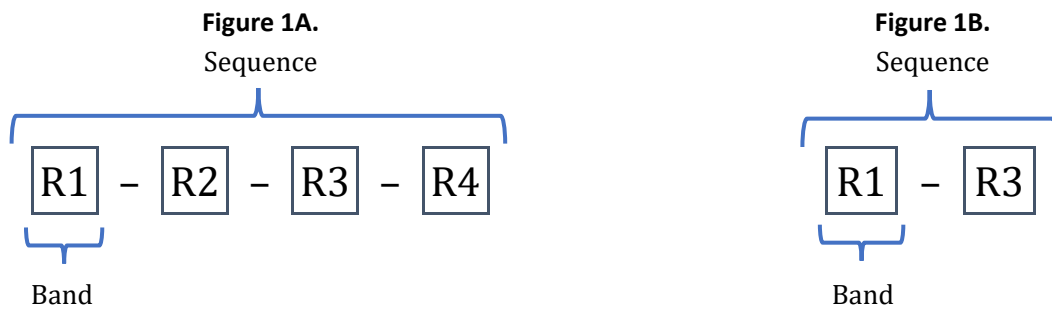
To maintain the panel concept of the NPS, the sample design for NPS 2014/2015 consisted of a combination of the original NPS sample and a new NPS sample. A nationally representative sub-sample was selected to continue as part of the “Extended Panel” while an entirely new sample, “Refresh Panel”, was selected to represent national and sub-national domains. Similar to the sample in NPS 2008/2009, the sample design for the “Refresh Panel” allows analysis at four primary domains of inference, namely: Dar es Salaam, other urban areas on mainland Tanzania, rural mainland Tanzania, and Zanzibar. This new cohort in NPS 2014/2015 will be maintained and tracked in all future rounds between national censuses.

## A. Panel Sequence

Not every NPS household or individual will have data present for all four rounds. In each round, circumstances such as tracking success, attrition through loss or refusal, and the revised sampling design in R4 contribute to a variety of *panel sequences* that represent distinct combinations of movement and participation over time. As both households and individuals are tracked over time, both *household panel sequences* and *individual panel sequences* are formed. These sequences correspond to the data sets of the same information level i.e. household panel sequences are observed in household-level datasets (e.g. education) and individual panel sequences are observed in individual-level datasets (e.g. housing, water, and sanitation).

### 1. Household Panel Sequences

Each household panel sequence represents a unique combination of household movement and participation over time. Sequences are composed of “bands” of data, which are determined by a successful interview and thus the presence of data in a particular round.



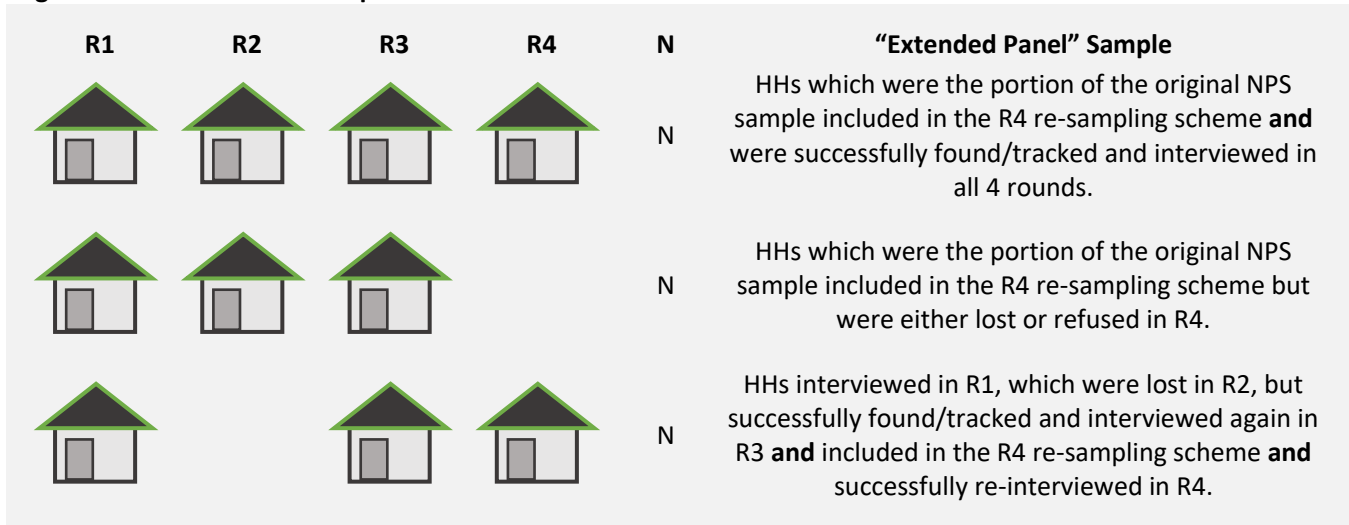
The concept of these sequences is integral for understanding the architecture of the universal panel datasets. While the following illustrations visually present “gaps” in data (i.e. from a lost household or refusal), no gap or placeholder is included in either the sequences or in the universal panel datasets (Figure 1B); sequences are compiled only of existing bands of data. While sequences differ within the “Extended Panel Sample”, “Refresh Panel Sample”, and “Initial Sample (R1 – R3)”, the eight sequences below create the universe of household panel sequences observable in the household-level universal panel datasets.

#### a) Extended Panel Sample

From the original NPS sample, a nationally representative sub-sample was selected to continue as the “Extended Panel” sample, with data collection attempts made in each of the four rounds. Of these, ### were successfully found/tracked and interviewed to completion in each of R1, R2, R3, and R4 (Figure 2, Example 1), regardless of individual member movement in or out of the household (as long as an original member was continuously present in the household for each successive round) or the physical movement from one structure to another structure.

The *household panel sequence* for the extended panel sample is thus generated from the progression of the originating household in R1 through R4 and consists of a maximum of four possible bands of data. In very few instances (##), a household was lost in R2 but found/tracked in the subsequent rounds and re-integrated into the panel (Figure 2, Example 3). As the revised sampling design in R4 was derived from the sample from R3, there will be no panel sequence of R1-R2-R4 in the universal datasets.

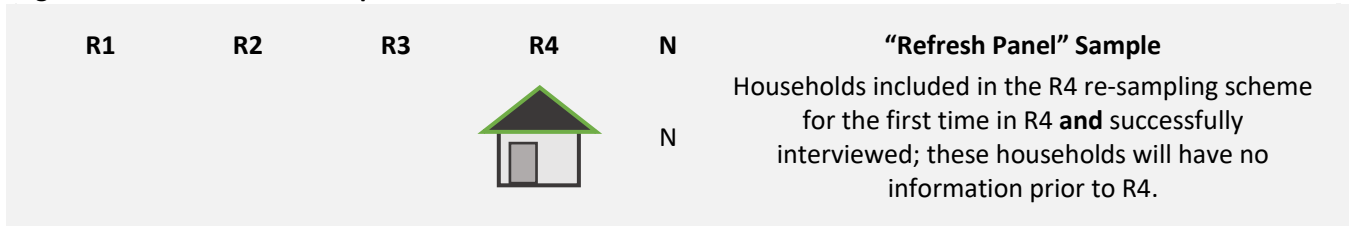
**Figure 2. Household Panel Sequences**



b) *Refresh Panel Sample*

As households in the “Refresh Panel” sample were interviewed for the first time in R4, the *household panel sequence* for these households is simply the single band of data available in R4 (Figure 3).

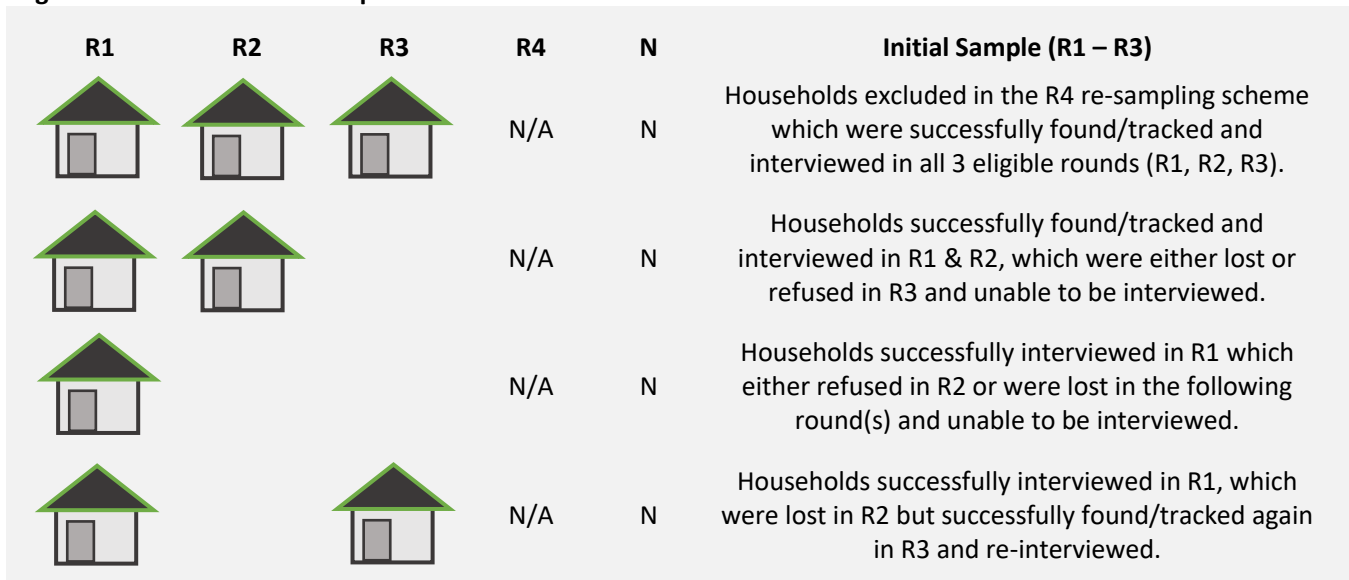
**Figure 3. Household Panel Sequences**



c) *Initial Sample (R1 – R3)*

The remaining sequences are for households interviewed under the initial sampling scheme (i.e. households *excluded* in the R4 revised sampling scheme). The *household panel sequence* for these households is thus generated from the progression of the original household in R1 through the following two rounds (Figure 4).

**Figure 4. Household Panel Sequence**

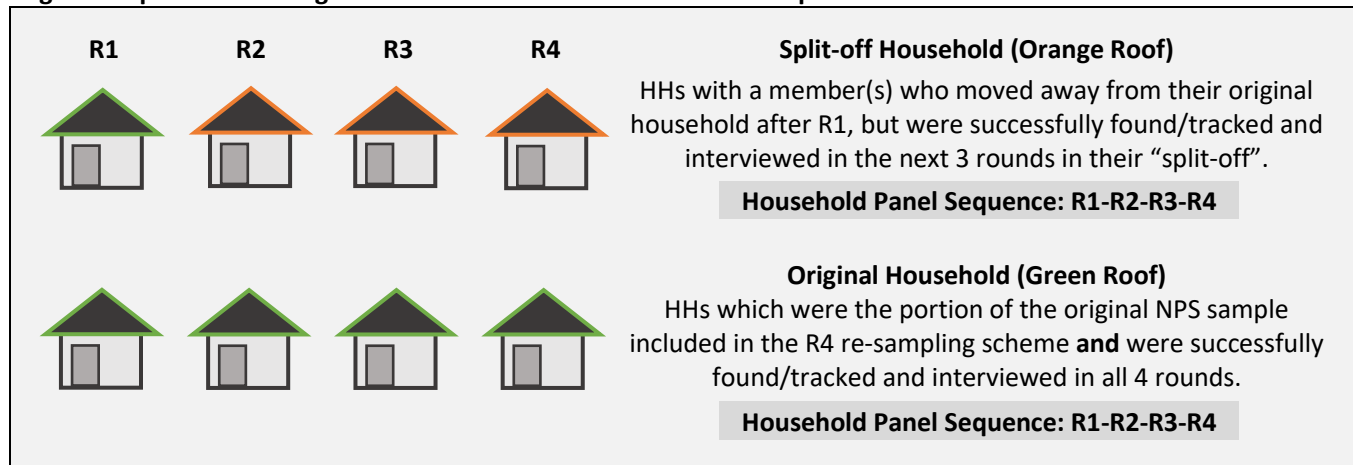


d) *Split-off Households*

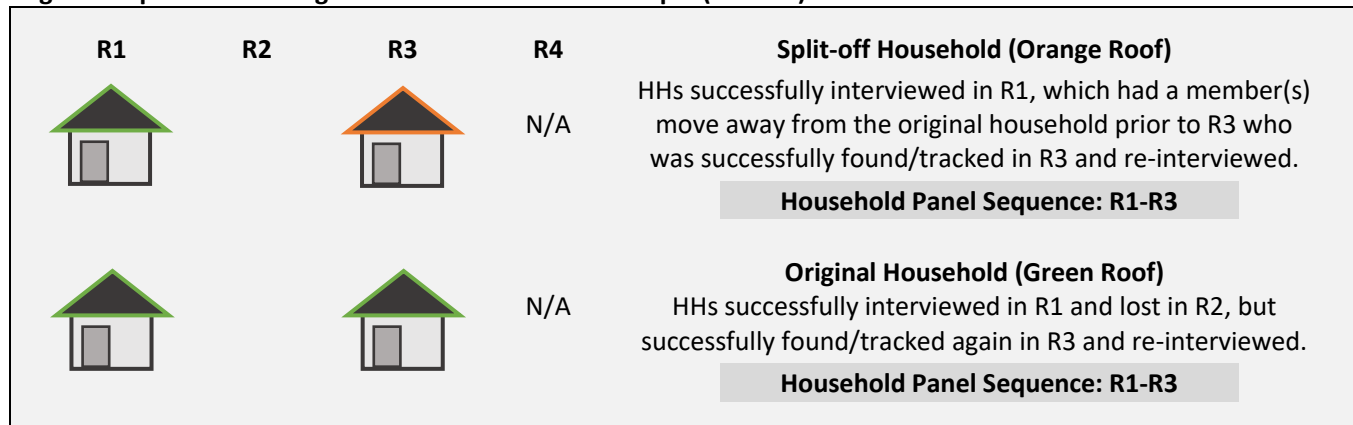
When a member(s) moves away from their original NPS household they form a “split-off” household, which is also tracked and interviewed as a part of the NPS sample. Household panel sequences must include an origination household, therefore the household panel sequence for split-off households is generated from a combination of the original household(s) and the subsequent split-off household(s). In other words, while R2-R3-R4 is technically a sequence of movement for a new *member* in a split-off household (i.e. a spouse that marries an original NPS member and now lives with them in a different household), it does not characterize a *household* panel sequence – the R1 household of the original NPS member now in the split-off household with the new member(s) would be included to create the household panel sequence (Figure 5, Example 1).

As a split-off can occur at any point in time, these households exist in both the extended panel sample and the initial sample. Since household panel sequences serve as a representation of movement and participation and not household make-up, it is important to note that the household panel sequences for a split-off household **do not** differ from the sequences for an original NPS household. Figure 5 and 6 illustrate two examples of a split-off household and the identical sequence for an original NPS household.

**Figure 5. Split-Off and Original Household in “Extended Panel” Sample**



**Figure 6. Split-Off and Original Household in Initial Sample (R1 – R3)**



## B. Data Set Layout

The NPS-UPQ datasets are arranged in a *long* format, sorted by a universal identifier (either at the individual or household level) and then by round (Table 3). Two variables were generated during development of the UPQ datasets to uniquely identify panel sequences after combining all four rounds of data: the Universal Panel Identifier (UPI4) for individuals and the Universal Panel Household Identifier (UPHI4).

The following variables shape the foundation of the NPS-UPQ datasets:

- *UP4I*: unique identifier for each individual panel sequence
- *UPHI4*: unique identifier for each household panel sequence
- *round*: numeric variable with values = 1 - 4 corresponding to the four rounds of NPS data; for panel sequences missing information (i.e. lost or refused) there will be no placeholder, the data line will simply not exist (Example: Table 3, *UPHI4* = 336 only presents data for R1 and R2)
- *r\_hhid*: full 14-, 16-, or 7-digit HHID corresponding to the round per data line

Module-level variables (i.e. hiq01) in the UPQ data sets utilize a combination of “h” for household and the module letter (i.e. “i”) as the prefix to distinguish them from round-specific data releases. Again, the question number in the UPQ variable name will not necessarily match the number of their round-specific variable; however, they will correspond directly to the UPQ variable names in the NPS-UPQ survey instrument.

**Table 3. Stata Dataset Excerpt – HH\_Module\_I – Housing, Water, & Sanitation**

	UPHI4	round	r_hhid	hiq01	hiq02
899	334	1	53020150090067	OWNER OCCUPIED	NO DOCUMENTATION AT ALL
900	334	2	5302015009006701	OWNER OCCUPIED	NO DOCUMENTATION AT ALL
901	334	3	3961-001	OWNER OCCUPIED	NO DOCUMENTATION AT ALL
902	335	1	53020150090067	OWNER OCCUPIED	NO DOCUMENTATION AT ALL
903	335	2	5302015009006703	OWNER OCCUPIED	NO DOCUMENTATION AT ALL
904	335	3	3692-001	OWNER OCCUPIED	NO DOCUMENTATION AT ALL
905	336	1	53020150090065	OWNER OCCUPIED	TITLE DEED FOR LAND
906	336	2	5302015009006501	OWNER OCCUPIED	TITLE DEED FOR LAND

### 1. Naming Convention for “r\_hhid”

Please see round-specific NPS BIDs for detailed explanations on the naming conventions of household IDs (HHID). The round-specific HHID variables are known as the following in their individual releases: “hhid” (R1), “y2\_hhid” (R2), “y3\_hhid” (R3), and “y4\_hhid” (R4). These have been condensed under the single variable “r\_hhid” in the UPQ data sets and are distinguishable only in combination with “round”.

For both originating and split-off households, the HHID in post-R1 rounds is generated using the lowest two-digit roster ID (e.g. 1=01; 3=03; 12=12) of the original member(s) in the household in the subsequent round. The other digits vary between rounds (i.e. R1 and R2 were longer numeric strings which were then minimized in R3 and maintained or recreated in R4 for extended and refresh panel households, respectively).

As data layout and round labeling is identical between originating and split-off households, the latter is identified only by their HHID, whose last two digits in the split-off round will be the lowest roster ID of the split-off member(s) and a value other than “01” (Example: Table 3, UPHI = 335, round = 2).

Figure 6 illustrates a member(s) who has moved from their original household and returned to the NPS sample in a non-sequential round (i.e. R1 and again in R3). While all previously stated panel concepts remain, the last three digits of the split-off HHID in R3 are in the 100’s to indicate **uncollected data** in R2 (Table 4).

**Table 4. Stata Dataset Excerpt – HH\_Module\_I – Housing, Water, & Sanitation**

	UPHI	round	r_hhid	up_hh_i01	up_hh_i02
11	5	1	55020180210068	OWNER OCCUPIED	NO DOCUMENTATION AT ALL
12	5	3	3922-108	OWNER OCCUPIED	LAND SALE AGREEMENT

## 2. Merging Between Data Sets

The design of the UPQ eliminates the need for user-generated merges *across* rounds for a particular module; however, merges between different modules will still occur (i.e. merging an individual's roster and education information, or food consumption and household sanitation information). The individual identification variable is the Universal Panel Identifier ("UPI4") and the household identification variable is Universal Panel Household Identifier ("UPHI4") in the data files.

The panel nature of these datasets means that most merging is simplified to a single, unique variable describing either an "individual" or a "household", either **UPI4** or **UPHI4**, in accompaniment with **round**. *Note: See previous section for a description of the generation of these identifying variables.* When merging or linking most individual-level data files, it is necessary to use only the individual identification variable, **UPI4**, and round (as each individual with that UPI4 will have between 1-4 data lines, each representing a round). When merging or linking most household-level data files, it is necessary to use only the household identification variable, **UPHI4**, and **round**<sup>5</sup>. All household level sections have been harmonized to provide a consistent number of observations in each module across like units of observation.

Additional variables required to merge data across sections depend on the particular sections being merged. In modules with multiple items (i.e. Module J1: Consumption of Food Over the Past One Week, or Module P: Credit, etc.), the specific item code will need to be included with the combination of **UPHI4** and **round**. The unique identifying combination of variables in each module can be found in the Section VII tables.

## 3. Obtaining Data

Data sets for the NPS-UPQ are available free of charge on the Tanzania National Bureau of Statistics website ([www.nbs.go.tz](http://www.nbs.go.tz)), as well as the World Bank LSMS-ISA website ([www.worldbank.org/lms-isa](http://www.worldbank.org/lms-isa)), and the World Bank Microdata Library ([microdata.worldbank.org](http://microdata.worldbank.org)). Inquiries pertaining to the data may be sent to the LSMS team at [lms@worldbank.org](mailto:lms@worldbank.org).

---

<sup>5</sup> The **r\_hhid** variable can be substituted for the **round** variable as the unique identifier. See Section VII: UPQ Module Information for exact unique identifiers in each module and all possible combinations.

## VII. UPQ MODULE INFORMATION

MODULE B: HOUSEHOLD MEMBER ROSTER			
<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Individual UPI4, round <u>or</u> UPI4, r_hhid	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_B.dta R1, R2, R3, R4
<p><b>Description:</b> Roster of household members, individual characteristics including: sex, age, relationship to the household head, panel member identification, presence in household, general occupation, parental status, place of birth, marital status, and spouse identification.</p> <p><b>Original Data Files:</b>  <b>R1:</b> SEC_B_C_D_E1_F_G1_U.dta      <b>R3:</b> HH_SEC_B.dta  <b>R2:</b> HH_SEC_B.dta      <b>R4:</b> HH_SEC_B.dta</p> <p><b>Key Notes:</b> For UPQ question <b>hb_06</b>: if an individual was not present in the previous round, the enumerator is instructed to record “99”. In “<b>UPQ4_HH_MOD_B.dta</b>,” value labels for this variable were adjusted to distinctly represent what round they were not present in, where “99” = “NOT SURVEYED IN HBS”, “999” = “NOT PRESENT IN Y1”, “9999” = “NOT PRESENT IN Y2”, and “99999” = NOT PRESENT IN Y3”.</p>			
MODULE C: EDUCATION			
<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Individual UPI4, round <u>or</u> UPI4, r_hhid	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_C.dta R1, R2, R3, R4
<p><b>Description:</b> Educational enrollment and attainment; literacy, education history/level attainment, current enrollment, school characteristics, temporary absence, education examination (participation and score) and education expenditures.</p> <p><b>Original Data Files:</b>  <b>R1:</b> SEC_B_C_D_E1_F_G1_U.dta      <b>R3:</b> HH_SEC_C.dta  <b>R2:</b> HH_SEC_C.dta      <b>R4:</b> HH_SEC_C.dta</p> <p><b>Key Notes:</b> None</p>			

MODULE D: HEALTH			
<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Individual UPI4, round <u>or</u> UPI4, r_hhid	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_D.dta R1, R2, R3, R4
<p><b>Description:</b> General health status and utilization of health services; source and financing of health treatments / hospitalization, disaggregated health expenditures, disability, bed net use, pregnancy, prenatal care and births, child health and ailments / diarrhea.</p> <p><b>Original Data Files:</b>  <b>R1:</b> SEC_B_C_D_E1_F_G1_U.dta      <b>R3:</b> HH_SEC_D.dta  <b>R2:</b> HH_SEC_D.dta      <b>R4:</b> HH_SEC_D.dta</p> <p><b>Key Notes:</b> In general, the NPS advises the household head or the respondent most knowledgeable about a particular module to respond to questions asked. Even in the case of individual-level modules (i.e. health), this method facilitates the most consistent and accurate reporting. In very few cases, this instruction differed across rounds, altering the terms used in the questions themselves. In R1, respondents 12 years and older were instructed to answer for themselves, so R1 questions are worded with respect to the individual answering the question, i.e. "Have you visited a health care provider...". In R2, R3, and R4, the household head/most knowledgeable respondent was instructed to answer on behalf of the individual, and thus the questions are worded as follows: "Has [NAME] visited a health care provider...". The UPQ questionnaire uses the latter terminology as these slight variations in wording are not anticipated to affect the interpretation. A footnote is provided in Module D of the questionnaire for the first instance but this will apply to the remainder of questions in the R1 questionnaire.</p>			

MODULE E: LABOUR			
<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Individual UPI4, round <u>or</u> UPI4, r_hhid	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_E.dta R1, R2, R3, R4
<p><b>Description:</b> Labour market participation; activities including unpaid apprenticeship, wage work, non-farm enterprise, and agricultural activity (including livestock and fishery activities) in the last seven days and in the last twelve months. Also includes unemployment and steps taken to find work, secondary activities, industry (TASCO) and occupation (ISIC) codes, wages earned, contract type, employer provide safety nets, and general domestic activities.</p> <p><b>Original Data Files:</b>  <b>R1:</b> SEC_B_C_D_E1_F_G1_U.dta      <b>R3:</b> HH_SEC_E.dta  <b>R2:</b> HH_SEC_E1.dta      <b>R4:</b> HH_SEC_E.dta</p> <p><b>Key Notes:</b> In general, the NPS advises the household head or occasionally the respondent most knowledgeable about a particular module to respond to the questions asked by the enumerator. Even in the case of individual-level modules (i.e. education, health), this facilitates the most consistent and accurate reporting. In very few cases, this instruction differed across rounds, thus altering the terms used in the questions themselves.</p> <p>In R2, questions were worded in respect to the individual answering the question on behalf of themselves, i.e. "Did you do any work of any type...". In R1, R3, and R4, the pronoun "[NAME]" was used in all questions and thus is the format used in the text of the UPQ questionnaire. Again, a footnote is provided in Module E of the questionnaire for the first instance but this will apply to the remainder of questions in the R2 questionnaire.</p> <p>In R1, questions in the self-employment section used the pronoun "you"/"your" instead of "[NAME]", similar to the language used in R2. The text of UPQ questions with this footnote is identical to the text in the questions in both R1 and R2 and is meant only to highlight the deviation from the language used throughout the remainder of R1.</p>			

MODULE F: FOOD CONSUMED OUTSIDE THE HOUSEHOLD			
<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Individual UPI4, round <u>or</u> UPI4, r_hhid	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_F.dta R1, R2, R3, R4
<p><b>Description:</b> Value of food consumed outside the home during the last seven days.</p> <p><b>Original Data Files:</b></p> <p><b>R1:</b> SEC_B_C_D_E1_F_G1_U.dta      <b>R3:</b> HH_SEC_F.dta  <b>R2:</b> HH_SEC_F.dta      <b>R4:</b> HH_SEC_F.dta</p> <p><b>Key Notes:</b> None</p>			

MODULE G: SUBJECTIVE WELFARE & CRIME			
<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Individual UPI4, round <u>or</u> UPI4, r_hhid	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_G.dta R1 ( <i>partial</i> ), R2, R3, R4
<p><b>Description:</b> Self-reported level of satisfaction with health, financial status, housing, job, services, and safety. Also includes perceived status at present, and as of 2 years ago. This section also collects information about any crime that the household may have been a victim of in the past 12 months.</p> <p><b>Original Data Files:</b></p> <p><b>R1:</b> Not present      <b>R3:</b> HH_SEC_G.dta  <b>R2:</b> HH_SEC_G.dta      <b>R4:</b> HH_SEC_G.dta</p> <p><b>Key Notes:</b> No specific module was dedicated to Subjective Welfare &amp; Crime in R1; however, Module E of R1 included the question regarding level of satisfaction with various components of life, which is found in the Subjective Welfare &amp; Crime module of R2, R3, and R4. In the UPQ survey instrument and UPQ datasets, this question has been removed from Module E (SEC_B_C_D_E1_F_G1_U.dta) and provided in Module G. This question can be identified in the UPQ Module G survey instrument by the R1 variable name (i.e. <b>seq49_1</b>, aligned with UPQ variable <b>he_134_1</b>).</p>			

SECTION H: FOOD SECURITY			
<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Household UPHI4, round <u>or</u> UPHI4, r_hhid	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_H.dta R2, R3, R4
<p><b>Description:</b> Information on the household's diet and food intake patterns, months of food insecurity, and distribution of food within the household.</p> <p><b>Original Data Files:</b></p> <p><b>R1:</b> Not present      <b>R3:</b> HH_SEC_H.dta  <b>R2:</b> HH_SEC_I1.dta, HH_SEC_I2.dta      <b>R4:</b> HH_SEC_H.dta</p> <p><b>Key Notes:</b> <i>For explanation of data derivation only:</i> the original R2 NPS datasets separated the monthly incidence question (<b>hh_i09_1_01 – hh_i09_3_12</b>) from the rest of the Food Security module. However, since all Food Security data were configured in a wide format, the two R2 datasets easily translated into the UPQ datasets.</p>			

SECTION I1: HOUSING, WATER, & SANITATION			
<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Household UPHI4, round <u>or</u> UPHI4, r_hhid	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_I1.dta R1, R2, R3, R4
<p><b>Description:</b> Tenure status, rental expenditure, estimated rental value, dwelling maintenance and improvement expenditures, dwelling characteristics, including size, construction materials, toilet facilities and means of garbage disposal. Also includes, main type of cooking and lighting fuel, access to water/drinking water and time spent collecting water. The questions on water/drinking water are asked separately for the rainy and dry season.</p> <p><b>Original Data Files:</b>  <b>R1:</b> SEC_H1_J_K2_O2_P1_Q1_S1.dta      <b>R3:</b> HH_SEC_I.dta  <b>R2:</b> HH_SEC_J1.dta                      <b>R4:</b> HH_SEC_I.dta</p>			

SECTION I2: HOUSING, WATER, & SANITATION			
<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Household UPHI4, round, itemid <u>or</u> UPHI4, r_hhid, itemid	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_I2.dta R2
<p><b>Description:</b> Additional detailed information on water sources reported, including type of container used for collection, amount and frequency of collection, and average price if purchased.</p> <p><b>Original Data Files:</b>  <b>R1:</b> Not present                              <b>R3:</b> Not present  <b>R2:</b> HH_SEC_J2.dta                      <b>R4:</b> Not present</p>			

SECTION J1: CONSUMPTION OF FOOD OVER THE PAST ONE WEEK			
<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Household UPHI4, round, hj_00 <u>or</u> UPHI4, r_hhid, hj_00	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_J1_1.dta – UPQ4_HH_MOD_J1_8.dta R1, R2, R3, R4
<p><b>Description:</b> Consumption of 61 key items/ item groups over the last 7 days. Quantity and value of consumed items from purchases, own-production and gifts are recorded.</p> <p><b>Original Data Files:</b>  <b>R1:</b> SEC_K1.dta                              <b>R3:</b> HH_SEC_J1.dta  <b>R2:</b> HH_SEC_K1.dta                      <b>R4:</b> HH_SEC_J1.dta</p> <p><b>Key Notes:</b> In R1-R3, item <b>hj_00</b> = 108 was the only item code for cereal products (“Wheat, barley, grain, and other cereals”). In R4 this item was parsed into two distinct items codes: <b>hj_00</b> = 1081 (“Wheat flour”), and <b>hj_00</b> = 1082 (“Barley, grain, and other cereals”). For the UPQ, the original cereals item code (R1-R3 = 108) was recoded to 1083 to encompass the three possible categories of cereal items.</p> <p>(continued)</p>			

The Section J1 dataset has been divided into 8 categorical datasets to control size limitations. The datasets and corresponding food items are as below:

UPQ4_HH_MOD_J1_1.dta:	“Cereals and Cereal Products”	<b>hj_00:</b> 101-112
UPQ4_HH_MOD_J1_2.dta:	“Starches”	<b>hj_00:</b> 201-207
UPQ4_HH_MOD_J1_3.dta:	“Sugar and Sweets”	<b>hj_00:</b> 301-303
UPQ4_HH_MOD_J1_4.dta:	“Pulses, Dry + Nuts and Seeds”	<b>hj_00:</b> 401, 501-504
UPQ4_HH_MOD_J1_5.dta:	“Vegetables + Fruits”	<b>hj_00:</b> 601-603, 701-704
UPQ4_HH_MOD_J1_6.dta:	“Meat, Meat Products, Fish”	<b>hj_00:</b> 801-810
UPQ4_HH_MOD_J1_7.dta:	“Milk and Milk Products” + “Oil and Fats” + “Spices”	<b>hj_00:</b> 901-903, 1001-1004
UPQ4_HH_MOD_J1_8.dta:	“Beverages”	<b>hj_00:</b> 1101-1108

## SECTION J3: CONSUMPTION OF FOOD OVER THE PAST ONE WEEK - FREQUENCY OF CONSUMED ITEMS

<b>Level of Observation:</b>	Household	<b>UPQ Data File:</b>	UPQ4_HH_MOD_J3.dta
<b>Unique Identifier(s):</b>	UPHI4, round, hj3_00 <b>or</b> UPHI4, r_hhid hj3_00	<b>Rounds Present:</b>	R2, R3, R4
<b>Description:</b> Number of days general food categories where consumed by the household in the last 7 days, includes: “Cereals, Grains and Cereal Products”, “Roots, Tubers, and Plantains”, Nuts and Pulses”, Vegetables”, Meat, Fish and Animal Products”, “Fruits” “Milk/Milk Products”, “Fats/Oils”, Sugar/Sugar Products/Honey”, and “Spices/Condiments”.			
<b>Original Data Files:</b>			
<b>R1:</b> Not present	<b>R3:</b> HH_SEC_J3.dta		
<b>R2:</b> HH_SEC_K2.dta	<b>R4:</b> HH_SEC_J3.dta		
<b>Key Notes:</b> None			

## SECTION J4: QUANTITY OF MEALS AND PEOPLE CONSUMING MEALS

<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Household UPHI4, round, hj4_00 <b>or</b> UPHI4, r_hhid, hj4_00	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_J4.dta R1, R2, R3, R4
<p><b>Description:</b> Number of people living outside of the household that shared meals within the household and then total number of meals shared over the last 7 days.</p> <p><b>Original Data Files:</b></p> <p><b>R1:</b> SEC_H1_J_K2_O2_P1_Q1_S1.dta      <b>R3:</b> HH_SEC_J4.dta  <b>R2:</b> HH_SEC_K3.dta                              <b>R4:</b> HH_SEC_J4.dta</p> <p><b>Key Notes:</b> None</p>			

## SECTION K: NON-FOOD EXPENDITURES – PAST ONE WEEK & ONE MONTH

<b>Level of Observation: Unique Identifier(s):</b>	Household UPHI4, round, hk_00 <u>or</u> UPHI4, r_hhid, hk_00	<b>UPQ Data File: Rounds Present:</b>	UPQ4_HH_MOD_K.dta R1, R2, R3, R4
--	--	---	-------------------------------------

**Description:**  
Total expenditure on non-food items during the last week or last month, including: public transportation, fuels, cellular phone credits, personal hygiene items, etc.

**Original Data Files:**

<b>R1:</b> SEC_L.dta	<b>R3:</b> HH_SEC_K.dta
<b>R2:</b> HH_SEC_L.dta	<b>R4:</b> HH_SEC_K.dta

**Key Notes:**  
Expenditure on “Repairs and maintenance to dwelling” in the past one month was only collected in R1 and R2.

## SECTION L: NON-FOOD EXPENDITURES – PAST TWELVE MONTHS

<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Household UPHI4, round, hl_00 <b>or</b> UPHI4, r_hhid, hl_00	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_L.dta R1, R2, R3, R4
--	--	---	-------------------------------------

**Description:**  
Total expenditure on non-food items during the past 12 months, including: household items, community contributions, fees and fines, marriage costs, clothing, etc.

**Original Data Files:**

<b>R1:</b> SEC_L.dta (items 301-317) SEC_M.dta (all other items)	<b>R2:</b> HH_SEC_L.dta <b>R4:</b> HH_SEC_L.dta
<b>R2:</b> HH_SEC_M.dta	

**Key Notes:**  
*For explanation of data derivation only:* the R1 Non-Food Expenditures – Past Twelve Months Module (Module M) included items 301-317 (with two content questions only) in the R1 Module L dataset (**SEC.L.dta**) as that module had the same two content questions; the remaining items (with three content questions) were kept in the Module M dataset (**SEC.M.dta**). In the UPQ dataset, the items were standardized across rounds rather than content questions; thus, the comprehensive list of items in the Non-Food Expenditures – Past Twelve Months Module will be included, and the content questions will be blank for items that are not applicable.

The wording of questions (and thus the variable label) varies nominally between items 301-324 and items 325-326; the former asks “Over the past 12 months, did you purchase or pay for any [...]?” while the latter asks “Over the past 12 months, did you gather, purchase, or pay for any [...]?” These have been combined under the **hl\_01** variable for sake of dataset simplicity. Please see the survey instrument for UPQ Module L for item specific question wording.

Likewise, these have been combined for the **hl\_02** variable as well, where the question for items 301-324 is “How much did you pay in total?” and “What was the estimated total value of [...] consumed?” for items 325-326.

## SECTION M: HOUSEHOLD ASSETS

<b>Level of Observation: Unique Identifier(s):</b>	Household UPHI4, round, hm_00 <b>or</b> UPHI4, r_hhid, hm_00	<b>UPQ Data File: Rounds Present:</b>	UPQ4_HH_MOD_M.dta R1, R2, R3, R4
<b>Description:</b> Quantity of key items owned by the household, age of item, purchase price, and current estimated value for items including, radio, telephones, appliances, furniture, cookware, vehicles, land, and agricultural tools.			
<b>Original Data Files:</b>  <div style="display: flex; justify-content: space-between;"> <span><b>R1:</b> SEC_N.dta</span> <span><b>R3:</b> HH_SEC_M.dta</span> </div> <div style="display: flex; justify-content: space-between;"> <span><b>R2:</b> HH_SEC_N.dta</span> <span><b>R4:</b> HH_SEC_M.dta</span> </div>			
<b>Key Notes:</b> None			

## SECTION N: FAMILY / HOUSEHOLD NON-FARM ENTERPRISES

<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Household UPHI4, round, hn_00 <u>or</u> UPHI4, r_hhid, hn_00	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_N.dta R3, R4
<p><b>Description:</b></p> <p>Details on non-farm businesses operated by the household during the last 12 months, including type of product or service provided, household members involved with the enterprise, value of current stock, operation location, length of time the enterprise has been operating, income earned, profits, business operating costs etc.</p> <p><b>Original Data Files:</b></p> <p><b>R1:</b> Not present                      <b>R3:</b> HH_SEC_N.dta  <b>R2:</b> Not present                      <b>R4:</b> HH_SEC_N.dta</p> <p><b>Key Notes:</b></p> <p>Filter questions for UPQ Module N are included in the <b>Filters.dta</b>. All households are included in the dataset regardless of their filter question to ensure proper household-level merges; however, please use the filter question(s) to determine if data are eligible but missing, or if they are missing because there are no non-farm enterprises and responses are thus not applicable.</p>			

## SECTION 01: ASSISTANCE AND GROUPS

<b>Level of Observation:</b>	Household	<b>UPQ Data File:</b>	UPQ4_HH_MOD_01.dta
<b>Unique Identifier(s):</b>	UPHI4, round, ho_00 <u>or</u> UPHI4, r_hhid, ho_00	<b>Rounds Present:</b>	R1, R2, R3, R4
<p><b>Description:</b>  Financial and in-kind assistance provided by both government and non-government institutions, including free food/maize distribution, food for work or cash for work programmes, inputs for work programmes, and school scholarships.</p>			
<p><b>Original Data Files:</b></p>			
<b>R1:</b> SEC_01.dta	<b>R3:</b> HH_SEC_01.dta		
<b>R2:</b> HH_SEC_01.dta	<b>R4:</b> HH_SEC_01.dta		

SECTION O2: ASSISTANCE AND GROUPS			
<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Household UPHI4, round	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_O2.dta R1, R2, R3, R4
<p><b>Description:</b> Detailed information on credit and savings groups (SACCOS) that household members are members of, including balances, contributions, and repayment time.</p> <p><b>Original Data Files:</b>  <b>R1:</b> SEC_O3.dta                      <b>R3:</b> HH_SEC_O2.dta  <b>R2:</b> HH_SEC_O2.dta                <b>R4:</b> HH_SEC_O2.dta</p> <p><b>Key Notes:</b> None</p>			

SECTION P: CREDIT			
<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Household UPHI4, round, hp_00 <u>or</u> UPHI4, r_hhid, hp_00	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_P.dta R1, R2, R3, R4
<p><b>Description:</b> Utilization of credit and loan institutions, including sources, amounts borrowed, and repayment plans.</p> <p><b>Original Data Files:</b>  <b>R1:</b> SEC_P2.dta                      <b>R3:</b> HH_SEC_P.dta  <b>R2:</b> HH_SEC_P2.dta                <b>R4:</b> HH_SEC_P.dta</p> <p><b>Key Notes:</b> Filter questions found in <b>Filters.dta</b></p>			

SECTION Q1: FINANCE			
<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Household UPHI4, round <u>or</u> UPHI4, r_hhid	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_Q1.dta R2, R3, R4
<p><b>Description:</b> Use of financial institutions and mobile-phone based money transfer services, income from rent or pensions, supplemental income sources, and banking history.</p> <p><b>Original Data Files:</b>  <b>R1:</b> Not present                      <b>R3:</b> HH_SEC_Q1.dta  <b>R2:</b> HH_SEC_Q.dta                <b>R4:</b> HH_SEC_Q1.dta</p> <p><b>Key Notes:</b></p>			

## SECTION Q2: FINANCE

<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Household UPHI4, round, sourceid <u>or</u> UPHI4, r_hhid, sourceid	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_Q2.dta R2, R3, R4
--	--	---	----------------------------------

**Description:**  
Remittances or financial assistance received in the last 12 months from sources residing in Tanzania or abroad, including socio-demographic characteristics of source, remittance channels, and use of cash/in-kind items.

**Original Data Files:**

<b>R1:</b> Not present	<b>R3:</b> HH_SEC_Q2.dta
<b>R2:</b> HH_SEC_Q.dta	<b>R4:</b> HH_SEC_Q2.dta

**Key Notes:**  
R3 & R4 collected Finance information at the household-level (i.e. one data row; dataset: **HH\_SEC\_Q1.dta**), and at the source-level (dataset: **HH\_SEC\_Q2.dta**). While most of the questions align across the three rounds, R2 collected information at household-level only. To enhance universality, R2 has been split into two datasets to mirror R3 & R4.

The resulting dataset **UPQ4\_HH\_MOD\_Q2.dta** will have no more than one source present for R2 households (based solely on questionnaire formatting and not actual source count), while R3 & R4 households may have more than one source as multiple sources were allowed.

## SECTION R: RECENT SHOCKS TO HOUSEHOLD

<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Household (Shock) UPHI4, round, hr_00 <b>or</b> UPHI4, r_hhid, hr_00	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_R.dta R1, R2, R3, R4
--	--	---	-------------------------------------

**Description:**  
Severity ranking of 16 common shocks, such as “drought or floods”, “severe water shortage”, and “death of member of household”. More detailed information and coping strategies are asked of the three most severe shocks.

**Original Data Files:**

<b>R1:</b> SEC_R.dta	<b>R3:</b> HH_SEC_R.dta
<b>R2:</b> HH_SEC_R.dta	<b>R4:</b> HH_SEC_R.dta

**Key Notes:**  
The original R1 Module R dataset only provided data for the specific shock(s) affecting each respective household – i.e. if the answer was “No” to a shock in the shock listing, no data line was recorded (rather than the comprehensive listing present for each household in the R2, R3, & R4 datasets and the first question recording “Yes”/“No”).

Therefore, households in R1 will have any number of data lines between -12 (19 is the most possible, but the most shocks any household reported was 12).

Add data line(s)? – if drop \_m==1, won’t match numbers for other datasets?  
Add data line(s) – Add one line? Add 3 lines? Add 19 lines?  
COME BACK TO THIS ---- ASK!

For this module (and others like it) it is important to consider the *possible* denominator when calculating rates. For example, **hr\_07\_1** had 20 possible response codes in R3, but only 15 in R4. While they maintain the same numeric response code value across years, a tabulation will report “0” for the 5 response codes not presented in R4. These are not truly “0”, rather were not possible in the round-specific questionnaire; thus, the UPQ survey instrument should be consulted prior to all analyses.

## SECTION S: DEATHS IN HOUSEHOLD

<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Household UPHI4, round, hs_00 <b>or</b> UPHI4, r_hhid, hs_00	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_S.dta R1, R2, R3, R4
<p><b>Description:</b> Deaths within the household, cause of death, duration of illness if applicable, and associated land or asset losses due to inheritance traditions.</p> <p><b>Original Data Files:</b></p> <p><b>R1:</b> SEC_S2.dta                      <b>R3:</b> HH_SEC_S.dta  <b>R2:</b> HH_SEC_S.dta                      <b>R4:</b> HH_SEC_S.dta</p> <p><b>Key Notes:</b></p>			

## SECTION V: ANTHROMPOMETRY

<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Individual UPI4, round	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_V.dta R1, R2, R3, R4
<p><b>Description:</b>          Anthropometric information of household members under the age of 15 and/or women of child bearing age (15-49 years old) who were able and willing to participate, including height, weight, and upper arm circumference measurements.</p> <p><b>Original Data Files:</b></p> <p><b>R1:</b> SEC_B_C_D_E1_F_G1_U.dta                      <b>R3:</b> HH_SEC_V.dta  <b>R2:</b> HH_SEC_U.dta                                      <b>R4:</b> HH_SEC_V.dta</p> <p><b>Key Notes:</b></p>			

## SECTION W1: CHILDREN LIVING ELSEWHERE

<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Individual UPI4, round <u>or</u> UPI4, r_hhid <u>or</u> round, r_hhid, r_id	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_W1.dta R1
<p><b>Description:</b> This section includes two filter questions for the remainder of the “Children Living Elsewhere” module – requesting information on female household members age 27 and older, and children age 15 and older living elsewhere.</p> <p><b>Original Data Files:</b></p> <p><b>R1:</b> SEC_B_C_D_E1_F_G1_U.dta                      <b>R3:</b> Not present  <b>R2:</b> Not present    <b>R4:</b> Not present</p> <p><b>Key Notes:</b> None</p>			

## SECTION W2: CHILDREN LIVING ELSEWHERE

<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Individual UPI4, child_id	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_W2.dta R1
<p><b>Description:</b> Details on children (&gt;15 years of age) living outside of the NPS household, including brief information on education, occupation, and remittances. Information in this module is completed by the mother of the child and not the child.</p> <p><b>Original Data Files:</b></p> <p><b>R1:</b> SEC_G2.dta                      <b>R3:</b> Not present  <b>R2:</b> Not present                      <b>R4:</b> Not present</p> <p><b>Key Notes:</b> Module W2 only contains information for women greater &gt;27 years of age who have children &gt;15 years of age living outside of the household. While individual-level module merges with <b>UPQ4_HH_MOD_W1.dta</b> will result in perfect merges, <b>UPQ4_HH_MOD_W2.dta</b> is a supplementary dataset resulting from the filter questions in Module W1.</p>			

## SECTION X1: GOVERNANCE

<b>Level of Observation: Unique Identifier(s):</b>	Household UPHI4, round <u>or</u> UPHI4, r_hhid	<b>UPQ Data File: Rounds Present:</b>	UPQ4_HH_MOD_X1.dta R1, R2
<p><b>Description:</b> This section collects general information on the primary respondent for the governance section.</p> <p><b>Original Data Files:</b>  <b>R1:</b> SEC_H1_J_K2_O2_P1_Q1_S1.dta      <b>R3:</b> Not present  <b>R2:</b> HH_SEC_H3.dta                      <b>R4:</b> Not present                HH_SEC_FILTERS.dta</p> <p><b>Key Notes:</b> None</p>			

## SECTION X2: GOVERNANCE

<b>Level of Observation:</b>	Household	<b>UPQ Data File:</b>	UPQ4_HH_MOD_X2.dta
<b>Unique Identifier(s):</b>	UPI4, round, hx2_00 <u>or</u> UPI4, r_hhid, hx2_00	<b>Rounds Present:</b>	R1, R2
<b>Description:</b> Details on knowledge of government officials (if known by respondent), including approval.			
<b>Original Data Files:</b>			
<b>R1:</b> SEC_H2.dta	<b>R3:</b> Not present		
<b>R2:</b> HH_SEC_H1.dta	<b>R4:</b> Not present		
<b>Key Notes:</b>			
None			

## SECTION X3: GOVERNANCE

<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Household UPI4, round, hx3_00 <u>or</u> UPI4, r_hhid, hx3_00	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_X3.dta R1, R2
<p><b>Description:</b> Information on the primary respondent's attendance at local governance meetings, religious affiliation, and participation in religious activities.</p> <p><b>Original Data Files:</b></p> <p><b>R1:</b> SEC_H3.dta                      <b>R3:</b> Not present  <b>R2:</b> HH_SEC_H2.dta                <b>R4:</b> Not present</p> <p><b>Key Notes:</b> None</p>			

## SECTION Y: VIOLENCE AGAINST WOMEN

<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Household (Individual) UPHI4, indidy1	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_Y.dta R1
<p><b>Description:</b> Attitudes about violence against women and the types of violence experienced.</p> <p><b>Original Data Files:</b></p> <p><b>R1:</b> SEC_I.dta                      <b>R3:</b> Not present  <b>R2:</b> Not present                      <b>R4:</b> Not present</p> <p><b>Key Notes:</b>  This module was only asked of women ages 15-50.</p>			

## SECTION Z1: CRIME & JUSTICE

<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Household UPHI	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_Z1.dta R1
<p><b>Description:</b> General reporting on if any crimes or disputes were suffered by the household during the last twelve months (see Section Z2) associated interactions with the police and justice system (see Section Z3).</p> <p><b>Original Data Files:</b>  <b>R1:</b> SEC_H1_J_K2_O2_P1_Q1_S.dta      <b>R3:</b> Not present  <b>R2:</b> Not present      <b>R4:</b> Not present</p> <p><b>Key Notes:</b> To properly allow for merges back to other UPQ datasets, the UPHI4 variable has been included and thus, there are duplicates at the <b>y1_hhid</b> level (because of the various panel combinations using a single R1 household). The UPHI4 variable will be the unique identifier for the module but analyses will most likely occur at the <b>y1_hhid</b> level.</p>			

SECTION Z2: CRIME & JUSTICE			
<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Household UPHI4, hz_00	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_Z2.dta R1
<p><b>Description:</b> Information on crimes or disputes suffered by the household during the last twelve months (Section Z2) and associated interactions with the police and justice system (see Section Z3).</p> <p><b>Original Data Files:</b>  R1: SEC_Q2.dta                      R3: Not present  R2: Not present                      R4: Not present</p> <p><b>Key Notes:</b> None</p>			

SECTION Z3: CRIME & JUSTICE			
<b>Level of Observation:</b> <b>Unique Identifier(s):</b>	Household UPHI4, hz_00	<b>UPQ Data File:</b> <b>Rounds Present:</b>	UPQ4_HH_MOD_Z3.dta R1
<p><b>Description:</b> Information on crimes or disputes suffered by the household during the last twelve months (see Section Z2) and associated interactions with the police and justice system (Section Z3).</p> <p><b>Original Data Files:</b>  R1: SEC_Q3.dta                      R3: Not present  R2: Not present                      R4: Not present</p> <p><b>Key Notes:</b> None</p>			

## VIII. APPENDIX

### A. APPENDIX I: NPS UPQ Naming Convention

R1		R2		R3		R4		UPQ	
A	Household Identification & Survey Staff Details	A	Household Identification & Survey Staff Details	A	Household Identification & Survey Staff Details	A	Household Identification & Survey Staff Details	A	Household Identification & Survey Staff Details
B	Household Member Roster	B	Household Member Roster	B	Household Member Roster	B	Household Member Roster	B	Household Member Roster
C	Education	C	Education	C	Education	C	Education	C	Education
D	Health	D	Health	D	Health	D	Health	D	Health
E	Labor	E	Labor	E	Labor	E	Labor	E	Labor
F	Food Outside the Household	F	Food Outside the Household	F	Food Consumption Outside the Household	F	Food Consumption Outside the Household	F	Food Consumption Outside the Household
		G	Subjective Welfare	G	Subjective Welfare	G	Subjective Welfare & Crime	G	Subjective Welfare & Crime
		I	Food Security	H	Food Security	H	Food Security	H	Food Security
J	Housing, Water & Sanitation	J	Housing, Water & Sanitation	I	Housing, Water & Sanitation	I	Housing, Water & Sanitation	I	Housing, Water & Sanitation
K	Consumption of Food Over Past One Week	K	Consumption of Food Over Past One Week	J	Consumption of Food Over Past One Week	J	Consumption of Food Over Past One Week	J	Consumption of Food Over Past One Week
L	Non-Food Expenditures	L	Non-Food Expenditures	K	Non-Food Expenditures	K	Non-Food Expenditures	K	Non-Food Expenditures
M	Non-Food Expenditures	M	Non-Food Expenditures	L	Non-Food Expenditures	L	Non-Food Expenditures	L	Non-Food Expenditures
N	Household Assets	N	Household Assets	M	Household Assets	M	Household Assets	M	Household Assets
				N	Family/Household Non-Farm Enterprises	N	Family/Household Non-Farm Enterprises	N	Family/Household Non-Farm Enterprises
O	Assistance and Groups	O	Assistance and Groups	O	Assistance and Groups	O	Assistance and Groups	O	Assistance and Groups
P	Credit	P	Credit	P	Credit	P	Credit	P	Credit
		Q	Finance	Q	Finance	Q	Finance	Q	Finance
R	Recent Shocks to Household Welfare	R	Recent Shocks to Household Welfare	R	Recent Shocks to Household Welfare	R	Recent Shocks to Household Welfare	R	Recent Shocks to Household Welfare
S	Deaths in Household	S	Deaths in Household	S	Deaths in Household	S	Deaths in Household	S	Deaths in Household
V	Household Re-Contact Information	V	Household Re-Contact Information	U	Household Re-Contact Information	U	Household Re-Contact Information	U	Household Re-Contact Information
U	Anthropometry	U	Anthropometry	V	Anthropometry	V	Anthropometry	V	Anthropometry
G	Children Living Elsewhere							W	Children Living Elsewhere
H	Governance	H	Governance					X	Governance
I	Violence Against Women							Y	Violence Against Women
Q	Crime & Justice							Z	Crime & Justice

## B. APPENDIX II: Questionnaire Footnotes

<b>MODULE B</b>
<p><sup>1</sup> R1 question: "IF THIS IS AN HBS HOUSEHOLD MEMBER, ENTER THE HBS ID NUMBER"</p> <p><sup>2</sup> R1 lists "UNPAID FAMILY WORK" twice; values/labels in R1 dataset indicate code 12 is "PAID FAMILY WORK" as in other rounds.</p> <p><sup>3</sup> R1 specified "ENTER 0 IF NONE"; removed from UPQ.</p> <p><sup>4</sup> "DISTRICT/COUNTRY NAME" and "REGION CODE" are included on R1 questionnaire, but variables are not included in R1 datasets so no R1 variable name is presented here.</p>
<b>MODULE C</b>
<p><sup>1</sup> The UPQ question has been altered to be applicable to all rounds. See round-specific questionnaires for exact wording.</p> <p><sup>2</sup> "ONE WAY ONLY" only specified in R3 &amp; R4 questionnaires.</p>
<b>MODULE D</b>
<p><sup>1</sup> In R1, respondents 12 years and older were instructed to answer for themselves, so R1 questions are worded with respect to the individual answering the question, i.e. "Have you visited a health care provider...". "[NAME]" was used in R2, R3, and R4 and thus is the format used in the text of the UPQ question. Only the first instance is noted but this applies to all questions in the R1 questionnaire.</p> <p><sup>2</sup> R2 question: "During the past 12 months, what measures were taken to improve [NAME]'s performance of activities?"</p> <p><sup>3</sup> "SPECIFY" was added to response code "OTHER" beginning in R2. The numeric response code (=7) remains the same across rounds.</p> <p><sup>4</sup> R1 &amp; R2 question: "Was this birth registered?"</p>
<b>MODULE E</b>
<p><sup>1</sup> In R2, the questions were worded in respect to the individual answering the question, i.e. "Did you do any work of any type...". "[NAME]" was used in R1, R3, and R4 and thus is the format used in the text of the universal panel question. Only the first instance is noted but this applies to the remainder of questions in the R2 questionnaire.</p> <p><sup>2</sup> The UPQ question has been altered to be applicable to both R3 &amp; R4. See round-specific questionnaires for exact wording.</p> <p><sup>3</sup> The instruction "(WAGE EMPLOYMENT)" is only specified in the R1 questionnaire; however, in each of R1 and R2 this question is included in the "Wage Jobs" section.</p> <p><sup>4</sup> See round-specific questionnaires for slight variations in wording; differences not expected to alter interpretation.</p> <p><sup>5</sup> The instruction "(if any)" is only specified in the R1 questionnaire, while the instruction "[ENTER '0' / '0' IF NEVER]" is only in R2.</p> <p><sup>6</sup> The instruction "[CODE: TASCO CODE]" is only present on the R3 &amp; R4 questionnaires.</p> <p><sup>7</sup> The instruction "[APART FROM SALARY]" is only specified in the R1, R3, &amp; R4 questionnaires.</p> <p><sup>8</sup> Maximum amounts of months, weeks, and hours are only specified on the R3 &amp; R4 questionnaires.</p> <p><sup>9</sup> Both "[wage]" and "[paid]" are specified in the R3 question, but are only in the section header in the R4 questionnaire.</p> <p><sup>10</sup> The R2 &amp; R3 questionnaires specify "wage" in the question, and in the section header in the R4 questionnaire.</p> <p><sup>11</sup> Questions in the self-employment section of the R1 questionnaire used the pronoun "you"/"your" instead of "[NAME]", similar to the language used throughout the R2 questionnaire. The text of UPQ questions with this footnote is identical to the text in the questions in both R1 and R2 and is meant only to highlight the deviation from the language used throughout the remainder of the R1 questionnaire.</p> <p><sup>12</sup> The instruction "IF NONE, WRITE '0'" is only in the R2 questionnaire.</p> <p><sup>13</sup> The instruction "[IF NONE, WRITE '0']" is only specified in R1 &amp; R2.</p> <p><sup>14</sup> Maximum amount of hours is only specified in R3 &amp; R4.</p> <p><sup>15</sup> Note: There is a question number 49 in Module E of R1. This question "seq49_1"- "seq49_6" will be found in UPQ Module G (as UPQ question hg_03_1-hg_03_9) as they align with the Subjective Welfare &amp; Crime question from R2, R3, &amp; R4.</p>
<b>MODULE F</b>
No footnotes
<b>MODULE G</b>
<sup>1</sup> Response (=8) was "DOES NOT APPLY" in R1 and "NOT APPLICABLE" in R2, R3, and R4; the UPQ uses "NOT APPLICABLE"
<b>MODULE H</b>
No footnotes

<b><u>MODULE I1</u></b>
<sup>1</sup> While both R1 and R2 report two sources, the instruction "UP TO 2 SOURCES" was only specified on the R2 questionnaire.
<sup>2</sup> "SPECIFY" was added to the "OTHER" response code starting in R2. The numeric response code value (=6) remains the same across rounds.
<sup>3</sup> R1 question: "What is the main toilet facilities for this household?" & R2 question: "What is the main toilet facilities usually used in this household?"
<sup>4</sup> "SPECIFY" was added to the "OTHER" response code starting in R2.
<sup>5</sup> R3: "How many other households share this toilet facility?"
<sup>6</sup> SPECIFY" was added to the "OTHER" response code starting in R2. The numeric response code value (=8) remains the same.
<sup>7</sup> "SPECIFY" was added to the "OTHER" response code starting in R2. The numeric response code value (=7) remains the same.
<sup>8</sup> Referred to as "RAINWATER" in R1-R3.
<sup>9</sup> "SPECIFY" was added to the "OTHER" response code starting in R2. Numeric response code values changed over time.
<sup>10</sup> R2 & R3 Qx: "How long does it take to get water from drinking water source to this dwelling in the rainy season?";
<sup>11</sup> R2 & R3 Qx: "Out of these [READ] minutes, how long do you spend waiting?"
<sup>12</sup> R2 & R3 Qx: "How long does it take to get water from drinking water source to this dwelling in the dry season?";
<sup>13</sup> R2 & R3 Qx: "Out of these [READ] minutes, how long do you spend waiting?"
<sup>14</sup> "BOIL" (R2 & R3) was referred to as "BOILING WATER" in R1. The numeric response code value remains the same; 15"MARK UP TO 2" was only specified in R2 & R3.
<sup>15</sup> "MARK UP TO 2" was only specified in R2 & R3.
<b><u>MODULE J1</u></b>
No footnotes
<b><u>MODULE J3</u></b>
No footnotes
<b><u>MODULE J4</u></b>
No footnotes
<b><u>MODULE K</u></b>
<sup>1</sup> "Mosque" was added to code 218 in R3 and R4
<b><u>MODULE L</u></b>
No footnotes
<b><u>MODULE M</u></b>
No footnotes
<b><u>MODULE N</u></b>
No footnotes
<b><u>MODULE O1</u></b>
No footnotes
<b><u>MODULE O2</u></b>
<sup>1</sup> The letter ID was not included in the R1 datasets and thus does not have a variable name like the other rounds.
<b><u>MODULE P</u></b>
<sup>1</sup> "in the last 12 months" was not included in the R2 or R1 question.
<sup>2</sup> The instruction "LIST UP TO THREE IN ORDER OF IMPORTANCE" was not specified in R2.
<b><u>MODULE Q1</u></b>
<sup>1</sup> R2 question: Which of the following remittances channels did you use in the last 12 months?

<b><u>MODULE Q2</u></b>
<sup>1</sup> R2 question: <i>What is your relationship to the sender?</i>
<sup>2</sup> R2 question: <i>"From what country did these remittances originate?"</i>
<sup>3</sup> R2 question: <i>Which of the following remittances channels did you use in the last 12 months?</i>
<sup>4</sup> R2 question: <i>When you receive the money from abroad what did you use it for?</i>
<b><u>MODULE R</u></b>
<sup>1</sup> Updated from "2" to "UP TO 3" to reflect maximum number allowed by rounds (=3), R3.
<b><u>MODULE S</u></b>
<sup>1</sup> The phrase "including those listed as "dead" in PRE-PRINTED TRACKING FORM" was added in R2.
<b><u>MODULE V</u></b>
No footnotes
<b><u>MODULE W1</u></b>
No footnotes
<b><u>MODULE W2</u></b>
<sup>1</sup> Some code values and value labels have been altered slightly between the R1 and UPQ versions (see "NEVER ATTENDED", F1-F4, and U2-U3).
<sup>2</sup> R1 codes differ between R1 questionnaire and R1 dataset; the UPQ has made the best effort to standardize and correct these.
<b><u>MODULE X1</u></b>
<sup>1</sup> The instruction "FROM THE PRE-PRINTED TRACKING FORM" was only specified in the R2 question.
<b><u>MODULE X2</u></b>
<sup>2</sup> The instruction "IF POST IS VACANT, WRITE '00'" was only specified in the R2 question.
<b><u>MODULE X3</u></b>
No footnotes
<b><u>MODULE Y</u></b>
No footnotes
<b><u>MODULE Z1</u></b>
No footnotes
<b><u>MODULE Z2</u></b>
<sup>1</sup> UPQ question wording changed from "above" to "in previous questions".
<b><u>MODULE Z3</u></b>
<sup>2</sup> UPQ question wording changed from "above" to "in previous questions".