NATIONAL BUREAU OF STATISTICS

Basic Information Document

National Panel Survey (2020-2021)

November, 2022

2020-2021

[UNITED REPUBLIC OF TANZANIA]

ACRONYMS

BHPS	-	British Household Panel Survey
BID	-	Basic Information Document
CAPI	-	Computer-Assisted Personal Interviewing
DEFF	-	Design Effect
EA	-	Enumeration Area
GPS	-	Global Positioning System
HHID	-	Household Identification
ILO	-	International Labour Organisation
ISIC	-	International Standard Industry Codes
JMP	-	Joint Monitoring Plan
LSMS-ISA	-	Living Standards Measurement Study-Integrated Surveys on Agriculture
MKUKUTA	-	National Strategy for Growth and Reduction of Poverty
MMMP	-	Monitoring Master Plan
NBS	-	Tanzania National Bureau of Statistics
NGO	-	Non-Governmental Organization
NPS	-	Tanzania National Panel Survey
PSID	-	Panel Study of Income Dynamics
SACCO	-	Savings and Credit Cooperative Organization
SDD	-	Sex Disaggregated Data
SE	-	Standard Error
TASCO	-	Tanzania Standard Classification of Occupation
UNICEF	-	The United Nations Children's Fund
VEO	-	Village Executive Officer
WHO	-	The World Health Organization

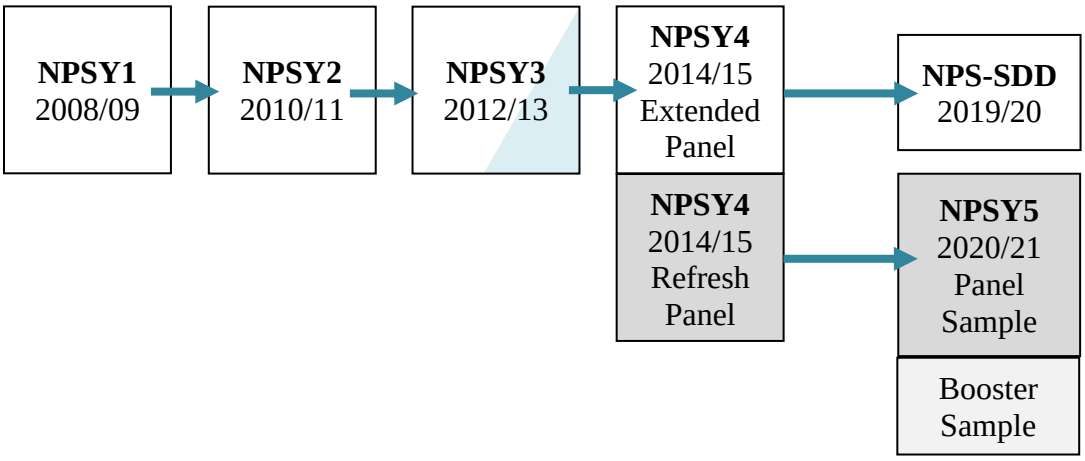
Table of Contents

Background 4
Survey Instruments 5
Sample Design 8
Implementation 8
Panel Attrition 11
Weighting / Expansion Factors 12
Data Set 13
Obtaining Data 14
Table 1: Household Questionnaire 15
Table 2: Agriculture Questionnaire 25
Table 3: Livestock Questionnaire 30
Table 4: Community Questionnaire 33
Appendix A: Calculation of Panel and Cross Section Weights 36
Appendix B: Confidential Information 43
Appendix C: Consumption Aggregate 44
1.0 The construction of the consumption aggregate 44
2.0 Price adjustment 47
3.0 Household composition adjustment 50
Appendix D: Tanzanian Educational System 52
Appendix E: Description of the Tanzania Standard Classification of Occupation (TASCO) 54
Appendix F: List of International Standard Industry Codes (ISIC) 70
Appendix G: Main Animal Diseases & Key Terms 74
Appendix H: Animal Breeding Types 76
Appendix I: List of Definitions 77
Appendix J: Anthropometry 78

Background

The National Panel Survey (NPS) was originally launched in Tanzania in 2008, with support from the Living Standards Measurement Survey – Integrated Surveys on Agriculture [LSMS-ISA[1]] program at the World Bank and other donors. The first round of the survey was conducted in 2008/09, the second round in 2010/11, the third round in 2012/13, and the fourth round in 2014/15. The 2020/21 NPS is the fifth round in 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 socioeconomic characteristics. All five rounds of the NPS have been implemented by the Tanzania National Bureau of Statistics (NBS).

In the fourth round (2014/15), the NPS sample consisted of a combination of the original NPS sample and a new sample. A nationally representative sub-sample (approximately 20 percent) was selected from the third round to continue as the original "Extended Panel" sample, while an entirely new sample, "Refresh Panel", was selected to also represent national and sub-national domains. An off-shoot survey, the NPS-SDD 2019/20 with sex-disaggregated data ("NPS-SDD 2019/20"), was undertaken by following the *entire* NPS 2014/15 "Extended Panel" sample. The NPS 2020/21 – the fifth round of the National Panel Survey – on the other hand followed the entire "Refresh Panel", in addition to including a set of new households considered an urban oversample, the "Booster Sample" group. The true panel sample in NPSY5 can therefore only be traced back to those NPS 2014/15 Refresh Panel households that were found and reinterviewed in the NPS 2020/21.



[1] The Living Standards Measurement Study (LSMS) is an ongoing research initiative within the Development Research Group of the World Bank with the goal of promoting and improving the collection of household level data in developing countries around the world. Further information can be found at www.worldbank.org/lsms. The LSMS-ISA project supports governments in seven Sub-Saharan African countries to generate nationally representative, household panel data with a strong focus on agriculture and rural development. Further information can be found at www.worldbank.org/lsms-isa.

The main objective of the NPS 2020/21 is to provide high-quality household-level data to the Tanzanian government and other stakeholders for monitoring poverty dynamics, tracking the progress of national development plans (including the Five-Year Development Plan (FYDP) II poverty reduction strategy[2] and its predecessor plans), and to evaluate the impact of other major, national-level government policy initiatives. As an integrated survey covering a number of different socioeconomic indicators, it compliments other more narrowly focused survey efforts, such as the Demographic and Health Survey (DHS) 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). Secondly, 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, as well as the linkages between agriculture and household welfare.

The Tanzania NBS was advised on technical issues related to survey design and implementation by the NPS Technical Committee, which included representatives from line ministries, such as the Ministry of Agriculture, Food Security and Cooperatives, and the Ministry of Finance. Funding for the NPS 2020/21 was provided by the European Union through the World Bank.

This document describes fundamental aspects of the NPS 2020/21, including the set of survey instruments, sample design, survey implementation, and the resulting data sets.

Survey Instruments

The NPS 2020/21 consists of four survey instruments: a Household Questionnaire, Agriculture Questionnaire, Livestock Questionnaire, and a Community Questionnaire.

[2] The FYDP II began in 2016/17 and finished in 2020/21, and integrated the frameworks of predecessor strategies, FYDP I and the National Strategy for Growth and Reduction of Poverty (MKUKUTA). The plan is the government strategy to meet the Millennium Development Goals (MDG) and other national development goals, with a current focus on growth and transformation alongside a reduction in poverty.

The Household Questionnaire comprises thematic sections. A detailed description of the contents of the questionnaire can be found in Box 1 of this report, including inclusion/exclusion requirements and level of data collection. This questionnaire format allows for the construction of a full consumption based welfare measure, permitting both distributional and incidence analysis. Data within the household questionnaire are primarily structured at the household-level, though data on education, health, labour, food outside of the household, welfare, and anthropometrics are collected at the individual-level, lending to even greater specificity on the characteristics of household units. This additional level of information will add value in the analysis of intra-household dynamics and revealing a more refined picture of welfare of Tanzania. To protect the confidentiality of respondents, sensitive information has been masked in or removed from the public household data files.

BOX 1: HOUSEHOLD QUESTIONNAIRE:

Section A:	Household Identification
Section B:	Household Member Roster
Section C:	Education
Section D:	Health
Section E:	Labour
Section F:	Food Consumed Outside the Household
Section G:	Subjective Welfare and Crime
Section H:	Food Security
Section I:	Housing, Water, and Sanitation
Section I2:	Handwashing
Section J:	Consumption of Food Over Past One Week
Section K:	Non-Food Expenditures – One Week/One Month
Section L:	Non-Food Expenditures – Twelve Months
Section M:	Household Assets
Section N:	Family / Household Non-Farm Enterprises
Section O:	Assistance and Groups
Section P:	Credit
Section Q:	Finance
Section R:	Recent Shocks to Household Welfare
Section S:	Deaths in Household
Section U:	Household Re-contact Information / Filters
Section V:	Anthropometry
Section WQT:	Water Quality Testing

The NPS 2020/21 also includes a robust instrument on agricultural activities (detailed in Box 2). It offers an essential data source to understand the dynamic role of agriculture to household welfare. Information on agriculture (inputs, production, and sales) is collected at both the plot and crop level, consistent with key phases in the agricultural value chain.

BOX 2: AGRICULTURE QUESTIONNAIRE:

Section 02:	Plot Roster
Section 03:	Plot Details
Section 04:	Annual Crops By Plot
Section 05:	Annual Crop Production and Sales
Section 06:	Permanent Crops By Plot
Section 07:	Permanent Crops Production and Sales
Section 10:	Processed Agricultural Products & By-Products
Section 11:	Farm Implements and Machinery
Section 12:	Extension

The NPS 2020/21 questionnaire likewise recognizes the importance of livestock activities to many households in Tanzania. As with the integrated survey instrument on agriculture activities, the NPS contains a robust instrument to capture details on these activities. The Livestock Questionnaire is administered to all households participating in these activities and asks about the inputs, outputs, labour, and sales related to these activities. Box 3 provides a more comprehensive list of the sections found within the Livestock Questionnaire.

BOX 3: LIVESTOCK QUESTIONNAIRE:

Section 02:	Livestock Stock Ownership
Section 03:	Animal Health
Section 04:	Feed, Water, Housing, Breeding
Section 4A:	Livestock Extension services
Section 05:	Livestock Labour
Section 06:	Milk
Section 07:	Animal Power and Dung
Section 08:	Other Livestock Products

The Community Questionnaire collects information on physical and economic infrastructure and events in surveyed communities[3]. Responses to this questionnaire are provided through group discussion among key informants within the community.

BOX 4: COMMUNITY QUESTIONNAIRE:

Section CA:	Community Identification
Section CB:	Access to Basic Services
Section CC:	Investment Projects
Section CD:	Land Use
Section CE:	Demographics, Land, and Livestock
Section CF:	Market Prices
Section CG:	Local Units

Each of the questionnaires were developed in collaboration with line ministries and donor partners, including the Technical Committee, over a period of several months. The NBS solicited feedback from various stakeholders regarding the survey content and design, paying due consideration to comparability with previous panel rounds.

It is important to note that question numbers might be non-sequential at times in the survey instruments because the questions in a section were not renumbered when questions were deleted from that section between the fourth and fifth round. The same question and section numbers were maintained from earlier rounds for consistency. Additionally, questions that were added in between rounds were given a letter suffix to distinguish them (i.e. 75a, 75b, etc.).

[3] Note that this is not a "community" in the sociological sense, but rather a mechanism to collect information about the areas where the households selected for the survey are located. In most rural areas, enumeration areas (EAs) are defined by village boundaries and therefore community refers to the village. In urban areas the distinction is less clear, and occasionally single community questionnaires were administered to adjoining EAs. Therefore, the number of EAs and community questionnaires will not be identical.

Sample Design

The NPS is based on a stratified, multi-stage cluster sample design which recognizes four analytical strata: Dar es Salaam, Other Urban areas in Mainland, Rural areas in Mainland, and Zanzibar. The sample design for the NPS 2020/21 targeted the sub-sample of households from the initial NPS 2014/15 cohort considered the "Refresh Panel". These specific households had never previously been a part of the NPS sample design. This sample consisted of 3,352 households from 419 clusters in the NPS 2014/15 that were tracked and interviewed in the NPS 2020/21. An additional "Booster Sample" of 545 households from major cities and urban areas (specifically, Mbeya, Arusha, Mwanza, Tanga, and Dodoma) was also interviewed to allow for improved estimates in urban centres.

In previous NPS rounds, the sample design included complete households that could not be interviewed in a particular year but were found in later rounds, *excluding* those households that had refused to be interviewed (i.e. a household that was interviewed in Round 1, lost in Round 2, and found again in Round 3). This situation does not exist in the NPS 2020/21 as they have only been included in, at most, two rounds.

The eligibility requirement for inclusion of a household in this round of the NPS and all others is defined as any household having at least one member aged 15 years and above, excluding live-in servants. Households with at least one eligible member were completely interviewed, including any non-eligible members present in the household.

Additionally, the final sample for NPS 2020/21 included any split-off household or eligible members identified during data collection (i.e. a previous NPS member who had moved or started another household in between rounds). Marriage and migration are the most common reasons for households splitting over time. Ultimately, the final sample size for NPS 2020/21 was 23,592 individuals in 4,709 households. Of these, 4,164 households allow for panel analysis as they have been found and interviewed in both NPS 2014/15 and NPS 2020/21, while the remaining 545 (in the "Booster Sample") will only have data available in the NPS 2020/21. The complete cohort interviewed in NPS 2020/21 will be maintained and tracked in all future waves of the NPS.

Implementation

Preparations:

The main training for the field team supervisors and enumerators took place in Iringa region in November-December, 2020 over a period of thirty (30) days. During this time, the field staff spent one week on the Household Questionnaire, a week for Agricultural Questionnaire, and a week on the Livestock Questionnaire, and tracking forms. The last week of the training was devoted to field

practice. Field practice were conducted in Iringa regions, where selected households from the previous rounds of the NPS were revisited to provide the team enumerators and supervisors practice with conducting tracking during fieldwork. After the fieldwork, extensive discussion and revisions were conducted with the participation of all team supervisors.

Over the training period, three tests were administered to the field teams. The goal was to gain feedback from the training sessions and to select the qualified enumerators. Overall, there were 54 enumerator candidates, with all the enumerators being selected. Interviewer manuals were developed with detailed instructions for field staff during training and as the main reference guide for the survey over the course of the fieldwork. At the end of the training, the enumerators were each provided with an interviewer manual in Kiswahili.

Field Work:

Fieldwork for the NPS 2020/21 began in the last week of December 2020 and was completed in January 2022. This 14-month span of fieldwork addresses any intra-year seasonality concerns and fluctuations. The survey was primarily implemented by eight mobile field teams, each composed of: one supervisor, seven enumerators, and one driver. Seven mobile field teams were responsible for different regions on the mainland and one team was responsible for all of Zanzibar.

Field teams visited each cluster for between 2-3 days for both panel and booster clusters. The questionnaires were administered to the selected households over the course of that time. This allowed the field team to make return visits to the household to complete the entire Household Questionnaire, Agriculture Questionnaire for farming households, and Livestock Questionnaire for households engaged in livestock activities. To ensure the depth and quality of each section of the survey, the questionnaire was administered across multiple respondents and to members most knowledgeable about each topic. For all the sampled tracking households, areas of all owned and/or cultivated agricultural plots were measured via GPS unless the household refused, the terrain was too difficult, or if the plot was more than one (1) hour drive from the location of the household. Anthropometric measurements were taken for all eligible individuals that were at home, not too ill, and willing to participate.

***Tracking*:**

Tracking consisted of three stages: (1) determining the current status of the households and members of the households from the previous survey; (2) determining the location of any households that have moved together with all its members to a new location; and (3) determining the location of any eligible members who have moved from their household from either or all rounds of the NPS to a new location. Eligibility for tracking of households and members was determined prior to the implementation of the survey. Members eligible for tracking were those

over the age of 15 and who are not a live-in house servant. Any members not meeting this criterion was not tracked and were only included in the data when located and interviewed with an eligible member. Furthermore, households eligible for tracking must have moved to a new location *within* Tanzania. If a household moved to a different country, it was not tracked.

If an entire household had moved from the original residence, teams were required to complete a "T-1" form designed to capture relevant information from key informants on the whereabouts of the household. The T-1 form contains information that would enable tracking of household to its new location. If a member or members of the household have moved from the original household, a "T-2" form was completed by the teams. Similar to the T-1 form, a T-2 form contains information on the location of the member(s) who have moved from the household. Once the tracking targets had been located, teams were required to interview the household as consistent with the eligibility requirements.

Within the tracking protocol, there are local and distance tracking cases. Local and distance tracking applies to both households and individuals. Local tracking is defined as cases where the tracking target is within one hour driving distance from the original cluster and at least one tracking member from the household is eligible for tracking. For local tracking cases, the teams are required to interview the tracking target before leaving the original cluster. Distance tracking occurs when the tracking target resides in a location that is more than one hour drive from the original cluster. In this case, the teams fill out the appropriate tracking form and send the information to NBS headquarters. Once at NBS headquarters, the distance tracking case is given to one of the field teams based on the location of the new residence of the household or individual, this assigned field team is then responsible for locating that household and conducting the interview.

Often households and members who have moved from their previous dwelling or household have relocated a great distance from their previous interview location. Given the inefficiency in searching for these members on a case-by-case basis across the country as they occur, field teams were not required to track households and members deemed to be too far away from the original location. Information on these cases was entered and sent electronically to the NBS headquarters, for compilation and review. After review, the tracking cases were assigned to the field teams based on the location of the new residence of the household or individual.

Distance tracking was done after the completion of the main data collection to allow enough time to accumulate a sufficient number of distance tracking targets. Furthermore, regular field teams would perform distance tracking whenever they had a break in their schedule. Three dedicated teams of four enumerators conducted the distance tracking fieldwork.

Data Processing & Management:

The NPS 2020/21 contains a robust multi-level quality assurance and data management system.

Great effort was placed on the development and utilization of this system by the NBS with technical assistance form the World Bank prior to the implementation of the NPS 2020/21 to assist in the management of the complex household panel survey and address the growing need for high quality timely data.

The NPS 2020/21 utilized a Survey Solutions Computer Assisted Personal Interviewing (CAPI) system developed by the Word Bank in collaboration with IT personnel from the NBS. This system was selected to increase the availability of data for review by managing staff as well as to provide regular and consistent quality assessment of data directly to the field staff. STATA software was utilized to perform complex aggregated checks. Furthermore, due to the panel nature of the survey, where applicable and appropriate, data was checked against previous round data.

Data files from completed clusters were transmitted to NBS headquarters via syncing to the Survey Solutions server using 3G USB modems. Received data files were downloaded at the headquarters, and regular checks performed to ensure the fieldwork was proceeding according to the schedule and that quality standards were met. During the course of field work, data was routinely checked at the aggregate level to identify any potential issues and where identified, additional checks were integrated into the CAPI and Stata systems.

Throughout the data processing, system versions of the data are archived at all key steps and all checking and cleaning syntax documented and archived.[4]

Panel Attrition

As with most panel surveys, a certain portion of panel respondents are not able to be re-interviewed over time. This attrition of panel respondents can lead to attrition bias where respondents drop out of the survey non-randomly and where the attrition is correlated with variables of interest. The Tanzania NPS has fortunately maintained low attrition over rounds (less than 4 percent for each of the first three rounds), thus minimizing the potential for attrition bias within the datasets. The attrition rate for the extended panel in NPS 2014/15 remained low, at just 4.9 percent in the 2019/20 NPS.

By the end of data collection for the fifth round of the NPS, 3,052 of the original 3,352 NPS 2014/15 “Refresh Panel” households had been located and successfully re-interviewed for a total household attrition rate of 9.1 percent.

[4] Due to the presence of confidential information within the datasets and the syntax used to process that data, this data and information will not be made available to the public.

Weighting / Expansion Factors

Assignment of urban and rural status for calculation of panel weights

Each round of the NPS requires an assessment of whether each household is in an urban or rural location, both for the calculation of panel weights and any subsequent analysis. The assignment of this location status is done by the NBS with assistance from the World Bank using geo-spatial location variables, mass location assignments, and other available location data.

In the NPS 2020/21, the location status designation for all households that remained in the same locations between 2014/15 and 2020/21 ("non-movers") remained the same as it had been in NPS 2014/15. Households that had moved from their original location to Dar es Salaam were automatically assigned an urban status as Dar es Salaam is a purely urban strata and therefore requires no distinctions for households in that location.

For households that had moved away from their original location to regions other than Dar es Salaam, geo-spatial analysis was used to determine the distance moved, and when possible an assessment was done of the location status of other nearby households. Unlike previous rounds of the NPS which employed *distance* tracking, the tracking protocol for NPS 2020/21 specified that households were only tracked when known to be in close proximity to their original location; thus the previous round's assignment of urban or rural was unlikely to change despite minor movement in location.

Weighting and expansion factors

In order to produce nationally representative statistics with the NPS data, it is necessary to apply weighting or expansion factors. For previous rounds of the NPS, panel survey weights adjust for differences in the probability of selection into the NPS 2008/09 sample for observations in various strata, 2008/09 households splitting into multiple households in NPS 2010/11 and NPS 2012/13, splitting even further in NPS 2014/15, and attrition between rounds of the survey.[5] A similar practice was used for panel survey weights in NPS 2020/21, based on their probability of selection in the NPS 2014/15 (as the Refresh Sample).

The first round of the NPS sample was a multi-stage clustered sample design. First stage sampling involved the selection of survey clusters with the probability of selection proportional to cluster size within a stratum. The sampling of these clusters was stratified along two dimensions: (i) eight administrative zones (seven on Mainland Tanzania plus Zanzibar as an eighth zone), and (ii) rural versus urban clusters within each administrative zone. The combination of these two dimensions

[5] The details of the sample design – including the sampling strata and the use of multiple sampling frames – are discussed in a separate document, *Sample Design for the National Panel Survey*, April 2009, available from NBS upon request.

yields 16 strata. In rural areas a cluster is defined as an entire village. In urban areas, a cluster is defined as a census enumeration area. As a general rule, the probability of selection was higher for clusters within strata where existing data sources showed that the variance of key variables of interest for the NPS (e.g., household consumption and maize production) were likely to be very high – implying the need for more observations to produce reliable estimates.

The expansion factors for the NPS 2020/21, variables ***y5_panelweight*** and ***y5_crossweight,*** have been integrated into Section A in the household dataset ("HH_SEC_A.dta"). The former is only available for the set of households that were a part of the NPS 2014/15 sample, while the latter provides a weight for the full cross-section of households collected in NPS 2020/21, regardless of whether they are a part of the panel or booster sample. Additionally, unique identifiers for the first-stage sampling units, ***clusterid***, and for the sampling strata, ***strataid*** can also be found in the Section A in the household datasets.

The methodology used to calculate the panel weights for the panel households in NPS 2020/21 was developed as part of the LSMS-ISA work program. Details on the methodology can be found in the paper: Himelein, Kristen. 2013. "Weight Calculations for Panel Surveys with Subsampling and Split-off Tracking." Statistics and Public Policy, vol (1), pp40-45 and in Appendix A.

Data Set

The NPS consists of several data files. Each data file pertains to a section of the questionnaire, or a set of sections that are for the same level of observation. The complete lists of data files as well as the unique identification variables in each module are listed in Tables 1- 4.

Unique Household Identifiers:

Households are identified by a nine-digit number, ***y5_hhid***, in the NPS 2020/21. This is generated using the combination of the seven-digit household identification number in NPS 2014/15 plus a two-digit number. The last two digits of the NPS 2020/21 household identifier is the lowest individual ID number from the NPS 2014/15 round out of all the members of that household that were reinterviewed in the NPS 2020/21. For example, an NPS 2014/15 household has the household identification number of "0218-001" and consisted of 6 members with individual ID numbers from 1-6. If in the NPS 2020/21, members 2, 3 and 4 remained in the household, then this household will have the unique identifier of "0218-001-02".

Merging between Round 4 data sets:

The household identification variable is ***y5_hhid*** in the data files. When merging or linking individual data files, it is necessary to use at minimum the household identification variable (***y5_hhid***) as well as the other unique identification variables pertinent to the data files being

merged. All household level sections have been harmonized, regardless of a household's or member's qualification for the section, to provide a consistent number of observations in each module across level-like units of observation.

Additional variables required to merge data across sections depend on which two sections are being merged. For example, merging Sections A and Section I (basic identification and household characteristics, respectively) require only the ***y5_hhid*** variable as it is the single unique identifier in both datasets. Merging Sections B and C (roster information with education levels, respectively) requires merging on both the ***y5_hhid*** and the individual's ID, ***indidy5***[6]. This combination of variables will be unique in the roster data file, as only one person per household can have a particular roster number. Similar patterns will be found in other combinations of data files, such as the plot number ***plotnum*** and crop ID ***cropid*** variables in the agricultural datasets.

Merging between the two rounds:

NPS 2020/21 households can also be linked back to their information in the previous round of the NPS (2014/15). To merge household-level datasets between the NPS 2014/15 and NPS 2020/21 it is necessary to use the ***y4_hhid*** variable found in the data for both rounds.

Merging of individuals in panel households across the NPS 2014/15 and NPS 2020/21 is also possible. It is important to note that due to the movement of household members, and particularly the joining of household members from one panel household to another, the Section A household identifier from the previous round does not always reflect the previous household of each member. To define each member's previous year's household, the ***y4_hhid*** variable reflecting each specific member's previous round household has been added to the Section B household roster. To facilitate accurate merging of panel members between rounds, a supplementary dataset, "NPSY5.PANEL.KEY", has also been provided. This dataset contains identification information for individuals, where applicable, across NPS 2014/15 and NPS 2020/21 an individual has been present.

Obtaining Data

Data and background documentation for the NPS 2020/21 is available on the Tanzania National Bureau of Statistics website (www.nbs.go.tz), as well as the World Bank Living Standards Measurement Study – Integrated Surveys on Agriculture website (https://www.worldbank.org/en/programs/lsms), and may be downloaded free of charge. Inquiries pertaining to the data may be sent to the LSMS team at lsms@worldbank.org.

[6] It is important to note that while the *y5_hhid* variable is constructed using individual IDs from Round 4, a new and unique individual identification number for each individual (both old and new) is provided regardless of their inclusion in Round 4 (*indidy5*).

Table 1: Household Questionnaire

SECTION A : HOUSEHOLD IDENTIFICATION			
Level of Observation :	Household	**Data File :**	HH_SEC_A
Unique Identifier :	y5_hhid	**Additional Info :**	

Description :
Household location variables, unique within panel round household identification variables, date and time of interview, analytic sampling weights, cluster identification, sampling strata identification, and status of survey.

Key Notes :
The variable **loc_info** distinguishes whether or not a household was in the same physical location as the previous round, NPSY4. For split-off households, this variable will be equal to "NO". For original households, if the answer is "NO", then new household location information will be provided. If the answer is "YES", then the household has confirmed they are in the same location and no new location information will be provided. For ease of use, these variables have been collated to a single set of location variables, regardless of a households' movement over time.

All sensitive identifying variables, such as names of the ward, village, mtaa, household head, tracking target, and the written description of the household location have been removed from the dissemination version of this dataset to preserve the confidentiality of the respondent.

SECTION B : HOUSEHOLD MEMBER ROSTER			
Level of Observation:	Individual	**Data File :**	HH_SEC_B
Unique Identifier :	y5_hhid, indidy5	**Additional Info :**	

Description :
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.

Key Notes :
Two sets of unique individual identifiers are present in the household roster; the *within* round (NPSY5) individual unique identifier and the *across* round (NPSY4/NPSY5) individual unique identifier. The within round individual unique identifier is composed of the unique NPSY5 nine-digit household identification, **y5_hhid**, and the roster row individual identification number, **indidy5**. The across round unique identifier is comprised of the seven digit NPSY4 unique household identifier, **y4_hhid**, and the previous round's roster row number, **hh_b06.** Both individual identifiers are present due to the rare instances where members from different panel households have merged into one household.

Brand new members, previously absent in any panel household, are indicated by a value of "99" in the **hh_b06**. If a respondent from NPSY4 moved to *a different* NPS household in NPSY5, their **hh_b06** will indicate their NPSY4 roster ID, and their **y4_hhid** will differ from the rest of the members in their new household as it links them to their original round four household.

The respondent's name (hh_b01) and date of birth (hh_b03_1, hh_b03_2), have been masked to protect respondent confidentiality. Age of children under 5 years in months can be found in the supplemental dataset, "NPSY5.CHILD.ANTHRO".

SECTION C : EDUCATION

Level of Observation :	Individual	**Data File :**	HH_SEC_C
Unique Identifier :	y5_hhid, indidy5	**Additional Info :**	(>=5 years) ; Appendix D

Description :
Educational enrolment and attainment; literacy, education history/level attainment, current enrolment, school characteristics, temporary absence, education examination (participation and score) and education expenditures.

Key Notes :
Section is administered to members 5 years of age or older.

See Appendix D for additional terminology and information on changes to the Tanzanian educational system potentially affecting comparability of educational attainment of household members born before or around 1970.

SECTION D : HEALTH

Level of Observation :	Individual	**Data File :**	HH_SEC_D
Unique Identifier :	y5_hhid, indidy5	**Additional Info :**	

Description :
General health status and utilization of health services; source and financing of health treatments / hospitalization, disaggregated health expenditures, disability, bednet use, pregnancy, prenatal care and births, child health and ailments / diarrhoea.

Key Notes :
Respondents less than 12 years of age answer for themselves; else an informed respondent provides information.

Questions on disability were included in NPSY2, NPSY4, NPS-SDD, and in NPSY5 but *not* in NPSY3.
Questions on pregnancy are asked only to women 14 through 49 years of age.
Question on child instance of diarrhoea were asked for members less than 5 years of age.

SECTION E : LABOUR

Level of Observation :	Individual	**Data File :**	HH_SEC_E1
Unique Identifier :	y5_hhid, indidy5	**Additional Info :**	(>= 5 years); Appendix E; F

Description :
Labour market participation; activities including unpaid apprenticeship, wage work, non-farm enterprise, and agricultural activity (including livestock activities) in the last seven days. Also includes unemployment and steps taken to find work, secondary activities, industry (TASCO) and occupation (ISIC) codes (Appendix E and Appendix F), wages earned, contract type, employer provided safety nets, and general domestic activities.

Key Notes :
Section is administered to household members 5 years of age or older.

Recorded written descriptions of industry and occupation have been masked to protect respondent confidentiality. These text fields are removed for confidentiality purposes, as well as the text names of the employing companies.

SECTION E2 : LABOUR (USUAL ACTIVITY)

Level of Observation :	Individual	**Data File :**	HH_SEC_E2
Unique Identifier :	y5_hhid, indidy5	**Additional Info :**	(>= 5 years); last 12 months

Description :
Usual labour activity during the past twelve months, including unpaid apprenticeships, wage work, non-farm enterprise, and agricultural activity (including livestock activities).

Key Notes :
Section is administered to household members 5 years of age or older.

SECTION E3 : OWN USE PRODUCTION OF GOODS

Level of Observation :	Individual	**Data File :**	HH_SEC_E3
Unique Identifier :	y5_hhid, indidy5	**Additional Info :**	(>= 5 years); past 7 days

Description :
Hunting, gathering, food preparation/preservation, household good crafting, water and firewood collection activities done by respondents in the last seven days and how much time has been spent on these activities.

Key Notes :
Section is administered to household members 5 years of age or older.

SECTION F : FOOD CONSUMED OUTSIDE THE HOUSEHOLD

Level of Observation :	Individual	**Data File :**	HH_SEC_F
Unique Identifier :	y5_hhid, indidy5	**Additional Info :**	Last 7 days

Description :
Value of food consumed outside the home during the last seven days.

Key Notes :
None

SECTION G : SUBJECTIVE WELFARE AND CRIME

Level of Observation :	Individual	**Data File :**	HH_SEC_G
Unique Identifier :	y5_hhid, indidy5	**Additional Info :**	(>=15 years)

Description :
Self-reported level of satisfaction with health, financial status, housing, job, services, and safety. Also includes perceived status at present, and as of two years ago. This section also collects information about any crime that the household may have been a victim of in the past twelve months and any associated police involvement.

Key Notes :
None

SECTION H : FOOD SECURITY

Level of Observation :	Household	**Data File :**	HH_SEC_H
Unique Identifier :	y5_hhid	**Additional Info :**	

Description :
Information on the household's diet and food intake patterns, months of food insecurity, and distribution of food within the household.

Key Notes :
None

SECTION I : HOUSING, WATER & SANITATION

Level of Observation :	Household	**Data File :**	HH_SEC_I
Unique Identifier :	y5_hhid	**Additional Info :**	

Description :
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.

Key Notes :
Major source for the questions on water and sanitation was the set of harmonized questions on drinking-water and sanitation for household surveys developed by the WHO/UNICEF Joint Monitoring Program (JMP) for Water Supply and Sanitation.

SECTION I2 : HANDWASHING

Level of Observation :	Household	**Data File :**	HH_SEC_I2
Unique Identifier :	y5_hhid	**Additional Info :**	

Description :
Information on handwashing practices of the household, including location, presence of water, and presence of soap, asked to the head of household.

Key Notes :
This module was introduced for the first time in NPS-SDD 2019/20, and for the second time in the NPSY5. There will be no panel information on handwashing available in NPSY4.

SECTION J1 : CONSUMPTION OF FOOD OVER THE PAST ONE WEEK

Level of Observation :	Food item	**Data File :**	HH_SEC_J1
Unique Identifier :	y5_hhid, itemcode	**Additional Info :**	Last 7 days

Description :
Consumption of 60 key items/ item groups over the last 7 days. These 60 items/item groups are divided into twelve broader categories; “Cereals and Cereal Products”, “Starches, Sugars and Sweets”, “Pulses, Dry”, “Nuts and Seeds”, “Vegetables”, “Fruits”, “Meat, Meat Products, Fish”, “Milk and Milk Products”, Oils and Fats”, “Spices and Other Foods”, and “Beverages”. Quantity and value of consumed items from purchases, own-production and gifts are recorded.

Key Notes :
Unlike previous rounds, no data was collected in the NPSY5 on the presence of iodine. Due to the incorporation of experiment into NPSY5, this module was administered to about 419 households.

Previous rounds of the NPS have had 59 items. In NPSY4, item code 108 (“wheat flour, barley grain and other cereals”) was split into two items: item code 1081 (“wheat flour”) and item code 1082 (“barley grain and other cereals”), for a resulting 60 items. The same codes remained in NPSY5.

SECTION JA1 : CONSUMPTION OF FOOD OVER THE PAST ONE WEEK

Level of Observation :	Food item	**Data File :**	HH_SEC_J1
Unique Identifier :	y5_hhid, itemcode	**Additional Info :**	Last 7 days

Description :
This is basically the same as that of SEC J1, except that this module included an expanded list of Non-Standard Units (NSUs). The number of food items was also expanded to 75 over the last 7 days. These 75 items/item groups are divided into twelve broader categories; “Cereals and Cereal Products”, “Starches, Sugars and Sweets”, “Pulses, Dry”, “Nuts and Seeds”, “Vegetables”, “Fruits”, “Meat, Meat Products, Fish”, “Milk and Milk Products”, Oils and Fats”, “Spices and Other Foods”, and “Beverages”. Quantity and value of consumed items from purchases, own-production and gifts are recorded.

Key Notes :
Unlike previous rounds, no data was collected in the NPSY5 on the presence of iodine.

Previous rounds of the NPS have had 59 items. In NPSY4, item code 108 (“wheat flour, barley grain and other cereals”) was split into two items: item code 1081 (“wheat flour”) and item code 1082 (“barley grain and other cereals”), for a resulting 60 items. The same codes remained in NPSY5.

SECTION J3 : FREQUENCY OF CONSUMED ITEMS

Level of Observation :	Food item category	**Data File :**	HH_SEC_J3
Unique Identifier :	y5_hhid, itemcode	**Additional Info :**	Last 7 days

Description :
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”.

Key Notes :
The categories/items composing the food categories in this section may differ from those in Section J1.

SECTION J4 : QUANTITY OF MEALS AND PEOPLE CONSUMING MEALS

Level of Observation :	Age group	**Data File :**	HH_SEC_J4
Unique Identifier :	y5_hhid, agegroup	**Additional Info :**	Last 7 days

Description :
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.

Key Notes :
None

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

Level of Observation :	Item	**Data File :**	HH_SEC_K
Unique Identifier :	y5_hhid, itemcode	**Additional Info :**	Last 7 days / Last 1 month

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.

Key Notes : None

SECTION L : NON-FOOD EXPENDITURES – PAST TWELVE MONTHS

Level of Observation :	Item	**Data File :**	HH_SEC_L
Unique Identifier :	y5_hhid, itemcode	**Additional Info :**	Last 12 months

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

Key Notes :
“Wood poles, bamboo” and “Grass for thatching roof or other use” are commonly consumed by households but not purchased. For these items, estimated total value is also recorded.

SECTION M : HOUSEHOLD ASSETS

Level of Observation :	Item	**Data File :**	HH_SEC_M
Unique Identifier :	y5_hhid, itemcode	**Additional Info :**	Last 12 months

Description :
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.

Key Notes :
Age of item, purchase price, and estimated value are new additions to the Household Assets section beginning in NPSY3. Age of item, purchase price and estimated value of item were not solicited for all items. Refer to the Household Questionnaire Section M for list of items with this restriction.

SECTION N : FAMILY / HOUSEHOLD NON-FARM ENTERPRISES

Level of Observation :	Enterprise	**Data File :**	HH_SEC_N
Unique Identifier :	y5_hhid, enterprise_id	**Additional Info :**	

Description :
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.

Key Notes :
Only households operating non-farm enterprises completed this module. The filter question for this module is **hh_n01a**, and though it is asked at the household level, it is included in this data file.

SECTION O1 : ASSISTANCE AND GROUPS

Level of Observation :	Assistance type	**Data File :**	HH_SEC_O1
Unique Identifier :	y5_hhid, assistance_id	**Additional Info :**	

Description :
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.

Key Notes :
The names of the organizations have been masked to protect respondent confidentiality.

SECTION O2 : ASSISTANCE AND GROUPS

Level of Observation :	Individual	**Data File :**	HH_SEC_O2
Unique Identifier :	y5_hhid, indidy5	**Additional Info :**	

Description :
Involvement in cooperative microfinance activities and financial details of membership, including contributions, withdrawal history, and repayment plans.

Key Notes :
Participating members in assistance groups will be distinguished by their **indidy5,** which can also be used to link to other individual level sections.

SECTION P : CREDIT

Level of Observation :	Loan	**Data File :**	HH_SEC_P
Unique Identifier :	y5_hhid, loanid	**Additional Info :**	

Description :
Utilization of credit and loan institutions, including sources, amounts borrowed, and repayment plans.

Key Notes :
Only households that borrowed cash, goods, or services in the last 12 months completed this module, though each household is included here to facilitate accurate merging between household-level modules. The filter question for this module is **hh_p01**, and though it was asked to at the household level, it is included in this data file.

Individual characteristics can be linked with household members who utilized loan/credit sources using the reported individual ID in Q4. The name of the lending institution is masked due to confidentiality purposes.

SECTION Q1 : FINANCE

Level of Observation :	Household	**Data File :**	HH_SEC_Q1
Unique Identifier :	y5_hhid	**Additional Info :**	

Description :
Use of financial institutions and mobile-phone based money transfer services, income from rent or pensions, supplemental income sources, and banking history.

Key Notes :
Names of banking institutions where households maintained accounts have been masked to protect confidentiality.

SECTION Q2 : FINANCE

Level of Observation :	Source of finance	**Data File :**	HH_SEC_Q2
Unique Identifier :	y5_hhid, sourceid	**Additional Info :**	

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.

Key Notes :
The names of the sources have been masked to protect respondent confidentiality.

SECTION R : RECENT SHOCKS TO HOUSEHOLD

Level of Observation :	Shock	**Data File :**	HH_SEC_R
Unique Identifier :	y5_hhid, shock_id	**Additional Info :**	

Description :
Severity ranking of 17 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.

Key Notes :
Shocks that consistently had very low level of occurrence in previous NPS rounds were removed from the list of shocks. Previous rounds of NPS had 19 shocks.

SECTION S : DEATHS IN HOUSEHOLD

Level of Observation :	Individual	**Data File :**	HH_SEC_S
Unique Identifier :	y5_hhid, personid	**Additional Info :**	

Description :
Deaths within the household, cause of death, duration of illness if applicable, and associated land or asset losses due to inheritance traditions.

Key Notes :
The name of the deceased household member is excluded for confidentiality purposes.

The **personid** variable is a within-module, constructed variable that only serves as a unique identifier for HH_SEC_S. It is not a linking variable with any other individual-level module.

The across rounds unique identifier is comprised of the seven-digit NPSY4 unique household identifier, **y4_hhid**, and the previous round's roster row number, **hh_s04_2**. This variable can be linked in the same way that **hh_b06** in Section B is used to link living members. Note that not all deceased household members will have been present in the fourth round as some may have both joined the household and died in the interim between survey rounds.

SECTION V : ANTHROPOMETRY

Level of Observation :	Individual	**Data File :**	HH_SEC_V
Unique Identifier :	y5_hhid, indidy5	**Additional Info :**	

Description :
Anthropometric information of household members under the age of 15 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.

Key Notes :
Unlike in previous NPS rounds where this section was administered to all members, in the NPSY5 this section was administered to only members that were under the age of 15 or women of child bearing age (15-49 years old).

Upper arm circumference measurements are only obtained for children less than 5 years of age.

CONSUMPTION AGGREGATES – 2020/21

Level of Observation :	Individual	**Data File :**	consumption_real_y5
Unique Identifier :	y5_hhid	**Additional Info :**	

Description :
Consumption aggregates for NPSY5, spatially adjusted in prices of survey year, 2020/21. In addition to food and non-food, it also includes rent and durables

Key Notes :
The methodology used for the NPS 2020/21 is identical to the methodology used for previous rounds of the NPS, except that the NPS 2020/21 consumption aggregates include rent and durables. Given that the NPSY5 is a panel of the refresh sample from 2014/15, we also reconstructed the aggregates for NPS 2014/15 to include rent and durables to make it comparable to the NPS 2020/21. Thus, any comparative welfare analysis should use the new consumption aggregates, consumption_real_y4 and consumption_real_y5.

CONSUMPTION AGGREGATES – 2014/15

Level of Observation :	Individual	**Data File :**	consumption_real_y4
Unique Identifier :	y4_hhid	**Additional Info :**	

Description :
Consumption aggregates for NPSY4, spatially adjusted in prices of survey year, 2014/15. In addition to food and non-food, it also includes rent and durables.

ANTHROPOMETRIC INDICATORS

Level of Observation :	Individual	**Data File :**	NPSY5.CHILD.ANTHRO
Unique Identifier :	y5_hhid, indidy5	**Additional Info :**	Appendix J

Description:
A dataset containing Z-scores for the anthropometric indicators “weight-for-age”, “height-for-age”, and “weight-for-height”. A binary variable indicating severity of malnourishment (<-2 SDs) is also provided for each indicator.

Key Notes: The unique individual identifier is composed of the unique NPSY5 household identification, **y5_hhid**, and the roster row individual identification number, **indidy5**. This is the same unique individual identifier combination that exists in the individual-level household data files.

While physical body measurements were collected in section “HH_SEC_V” for all household members under 15 and women of child-bearing age (15-49 years), Z-scores for anthropometric indices are only provided for children 0-59 months.

See Appendix J for a detailed description of calculation and application of anthropometric indices.

Table 2: Agriculture Questionnaire

All households that answered Question 1, 2, 3, *or* 4 in Section U2 as "YES" should appear in this module. This means that a household cultivated a plot(s) or owned a plot(s) that they did not cultivate, at some point in the last 12 months. The cover page is not included in the dataset because they do not contain any additional information beyond what is included in the household questionnaire cover page.

Please note that Section 8 for "Input Vouchers" and Section 9 "Outgrower Schemes" were removed in the Agriculture questionnaire in NPS 2014/15 and remain removed for NPS 2020/21. However, the numbering of the sections in the questionnaire have remained the same for consistency across NPS rounds.

SECTION 01 : LAND PLOT ROSTER			
Level of Observation :	Plot	**Data File :**	AG_SEC_01
Unique Identifier :	y5_hhid, plot_id	**Additional Info :**	

Description :
Plot roster of all previously owned and new plots owned by family, including location, tenure system, and use.

Key Notes :
In previous rounds, this data file has been called HH_SEC_B1B.dta.

SECTION 02 : PLOT ROSTER			
Level of Observation :	Plot	**Data File :**	AG_SEC_02
Unique Identifier :	y5_hhid, plot_id	**Additional Info :**	

Description :
Roster of all plots owned or cultivated by the household, including measurement information as calculated by GPS and farmer's estimate, GPS coordinates, weather conditions at measurement, and reason for missing GPS.

Key Notes :
If applicable, the NPSY4 plot ID of a NPS 2020/21 plot is recorded in **ag2a_05**, in a similar fashion to **hh_b06** in Household Module B which links a NPSY5 member to NPSY4. This will need to be combined with the **y4_hhid** variable found in Module A.

The plot areas were measured using Garmin eTrex HC series GPS devices. The survey protocol indicates that all plots should be measured as long as they were with one hour's transportation (either on foot, by bicycle / motorbike, or, if possible, by vehicle) from the household. The only other acceptable reason for the plot not to be measure would be if the household refused.

The plot names, descriptions, and GPS coordinates are excluded for confidentiality reasons.

SECTION 3A/3B : PLOT DETAILS

Level of Observation :	Plot	**Data File :**	AG_SEC_3A/3B
Unique Identifier :	y5_hhid, plot_id	**Additional Info :**	

Description :
Detailed information on usage of plot, main cultivated crops, decision-makers in household, soil type and quality with a focus on erosion, sources of irrigation, ownership status of plot, rental value, usage patterns of fertilizers, and agricultural inputs obtained on credit. Household and hired labour for farming activities is also reported.

Key Notes :
Identical questions are asked in "AG_SEC_3A" and "AG_SEC_3B", with the exception of Q75-85 which gather brief information on their rainy season counterpart's plot usage. Questions in 3A again refer to plots owned or cultivated during the last long rainy season, while 3B refers to the last completed short rainy season.

SECTION 4A/4B : ANNUAL CROPS BY PLOT

Level of Observation :	Plot-crop	**Data File :**	AG_SEC_4A/4B
Unique Identifier :	y5_hhid, plot_id, cropid	**Additional Info :**	

Description :
Crop planting patterns, intercropping, area and quantity harvested, associated losses, crop seeds purchased along with associated values, source and type of seed for all annual crops.

Key Notes :
"AG_SEC_4A" is asked of the long rainy season, while "AG_SEC_4B" is asked of the most recent short rainy season (dependent on date of interview). This section should only include annual crops, but some permanent and fruit crops were also captured.

SECTION 5A/5B : ANNUAL CROP PRODUCTION AND SALES

Level of Observation :	Crop	**Data File :**	AG_SEC_5A/5B
Unique Identifier :	y5_hhid, cropid	**Additional Info :**	

Description :
Questions on quantity of crops sold, value of sales, customers crops sold to, average distance that crops were transported to for sale, post-harvest losses, how crop residue was handled, method and duration for which crop was stored.

Key Notes :
"AG_SEC_5A" is asked of the long rainy season, while "AG_SEC_5B" is asked of the most recent short rainy season (dependent on date of interview).

In some cases crops produced by the household were not represented in the list of response codes. In these cases crops were classified as "Other" with a corresponding code.

SECTION 6A : FRUIT CROPS BY PLOT

Level of Observation :	Plot-crop	**Data File :**	AG_SEC_6A
Unique Identifier :	y5_hhid, plot_id, cropid	**Additional Info :**	

Description :
Number of fruit trees planted on the plot, when these were planted, presence of intercropping, quantity produced, loss before and after harvest, quantity sold, associated value and location sold, method and quantity of crop stored are asked in this section.

Key Notes :
This section should include only fruit trees but due to a small number of interviewer errors, some fruits are included in Section 4 instead while some permanent crops are included in this section. Note that in the English version of the questionnaire, both peaches and plums appear twice in the crop listing. This is due to the fact that some fruits common in Tanzania have the same English translation, while the names are different in Swahili.

Note that this section should include only fruits but due to a small number of interviewer errors, some fruits are included in Section 4 instead and some permanent crops are also included in this section.

SECTION 6B : PERMANENT CROPS BY PLOT

Level of Observation :	Plot-crop	**Data File :**	AG_SEC_6B
Unique Identifier :	y5_hhid, plot_id, cropid	**Additional Info :**	

Description :
Number of permanent crops planted on the plot, when these were planted, how many were planted in the past 12 months, intercropping activities, quantity produced, losses before and after harvest, quantity sold, associated value and location sold, method and quantity of crop stored are asked in this section.

Key Notes :
Questions in “AG_SEC_6A” and “AG_SEC_6B” are identical, however 6A is asked only of fruit trees while 6B is asked of permanent trees/crops.

SECTION 7A : FRUIT CROP PRODUCTION AND SALES

Level of Observation :	Crop	**Data File :**	AG_SEC_7A
Unique Identifier :	y5_hhid, cropid	**Additional Info :**	

Description :
Quantity of crop sold, associated value and location sold, post production losses and method and quantity of crop stored are included.

Key Notes :
None

SECTION 7B : PERMANENT CROP PRODUCTION AND SALES

Level of Observation :	Crop	**Data File :**	AG_SEC_7B
Unique Identifier :	y5_hhid, cropid	**Additional Info :**	

Description :
Quantity of crop sold, associated value and location sold, post production losses and method and quantity of crop stored are included.

Key Notes :
Although 7B was intended to be a direct continuation of Section 6B and therefore include all crops listed in 6B, crops that were used specifically for own consumption were often not reported. Consequently, Section 7B has fewer observations than what is reported in Section 6B.

SECTION 10 : PROCESSED AGRICULTURAL PRODUCTS AND BY-PRODUCTS

Level of Observation :	Crop	**Data File :**	AG_SEC_10
Unique Identifier :	y5_hhid, cropid	**Additional Info :**	

Description :
Information on crops, by-product names and quantity produced, amount of crop used as input, quantity sold, associated prices and buyers and costs incurred due to labor/other inputs are included in this section.

Key Notes :
Crops listed in this section should have been harvested by the household, and therefore should also appear in Section 5 and Section 7. However, due to interviewer error, some of the crops found in this section (primarily maize) were most likely purchased instead of harvested, and therefore will not appear in Sections 5 and 7.

SECTION 11 : FARM IMPLEMENTS AND MACHINERY

Level of Observation :	Item	**Data File :**	AG_SEC_11
Unique Identifier :	y5_hhid, itemid	**Additional Info :**	

Description :
Detailed information on the number of farm implements and machinery used or owned by the household in the past 12 months along with associated value if sold, whether the item was used, reasons for no usage, whether any of these items were rented or borrowed for use in the last twelve months and associated rents paid.

Key Notes :
None

SECTION 12A : EXTENSION SERVICES

Level of Observation :	Extension service	**Data File :**	AG_SEC_12A
Unique Identifier :	y5_hhid, sourceid	**Additional Info :**	

Description :
Any extension services or advice that the household received for agricultural or livestock activities in the past 12 months through government extension, NGOs, Cooperative/Farmer's Association, or Large Scale Farmers, including what activity advice was sought for, subjective rating for advice received, and price paid for receiving advice.

Key Notes :
None

SECTION 12B : EXTENSION SERVICES

Level of Observation :	Extension service	**Data File :**	AG_SEC_12B
Unique Identifier :	y5_hhid, sourceid	**Additional Info :**	

Description :
Any extension services or advice that the household received for agricultural or livestock activities in the past 12 months through government extension, NGOs, Cooperative/Farmer's Association, Large Scale Farmers, Radio/television, Publications or neighbours including what activity advice was sought for, subjective rating for advice received, and price paid for receiving advice.

Key Notes :
Note that in both datasets in Section 12 (A and B), the variables to uniquely identify observations are **y5_hhid** and **sourceid**. However, these two variables are not equivalent and it is not possible to merge the two datasets using these variables.

Table 3: Livestock Questionnaire

Sections 2-8 of the Livestock Questionnaire are asked of all households that answered Question 5 in Section U2 as "YES". Unlike previous rounds of the NPS, there was no data collection on Fisheries. The cover page is not included in the dataset because they do not contain any additional information beyond what is included in the household questionnaire cover page.

Note that Section 2 and Sections 3-7 ask for information about a household's livestock at different levels. Section 2 asks for information at the level of individual animal types (i.e. cows, chickens, etc.) while Sections 3-7 ask information at the broader level of aggregated animal types (i.e. large ruminants, small ruminants, etc.).

SECTION 02 : LIVESTOCK STOCK			
Level of Observation :	Livestock animal	**Data File :**	LF_SEC_02
Unique Identifier :	y5_hhid, lvstckid	**Additional Info :**	

Description :
For 16 animal categories (including "Other"), this section asks questions on topics related to ownership, purchases, gifts received, diseases and animals lost because of them, thefts, sales and associated earnings, slaughtering and associated earning in the last 12 months.

Key Notes :
None

SECTION 03 : ANIMAL HEALTH			
Level of Observation :	Livestock group	**Data File :**	LF_SEC_03
Unique Identifier :	y5_hhid, lvstckid	**Additional Info :**	Appendix G

Description :
Detailed information on the health of animals in the last 12 months related to diseases affecting the animals, vaccinations provided, de-worming treatments, preventative and treatment measures taken against tick borne diseases, and associated spending for any preventative and treatment measures.

Key Notes :
Questions in this section are only asked of large and small ruminants, pigs and poultry. Additional information on main animal diseases for which vaccination is available in Appendix G.

SECTION 04 : FEED, WATER, HOUSING, BREEDING

Level of Observation :	Livestock group	**Data File :**	LF_SEC_04
Unique Identifier :	y5_hhid, lvstckid	**Additional Info :**	Appendix H

Description :
This section asks about fodder and water costs for the animals, type of housing used for the animals, and breeding strategies used by the household.

Key Notes :
Questions in this section are only asked of large and small ruminants, pigs and poultry. Additional information on breeding strategies may be found in Appendix H.

SECTION 05 : LIVESTOCK LABOUR

Level of Observation :	Livestock group	**Data File :**	LF_SEC_05
Unique Identifier :	y5_hhid, lvstckid	**Additional Info :**	

Description :
This section gathers information about the household members responsible for activities associated with upkeep of animals, the months in which the household hired help (if any) to assist in the upkeep of animals, and any associated costs with hired labor.

Key Notes :
Questions in this section are only asked of large and small ruminants.

SECTION 06 : MILK

Level of Observation :	Livestock group	**Data File :**	LF_SEC_06
Unique Identifier :	y5_hhid, lvstckid	**Additional Info :**	

Description :
This section asks about characteristics of milk production by the animals and the amount of milk produced that were consumed by the household, sold, or processed in the last 12 months.

Key Notes :
None.

SECTION 07 : ANIMAL POWER AND DUNG

Level of Observation :	Livestock group	**Data File :**	LF_SEC_07
Unique Identifier :	y5_hhid, lvstckid	**Additional Info :**	

Description :
Detailed information related to amount of dung produced by the animals, characteristics of the use of the dung, associated earnings from any sales of dung, and use of animals for transport or ploughing.

Key Notes :
Questions in this section are only asked of large and small ruminants.

SECTION 08 : OTHER LIVESTOCK PRODUCTS

Level of Observation :	Product	**Data File :**	LF_SEC_08
Unique Identifier :	y5_hhid, productid	**Additional Info :**	

Description :
Information is collected on livestock by-products, including the quantity produced, the quantity sold, value of sold goods, buyers of by product sold.

Key Notes :
None

Table 4: Community Questionnaire

For purposes of this survey, a "community" is defined as the village in rural areas and the mtaa in which the cluster is located in urban areas. The community questionnaire was administered to a group of local leaders determined by the field supervisors. In general, in rural areas this group included the ward executive officer, village chairperson and the village executive officer (VEO), as well as other members from the village council. In urban areas the group included the ward executive officer, mtaa chairperson and possibly other local leaders. Note that not all sample clusters have a corresponding community questionnaire. Particularly in urban areas, clusters within the same ward share the same administration and therefore community level information.

In addition, individuals that moved to new communities would not have corresponding community information as this information was only collected for the originally selected EAs.

SECTION A : COMMUNITY IDENTIFICATION			
Level of Observation :	Community	**Data File :**	CM_SEC_A
Unique Identifier :	id_01 id_02 id_03 id_04 id_05 *or interview__key*	**Additional Info :**	

Description :
Community identification information including region, district, ward, regional capital identifier, location of market price information, and enumeration area, as well as survey staff information such as interviewers ID code, supervisor, and direct observation questions.

Key Notes :
Sensitive identifying variables, such as name of village, GPS coordinates, and the names of the interviewer, supervisor, and data entry operators have been removed from the dissemination version of this dataset to preserve confidentiality.

The unique identifier for Section A can be either the combination of location variables (region + district + ward + ea + mtaa, or the single, data-entry generated identifier **interview__key**. Please note that this is not comparable to the interview__key variable provided in household-level and individual-level datasets.

SECTION B : ACCESS TO BASIC SERVICES			
Level of Observation :	Basic service	**Data File :**	CM_SEC_B
Unique Identifier :	interview__key service_id	**Additional Info :**	

Description :
Information on access to basic services in terms of distance and associated transportation costs for these services.

Key Notes :
Names of services/institutions have been dropped from the public data file to preserve confidentiality.

SECTION C : INVESTMENT PROJECTS

Level of Observation :	Project	**Data File :**	CM_SEC_C
Unique Identifier :	interview__key project_id	**Additional Info :**	

Description :
Sources of funds and associated amounts for recent construction projects such as road construction/maintenance, market construction/maintenance, water supply such as wells and pumps, school construction and maintenance at pre-primary, primary and secondary levels, health and veterinary services, irrigation schemes and grain storages.

Key Notes :
None

SECTION D : LAND USE

Level of Observation :	Community	**Data File :**	CM_SEC_D
Unique Identifier :	interview__key	**Additional Info :**	

Description :
Land use related issues with estimated percentages of how different types of village land are used (cultivation, forest, pasture, wetland, residential, business).

Key Notes :
Occasionally key informants did not know all the information asked in the questionnaire.

SECTION D2 : MAJOR LAND RELATED EVENTS

Level of Observation :	Event	**Data File :**	CM_SEC_D2
Unique Identifier :	interview__key, event_id	**Additional Info :**	

Description :
Information on major land related events, including appropriation of land for outside investors, district or central government declaring land as “reserve land”, and allocation of land cultivated/inhabited by villagers for public use, as well as reasons for re-allocation of land (if any), number of households affected, and associated compensation.

Key Notes :
None

SECTION E : DEMOGRAPHICS, LAND, AND LIVESTOCK

Level of Observation :	Community	**Data File :**	CM_SEC_E
Unique Identifier :	interview__key	**Additional Info :**	

Description :
Participants utilizing SACCOs and other farmer cooperative groups, activities undertaken by cooperatives, the timing and quantity of the masika and vuli rainy seasons, detailed information on maize seed suppliers and sales, and the migration patterns of community members due to livestock activities.

Key Notes :
The name of the nearest supplier of improved maize seeds has been dropped for confidentiality purposes.

SECTION F : MARKET PRICES

Level of Observation :	Item	**Data File :**	CM_SEC_F, CM_SEC_F_ID
Unique Identifier :	interview__key, item_id	**Additional Info :**	

Description :
Market prices for the surveyed communities, reported for both the village level and the district capital area.

Key Notes :
Section F is split into two datasets: CM_SEC_F, which provides detailed location information on the community the market prices were collected at and CM_SEC_F_ID, which provides unit, weight, and price data per item. The unique identifier for CM_SEC_F is **interview__key**, while for the corresponding CM_SEC_F_ID dataset it is the combination of **interview__key** and **item_id.**

The GPS coordinates are removed for confidentiality purposes.

SECTION G : LOCAL UNITS

Level of Observation :	Item-Observation number	**Data File :**	CM_SEC_G
Unique Identifier :	interview__key, item_id, obs_id	**Additional Info :**	

Description :
Records the local units used for certain items in the surveyed communities. Similar to Section CF, the information is collected both at the village level and the district capital area. The kilogram or liter equivalent for the local units is collected, in addition to the price of the item in that local unit.

Key Notes :
The unique identifier is a combination of the community location variables plus the food item code, food item name, and food item number. This is necessary as each food item (**item_id**) is allowed up to three separate responses, or observations, of local units (**obs_id**).

Appendix A: Calculation of Panel and Cross Section Weights

In order to produce representative estimates from the successfully interviewed sample, weights must be applied to the information provided by sampled households. For the NPSY5, two sets of weights were produced: (1) a panel weight for the set of households that were interviewed in NPSY4 and NPSY5 and (2) a cross section weight for the pooled sample of panel households as well as the "booster sample" of urban households freshly interviewed in NPSY5. The procedure for each of these weights is slightly different and thus documented here separately.

Panel Weights

The procedure for calculating the panel weights builds upon the procedures implemented in NPSY4 as well as NPSY1-NPSY3. The panel weights are developed in the following steps:

1) Begin with the "base weights", or those calculated during the previous round of the survey
2) Incorporate fair-share weights for composition changes
3) Derive attrition adjusted weights for all individuals, including split-off[7] households, then aggregate these weights to the household level
4) Calibrate pooled weights to known population totals
5) Trim outliers

Each of these steps in discussed in detail below.

1) Base Weights from 2014/15 NPSY4 Refresh Sample

The panel weight calculations are based on the 2014/15 household weights for the Refresh Sample households. These weights are based on the inverse probability of selection, EA level non-response correction, trimming of outlier weights, and a post-stratification correction[8]. These probability weights form the first component of the NPSY5 calculations:

$$W_1 = W_{2014}^R$$

Where W_1 are the base weights for the NPSY5 panel sample and W_{2014}^R are the final weights for the NPS4 Refresh Sample. The next step in a panel weight calculation would typically be the generation of "shadow weights", or "phantom weights", for those not in the previous year but still eligible as part of the sample. These weights are equivalent to what their weight would have been in the previous year, if present. In the NPSY5 panel sample, no shadow weights were calculated as no household was tracked if it had not been found and interviewed in the 2014/15 round.

2) Fair Share Correction

[7] 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. Since this distinction is arbitrary, however, there is no mathematical difference between the parent and split households.

[8] See BID for the NPS4 for further details on the calculations of the base weights for the Refresh Sample.

Based on the tracking protocols, the split-off rules for the NPSY5 allow for the incorporation of people who now live with original sample members. For example, a young adult living with his parents in 2014, may have formed a new household in 2018, getting married and having 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.

In an ideal world, it would be possible 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 ½ 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.[9]

$$a_1 = \begin{cases} 1 & \text{otherwise} \\ \frac{1}{2} & \text{if new members} \end{cases}$$

Then the adjusted weights would be:

$$W_2 = W_1 * a_1$$

A limitation of the panel methodology is that the represented population is not identical to the 2020 Tanzanian 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 2020 Tanzanian population to permit the desired cross-sectional estimates (using the panel sample).

[9] 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 2014.

3) Attrition Correction Factor

All household panel surveys must tackle the problem of attrition, sample members selected for follow-up interview which cannot be located and/or interviewed. The methodology used to adjust weights for attrition in the TZNPS 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).

The attrition correction in the case of the TZNPS 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. Therefore, in the TZNPS, the second methodology is employed.

To obtain the attrition adjustment factor, the probability that a sample household was successfully re-interviewed in the NPSY5 round is modeled with the linear logistic model at the level of the individual. A binary response variable is created by coding the response disposition for eligible households that do not respond in the second round as 0, and households that do respond as 1[10].

Next, a logistic response propensity model is fitted, using household and individual characteristics measured in NPSY4 as covariates. Included covariates are:

- gender
- age
- marital status
- current school attendance
- years of education
- labor force participation
- disability status
- household size
- rural / urban status
- household per capital consumption expenditures
- residence in agricultural (crop) household
- residence in a livestock household
- residence in a fishing household
- residence in household owning non-farm enterprise
- residence in household receiving transfer income
- residence in dwelling with improved walls
- residence in dwelling with improved roof
- residence in dwelling with improved floor
- residence in a household with receiving wage income

[10] 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.

- residence is a titled dwelling
- residence in a rental dwelling
- residence in household with at least one member owning mobile phone
- region of residence[11]

In a few limited cases, values of unit-level variables were missing from the 2014/15 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 TZNPS strata (urban/rural and region). Overall, 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 2020/21 survey. These response probabilities were then aggregated to the household level (by calculating the mean). In principle, the inverse of this household-level predicted probability could serve as the adjustment factor applied to the weights. However, in order to reduce potential distortions due to outliers, all units were sorted into 10 response classes according to the deciles of predicted response propensity and the average predicated probability within each class taken as the adjustment factor. Then the adjusted weights would be:

$$W_3 = W_2 \times \frac{1}{\bar{\hat{p}}_c}$$

Where $\bar{\hat{p}}_c$ is the average predicted probability of response for response class/decile c and $c = 1, \ldots, 10$.

4) Calibration

The weight obtained after the first three adjustment (W_3) was then calibrated according to official demographic and population projections for 2021 provided by NBS. This calibration step will help to reduce the overall standard errors enabled the weighted population totals from the survey to correspond with official population estimates. The calibration was applied using the *ReGenesees* package in R using a calibration model that minimizes the distance between the base weights and calibrated weights using a logit distance function. The population parameters there were used are (1) number of households, number of males, and number of females for 30 regions, and (2) the breakdown of the population of males and females by age group, separately for the Mainland and Zanzibar[12]. The population projection values used are presented in Table A1 and A2 below.

5) Trimming

Lastly, the calibrated weights were **trimmed** at the at the 1st and 98th percentiles. The trimming was performed using the *ReGenesees* package in R which adjusts the weights within given bounds, minimizing the deviation from the estimates obtained from the calibration in step 4 and maintaining the calibration parameters.

[11] Zanzibar regions are aggregated into Unguja (region 51, 52, and 53) and Pemba (region 54 and 55) islands due to a lack of variable with the disaggregated categories.

[12] Population estimates disaggregated for rural and urban areas would have been preferred but were not available.

Table A1: Regional Population Projections for 2021

Region	Number of households	Number of males	Number of females
DODOMA	593,406	1,339,007	1,390,661
ARUSHA	479,225	1,052,779	1,103,732
KILIMANJARO	464,407	973,346	1,023,606
TANGA	533,923	1,223,641	1,285,798
MOROGORO	636,195	1,381,129	1,418,131
PWANI	315,644	666,641	690,630
DAR ES SALAAM	1,381,660	2,699,021	2,827,617
LINDI	275,732	505,698	542,085
MTWARA	407,412	719,218	788,208
RUVUMA	376,679	826,638	868,419
IRINGA	280,316	572,248	605,079
MBEYA	835,651	1,743,598	1,849,702
SINGIDA	331,013	868,646	885,724
TABORA	531,866	1,581,506	1,609,688
RUKWA	262,001	640,208	669,799
KIGOMA	508,521	1,416,163	1,482,405
SHINYANGA	348,174	1,010,417	1,043,812
KAGERA	713,456	1,654,735	1,698,506
MWANZA	698,911	1,966,050	2,017,743
MARA	444,671	1,215,805	1,274,350
MANYARA	372,587	973,340	964,110
NJOMBE	208,520	407,869	447,063
KATAVI	150,393	418,509	423,691
SIMIYU	350,507	1,179,705	1,238,790
GEITA	416,248	1,259,287	1,279,827
KASKAZINI UNGUJA	49,480	116,849	120,656
KUSINI UNGUJA	31,294	70,348	70,473
MJINI MAGHARIBI UNGUJA	144,237	361,591	388,442
KASKAZINI PEMBA	57,768	150,305	155,868
KUSINI PEMBA	52,421	137,941	145,135

Source: NBS

Table A2: Age Group Population Projections for 2021

Age group	Mainland		Zanzibar	
	Men	Women	Men	Women
0 -5	5,864,271	5,754,502	163,129	160,198
6-10	3,890,487	3,848,621	116,389	114,770
11-15	3,479,171	3,465,986	99,621	99,510
16-20	2,925,980	2,953,170	82,785	83,775
21-25	2,571,651	2,589,549	78,099	79,652
26-30	1,974,347	2,254,736	61,818	73,218
31-35	1,586,336	1,903,870	50,398	62,125
36-40	1,416,885	1,675,436	40,762	50,684
41-45	1,199,282	1,284,738	34,812	39,464
46-50	993,420	1,071,106	30,189	35,720
51-55	726,226	731,946	22,825	23,516
56-60	572,981	677,629	20,440	25,756
61-65	385,354	386,748	14,113	11,574
66-70	288,939	330,966	9,915	9,273
71-75	180,751	196,291	5,710	4,825
76 and older	239,123	303,882	6,029	6,514

Source: NBS

Cross Section Weights

The NPSY5 included a "booster" sample of urban households that were visited for the first time in the NPSY5 round. In order to take advantage of this additional sample, a second set of weights needed to be calculated which applied to the sample of NPSY5 panel households as well as the urban "booster" sample. These weights we shall refer to as the "cross section" weights though the combined sample may serve as a new panel in future round of the NPS.

As for the panel weights, the calculation of the cross section weights proceeded in several steps:

1) Calculate the design weights for the urban booster sample
2) Implement a pooling adjustment to the booster sample design weight and the panel weight
3) Calibrate to known population totals
4) Trim outliers

1) Design weights for the urban booster sample

The first step to is establish the base weights for the urban booster sample. The booster sample followed a standard two-stage design with probability proportional to size in the first stage (EA selection) and simple random sample in the second stage (household selection). The design weights are the inverse of the product of these two probabilities of selection:

$$W_{se,1}^{b} = m_s \frac{L_s}{l_s} \times \frac{N_{se}}{n_{se}}$$

Where $W^b_{se,1}$ is the base weight for the booster sample, m_s is the number of EAs selected in stratum s, L_s is the total number of households in the stratum, l_s is the total number of households in the EA according to the frame, N_{se} is the total number of households identified in the listing in EA e, and n_{se} is the number of households selected in the EA. The strata for the booster sample are the 5 regions covered by the oversample (specifically, urban Mbeya, Arusha, Mwanza, Tanga, and Dodoma).

Note that there was no response for the booster sample and thus no nonresponse adjustment was implemented.

2) Pooling adjustment

Now the weights from the booster and panel sample need to be combined. One common approach which was implemented here is to apply a **pooling adjustment.** This adjustment ensures an appropriate balance of weights is maintained between the booster sample and panel same in the relevant urban regions. In this case, the adjustment is simply the relative share of the sample in each stratum:

$$w^b_{se,2} = w_{sef} \times \frac{n_{sf}}{\sum_f n_{sf}}$$

Where n_{sf} is the number of observations in the stratum from sample $f = \{panel, booster\}$ and $\sum_f n_{sf}$ is the total number of (pooled) observations in the stratum, and w_{sef} is the base weight with the base weight for the booster sample being the design weight and the base weight for the panel sample is the final panel weight calculated above.

3) Calibration

Following application of the pooling adjustment, the weights were calibrated following the exact sample process and parameters as for the panel weights above.

4) Trimming

Lastly, the calibrated weights were **trimmed** at the at the 1st and 98th percentiles. The trimming was performed using the *ReGenesees* package in R which adjusts the weights within given bounds, minimizing the deviation from the estimates obtained from the calibration in step 4 and maintaining the calibration parameters.

Appendix B: Confidential Information

To maintain the confidentiality of respondents, certain parts of the NPS database have not been made publicly available. The confidential variables pertain to (i) names of the respondents in the household and community questionnaires, (ii) village and constituency names, (iii) descriptions of household dwelling and agricultural plot locations, (iv) phone numbers of household members and their reference contacts, (v) GPS-based household and agricultural plot locations, (vi) names of the children of the head/spouse living elsewhere, (vii) names of the deceased household members, and (viii) other directly-identifying information i.e. names of banks.

Appendix C: Consumption Aggregate

Consumption rather than income is used to measure the welfare of households in NPS. This Appendix outlines the principles involved in the construction of the consumption measure and describes the components and estimation procedure of the nominal household consumption expenditures. While the methodology used to construct the consumption aggregates for the NPS 2020/21 broadly follows that of previous rounds, the NPS 2020/21 included rent and durables, which previous waves did not consider. The inclusion of rent and durables in NPS 2020/21 follows plans to include these components in the consumption aggregates of the upcoming new poverty surveys. Given that the NPS 2020/21 is a panel of the refresh sample from 2014/15, we also reconstructed the aggregates for NPS 2014/15 to include rent and durables. Thus, any comparative welfare analysis using the NPS 2014/15 and NPS 2020/21 should use the new consumption aggregates, ***consumption_real_y4*** and ***consumption_real_y5*** respectively. Note, however, that due to the inclusion of the rent and durables, these two aggregates (***consumption_real_y4*** and ***consumption_real_y5***) are not comparable to the ones constructed for earlier rounds, including the one in *consumptionNPS4* for NPS 2014/15 released earlier.

1.0 The construction of the consumption aggregate

Creating the consumption aggregate is guided by theoretical and practical considerations. First, it must be as comprehensive as possible given the available information. Omitting some components assumes that they do not contribute to people's welfare or that they do not affect the welfare ranking of the population. Second, market and non-market transactions are to be included, which means that purchases are not the sole component of the indicator. Third, expenditure is not consumption. For perishable goods, mostly food, and for frequently used non-food items, it is usual to assume that all purchases are consumed. However, for other goods and services, such as housing or consumer durable goods, imputations have to be made to approximate consumption. Fourth, a common reference period should be chosen. Typically, each consumption module in a survey has a different reference period, for instance, education could refer to the last 12 months, food could refer to the last week, and health could refer to the last month. Following common practice in Tanzania, consumption will be adjusted and reported for the period of 28 days.

1.1 Food component

A few general principles are applied in the construction of this component. First, all possible sources of consumption are included. This means that the food component comprises not only consumption from purchases in the market or from meals eaten away from home but also food that was produced by the household or received as a gift or a payment for a work, etc. Second, only food that was actually consumed, as opposed to total food purchases or total home-produced food, enters into the consumption aggregate. Third, non-purchased consumed food needs to be valued and included in the welfare measure. The NPS gathers information on the amount spent on

purchases and on the quantity purchased for all food items. A measure of prices, called unit values, can be obtained by dividing the monetary amount spent by the quantity purchased. Then the unit values are multiplied by quantity of food-item consumed by a household from different sources: self-produced, received as a gift or as a payment. The food component of the aggregate is a sum of imputed values of food consumption, across all food items the household consumed in the last 7 days, from all sources. The food consumed outside of home is valued as a sum of monetary expenditures on meals by all members of household.

1.2 Non-food component

Expenditure data on a wide range of non-food items are available in the NPS, such as water, kerosene, electricity, health, transportation, communications, recreation, education, furnishings, personal care, etc. Unlike food, the NPS only collects data on purchases of non-food items, that is, the survey assumes that the consumption of non-food goods and services coming from own-production, from gifts or from other sources is negligible and can be ignored. In addition, the NPS does not gather information on quantities purchased because most non-food items are too heterogeneous to try to calculate prices.

Each non-food component is associated with a particular reference period, which reflects the frequency of that purchase or consumption. For instance, expenses on public transportation are collected for the last seven days, expenses on mobile phones and personal care are collected for the last month, and expenses on furnishings and small appliances for the last twelve months.

The information about some non-food goods and services needs to be excluded from the consumption aggregate because those items are not consumption. Payments of mortgages or debts are financial transactions and not consumption. Losses to theft are neither expenditure nor consumption. Remittances to other households are expenditures but not consumption. Expenditures on marriages, dowries, births and funerals are consumption but given their sporadic nature and the fact that the reported amounts are typically rather large, this consumption is left out to avoid overestimating the true level of welfare of the household.

Education expenditures are important component of the consumption aggregate and include all education related expenses from pre-school to tertiary education levels: school fees, uniform, textbooks, meals and lodging, transport, private tutoring and other expenses incurred while obtaining education. Education expenses were recorded for most recent school/academic year.

The NPS captures the health care seeking behaviour of households. Health expenditures are recorded on consultations, medicines, laboratory exams, hospitalization charges, transport and other out of pocket cost related to the health care. Elsewhere the motivation for excluding the health-related expenditures from aggregate is linked to consideration of health cost as a

"regrettable necessity". If a member of household falls ill and incurs medical expense this will increase total expenditures and therefore household's assumed level of welfare when in fact, the opposite may be the case. Thus, to avoid the bias in ranking of households the decision was to exclude hospitalization or extraordinary medical related cost, but include current health care cost, like regular medicine, consultations, etc.

Non-food expenditures are valued at the purchase or self-reported acquisition value. Depending on the recall period the expenditures are annualized by a factor of (365/7) for 7-day recall items or (12/1) for one month recall items, except for a few items, like education which were calculated without annualization

1.3 Durable goods

Utilization of consumer durable goods (versus agricultural or other productive equipment), such as telephones, beds, bicycles, motorcycle, cars, air conditioners, etc., is an important component of the household welfare. Given that these goods last for many years, the expenditure on purchases is not the proper measure to consider. The right measure to estimate, for consumption purposes, is the stream of services and benefits that households derive from the use of all consumer durable goods in their possession over the relevant reference period, normally over a year. This flow of utility is unobservable, but it can be assumed to be related to the value of the good and thus statistically imputed. The imputation is based on the hypothetical experiment whereby, on the one hand, household sells the good and obtains the interest gains and on the other hand, household benefits from using the good, but foregoes interest gain, and incurs the depreciation of the good. Difference between these two components reflect the cost the household is willing to pay to utilize the durable good. Mathematically the value of durables consumption could be approximated by the following formula: $p_t(r_t + \delta_t)$, where p_t is current value of the durable item, r_t is real interest rate (i.e., adjusted for inflation) and δ_t is annual depreciation rate for the durable item, which needs to be also imputed: $\delta = 1 - \left(\frac{p_t}{p_0}\right)^{1/t}$.

Information on the number of the consumer durable goods owned, their age, and their likely value (current and original) is required to estimate the user cost of durables. Unfortunately, rounds 1 and 2 of the NPS only provide data on the number of durable goods owned by the household, while rounds 3, 4, and 5 asked for all required information - the number owned, age, and value. Calculating this consumption component in previous rounds would have involved making assumptions about their age, current value and lifespan. This might have resulted in an extremely imprecise estimation, thus it was decided to exclude this component from the consumption aggregate in previous rounds. However, given availability of data and the importance of the durables for consumption aggregate, the flow of services from the use of consumer durables good was imputed and included in the NPS 2014/15 and NPS 2020/21 to ensure comprehensiveness of the welfare aggregate.

1.4 Housing

Living in a good dwelling with good housing conditions is considered to be an essential part of people's living standards. As in the case of durable goods, the objective is to measure the flow of services received by the household from occupying its dwelling. Housing cost is defined as (implicit) value or benefit that household receives from occupying a dwelling and not the expenditures on purchasing the dwelling itself. When a household lives in a rented dwelling, and provided that rental markets function well, that value would be the actual rent paid. If enough households rent dwellings, imputations can be made for those households that own their dwelling. It is common to include a question for homeowners asking them to provide the hypothetical rent they would pay for renting their dwelling. In principle, these self-reported rents can be used to value the consumption the household gets from occupying its dwelling, but these amounts are not always credible or usable, particularly in rural areas where very few households rent. In Rounds 4 and 5, both actual and potential rents are reported, which makes it possible to impute or predict the housing rent, based on hedonic regression model. The dependent variable is actual rent paid in logarithmic form, regressed on a set of housing characteristics variables like, location, number of rooms, material of roof, material of floor, material of wall, amenities/utilities (toilet, water sources, garbage collection etc.). The imputed rent is a predicted value of housing from regression, that was transformed back into TSh terms from log form, using Duan-Smearing method and applied to the households that own the housing or do not pay rent for dwelling. The challenge in estimating the rent is that rental market in some areas specifically in rural areas is quite thin. To overcome this, the hedonic model is applied at more aggregated strata levels, instead of disaggregated region levels.

2.0 Price adjustment

To ensure inter-household utility comparability, nominal consumption expenditures of the household must be adjusted for cost-of-living differences across space and time. To this end, temporal and spatial price adjustments are implemented to render consumption in real terms. Temporal differences are associated with the duration of the fieldwork (TSh 1,000 in January 2021 may not have the same value as in August 2021) as well as with the different recall periods (TSh 1,000 spent in the last month may not have the same value as in the last quarter or in the last year)[13]. Spatial differences are associated with the location of households interviewed in the survey (the purchasing power of TSh 1,000 in Dar es Salaam may be different than in Ruvuma).

[13] In the NPS 2020/21, the number of interviews conducted in December 2020 and January 2022 were negligible; therefore, these have been reallocated to the January 2021 and December 2021 months, respectively, for simplicity of quarter definitions.

The price data required to construct the price index could come partly or fully from the NPS. A price index is a combination of prices and budget shares in a base and a comparison period. The budget shares are the weights that each commodity has in the index and are equivalent to their share in the cost of the bundle being analysed. The NPS can provide information on budget shares for all items, but information on prices (unit values) only available for food items. Two possible price indices could be constructed: a price index based only on food items (the assumption would be that non-food prices follow the same temporal and spatial differences as food items) or a price index that takes into account both food and non-food by combining information from the survey (food prices and weights for food and non-food items) and the official consumer price index (for non-food prices).

Fisher price indices based only on food items were employed to adjust the nominal consumption aggregate for spatial and temporal price differences. Fisher price indices do a better job than Laspeyres or Paasche price indices at capturing differences in consumption patterns across domains as a consequence of differences in relative prices. They also avoid overstating or understating the true inflation (as would be the case with Laspeyres and Paasche respectively).[14] Price indices were estimated by stratum (an area) and quarter (a period of three consecutive months) and the base period comprises the entire period of each round of the NPS – that is, price indices were calculated separately for each round. A price index by stratum and month would have been ideal, but complications arose with the sample size because in some combinations of stratum and month only a few households were interviewed. In this context, price indices by stratum and quarter is the second best solution. Fisher price indices by stratum and quarter were constructed using the following formula:

$$F_i = \sqrt{L_i P_i}$$

where i is a combination of stratum and quarter, L refers to a Laspeyres price index and P refers to a Paasche price index. The Laspeyres and Paasche price indices are defined as

$$L_i = \sum_{k=1}^{n} w_{0k}\left(\frac{p_{ik}}{p_{0k}}\right), P_i = \left[\sum_{k=1}^{n} w_{ik}\left(\frac{p_{ik}}{p_{0k}}\right)^{-1}\right]^{-1}$$

where w_{0k} is the average household budget share of item k in the country, w_{ik} is the average household budget share of item k in stratum and quarter i, p_{0k} is the national median price of item k and p_{ik} is the median price of item k in stratum and quarter i.

[14] See Deaton and Tarozzi (2000).

Food items that had been purchased by at least 10 households by stratum and quarter (i.e., available record of 10 transactions) were included in the construction of the price indices. Median unit values were estimated for the price indices because the median is less sensitive to outliers than the mean.

Table 2.1 shows the Fisher food price indices for each round of the NPS. Spatial price differences across strata remain fairly constant over time. The most expensive stratum is Dar es Salaam whereas the cheapest is rural areas in mainland. The cost of living in other urban areas in mainland and Zanzibar is relatively similar. Temporal price differences across quarters are noticeably larger during the NPS 2010/2011, thus reflecting a higher inflation in the second round compared to the first round.

Table 2.1: Fisher food price indices by stratum and quarter, NPS 2008/2009, NPS 2010/2011, NPS 2012/2013, NPS 2014/2015 and NPS-SDD 2019/2020

NPS 2008/2009	Oct-Dec 2008	Jan-Mar 2009	Apr-Jun 2009	Jul-Sep 2009
Dar es Salaam	1.08	1.18	1.20	1.15
Other urban	1.00	1.04	1.04	1.04
Rural	0.92	0.86	0.92	0.96
Zanzibar	1.03	1.06	1.07	1.07

NPS 2010/2011	Oct-Dec 2010	Jan-Mar 2011	Apr-Jun 2011	Jul-Sep 2011
Dar es Salaam	1.05	1.14	1.17	1.18
Other urban	0.90	0.97	1.06	1.08
Rural	0.87	0.86	0.98	1.02
Zanzibar	0.89	0.98	1.06	1.07

NPS 2012/2013	Oct-Dec 2012	Jan-Mar 2012	Apr-Jun 2013	Jul-Sep 2013
Dar es Salaam	1.12	1.17	1.13	1.07
Other urban	0.99	1.04	1.02	0.93
Rural	0.95	0.94	1.00	0.93
Zanzibar	0.88	0.91	0.93	0.99

NPS 2014/2015	Oct-Dec 2014	Jan-Mar 2015	Apr-Jun 2015	Jul-Sep 2015
Dar es Salaam	1.00	1.09	1.15	1.18
Other urban	0.93	0.98	1.01	1.05
Rural	0.94	0.91	0.98	0.94
Zanzibar	0.89	0.86	0.85	0.99

NPS 2020/2021	Jan-Mar 2021	Apr-Jun 2021	Jul-Sep 2021	Oct-Dec 2021
Dar es Salaam	1.11	1.03	1.07	1.11
Other urban	1.02	0.99	1.00	1.05
Rural	0.97	0.99	0.96	1.00
Zanzibar	1.01	1.07	1.06	1.04

3.0 Household composition adjustment

The final step in constructing the welfare indicator involves going from a measure of standard of living defined at the household level to another at the individual level. Ultimately, the objective is to make comparisons across individuals and not across households. Two types of adjustments have to be made to correct for differences in composition and size. The first relates to demographic composition. Household members have different needs based mainly on their age and gender, although other characteristics can also be considered. Equivalence scales are the factors that reflect those differences and are used to convert all household members into "equivalent adults". For instance, children are thought to need a fraction of what adults require, thus if a comparison is made between two households with the same total consumption and equal number of members, but one of them has children while the other is comprised of only adults, it could be expected that the former will have a higher individual welfare than the latter. While there is no agreement on a consistent methodology to calculate these scales, it is important to maintain the same scale across rounds. Most scales are based on nutritional grounds, but while a child may need only 50% of the food requirements of an adult, it is not clear why the same scale should be carried over non-food items. It may very well be the case that the same child requires a larger proportion than the adult in education or clothing.[15]

The second adjustment focuses on the economies of scale in consumption within the household. The motivation for this is the fact that some of the goods and services consumed by the household have characteristics of "public goods". A good is said to be public when its consumption by a member of the household does not necessarily prevent another member from consuming it as well. Examples of these goods could be housing and durable goods. For example, one member watching television does not preclude another from watching too. Larger households may need to spend less to be as well-off as smaller ones. Hence, the bigger the share of public goods in total consumption, the larger the scope for economies of scale. On the other hand, private goods cannot be shared among members – once one household member has consumed them, no other member can. Food

[15] See Deaton and Muellbauer (1986) or Deaton (1997).

is the classic example of a private good and, for instance, in poor economies, where food represents a sizeable share of the household budget, little room exists for economies of scale.

Welfare analysis in Tanzania employs an adult-equivalent scale to implement these two adjustments (see Table 3.1). In general, children are thought to consume less than adults and women less than men. An alternative and common practice would have been to use a per capita adjustment for household composition. This is a special case of both adjustments and implies that children consume as much as adults and there is no room for economies of scale. In other words, all members within the household consume equal shares of the total consumption and costs increase in proportion to the number of people in the household. In general, per capita measures will underestimate the welfare of households with children with respect to families with no children, and the welfare of large households with respect to families with a small number of members.

Table 3.1: Adult-equivalent scale by gender and age used in NPS

Age (years)	Male	Female
0-2	0.40	0.40
3-4	0.48	0.48
5-6	0.56	0.56
7-8	0.64	0.64
9-10	0.76	0.76
11-12	0.80	0.88
13-14	1.00	1.00
15-18	1.20	1.00
19-59	1.00	0.88
60 and more	0.80	0.72

Appendix D: Tanzanian Educational System

Tanzania has 13 years of formal schooling – D1 to D7 and F1 to F6.

D1 - Standard I (1st year)
D2 - Standard II (2nd year)
D3 - Standard III (3rd year)
D4 - Standard IV (4th year)
D5 - Standard V (5th year)
D6 - Standard VI (6th year)
D7 - Standard VII (7th year)

F1 - Form I (8th year)
F2 - Form II (9th year)
F3 - Form III (10th year)
F4 - Form IV (11th year)
F5 - Form V (12th year)
F6 - Form VI (13th year)

Prior to independence, there used to be a D8 – Grade 8. Additionally, all classes used to be taught in English but following independence, the Tanzanian primary education system switched to being based in Swahili. Until the early 1970s, a student was required to take an exam after Standard IV in order to proceed to Standard V.

For a student to proceed to a government secondary school, the student has to receive a passing grade on the Primary School Leaving Exam, which is taken after Standard VII. Otherwise, the student can continue education in a private secondary school.

If an individual does not proceed to Form I (F1), they can take the MS+ Course. This is a vocational course – for jobs such as carpentry - that ranges from three months to a year. For a student to proceed to Form V (F5), they must take the Form IV (F4) national level exam, which is also known as O+. It is important to note that the O+ does not constitute an extra year of school. It is simply a required final exam an individual must take to advance educationally. So an individual could have finished Form IV (F4) but have failed the O+, thus not proceeding to Form V (F5). Students must also take a national level exam, A+, after completing For VI (F6). If they pass the A+ exam with a certain grade, they go directly to University (U1 through U5).

If one does not pass the O+ exam, one can take a certificate course at a technical school.

If one does not pass the A+ exam, one can do the Diploma course or choose to not pursue further education. If the individual completes the Diploma course, they can then enroll into University.

U1 – 14th year
U2 – 15th year
U3 – 16th year
U4 – 17th year
U5 – 18th year
U5+ – 18th plus year

The Diploma course can range from one to three years. Acquiring a Diploma degree in Tanzania can qualify an individual to be a primary school teacher. However, participating in the Diploma course does not technically add additional years of education to an individual's record. This is because universities treat A+ certification and Diploma's equally for admission. Therefore, to calculate the number of educational years an individual, who attended the Diploma course, has is to add one year to their current university level (U1 through U5+). For example, a student with a Diploma who is in U2 would have 16 years of schooling.

Appendix E: Description of the Tanzania Standard Classification of Occupation (TASCO)

In Section E of the Household Questionnaire (HH_SEC_E1.dta), the TASCO codes are used. Respondents were asked to describe what kind of job/work they did. Based off the respondent's description, the TASCO codes were assigned. Depending on the specificity of the job/work description affects if there is a two or three digit TASCO code. Respondents were asked to be specific as possible but in some cases, their responses did not allow for a three digit TASCO code to be assigned. The following list is all of the potential TASCO codes and those used within the survey.

MAJOR GROUP 1: MANAGERS

- 11 Legislators, Chief executives and senior officials
  - 111 Legislators and senior official
    - 1111 Legislators
    - 1112 Senior government executive officials
    - 1113 Traditional chiefs and heads of village
    - 1114 Senior officials of special-interest organizations
    - 1119 Senior government administrative other specialised services
  - 112 Managing directors and chief executives
    - 1120 Managing directors and chief executives
- 12 Administrative and commercial managers
  - 121 Business services and administration managers
    - 1211 Finance managers
    - 1212 Human resource managers
    - 1213 Policy and planning managers
    - 1214 Business services mmanagers
    - 1215 Administrative managers
    - 1219 Business services and administration managers not elsewhere classified
  - 122 Sales, marketing and development managers
    - 1221 Sales and marketing managers
    - 1222 Advertising and public relations managers
    - 1223 Research and development managers
- 13 Production and specialized services managers
  - 131 Production managers in agriculture, forestry and fisheries
    - 1311 Agricultural and forestry production managers
    - 1312 Aquaculture and fisheries production managers
  - 132 Manufacturing, mining, construction, and distribution managers
    - 1321 Manufacturing managers
    - 1322 Mining managers
    - 1323 Construction managers
    - 1324 Supply, distribution and related managers

133 Information and communications technology service managers
1330 Information and communications technology service managers
134 Professional services managers
1341 Child care services managers
1342 Health services managers
1343 Aged care services managers
1344 Social welfare managers
1345 Education managers
1346 Financial and insurance services branch managers
1347 Legal managers
1348 Property manager
1349 Professional services managers not elsewhere classified
14 Hospitality, retail and other services managers
141 Hotel and restaurant managers
1411 Hotel managers
1412 Restaurant managers
142 Retail and wholesale trade managers
1420 Retail and wholesale trade managers
143 Other services managers
1431 Sports, recreation and cultural centre managers
1432 Personal care, cleaning and related services managers
1433 Safety and security managers
1439 Services managers not elsewhere classified

MAJOR GROUP 2: PROFESSIONALS

21 Science and engineering professionals
211 Physical and earth science professionals
2111 Physicists and astronomers
2112 Meteorologists
2113 Chemists
2114 Geologists and geophysicists
212 Mathematicians, actuaries and statisticians
2120 Mathematicians, actuaries and statisticians
213 Life science professionals
2131 Biologists, botanists, zoologists and related professionals
2132 Farming, forestry and fisheries and related professionals
2133 Environmental protection professionals
214 Engineering professionals (excluding electro-technology)
2141 Industrial and production engineers
2142 Civil engineers
2143 Environmental engineers
2144 Mechanical engineers
2145 Chemical engineers

2146 Mining engineers, metallurgists and related professionals
2149 Engineering professionals not elsewhere classified
215 Electro-technology engineers
2151 Electrical engineers
2152 Electronics engineers
2153 Telecommunications engineers
216 Architects, planners, surveyors and designers
2161 Building architects
2162 Landscape architects
2163 Product and garment designers
2164 Town and traffic planners
2165 Cartographers and surveyors
2166 Graphic and multimedia designers
22 Health professionals
221 Medical doctors
2211 Generalist medical practitioners
2212 Specialist medical practitioners
222 Nursing and midwifery professionals
2221 Nursing professionals
2222 Midwifery professionals
223 Traditional and complementary medicine professionals
2230 Traditional and complementary medicine professionals
224 Paramedical practitioners
2240 Paramedical practitioners
225 Veterinarians
2250 Veterinarians
226 Other health professionals
2261 Dentists
2262 Pharmacists
2263 Environmental and occupational health and hygiene professionals
2264 Physiotherapists
2265 Dieticians and nutritionists
2266 Audiologists and speech therapists
2267 Optometrists and ophthalmic opticians
2269 Health professionals not elsewhere classified
23 Teaching professionals
231 University and higher education teachers
2310 University and higher education teachers
232 Vocational education teachers
2320 Vocational education teachers
233 Secondary education teachers
2330 Secondary education teachers
234 Other teaching professionals
2341 Education methods specialists
2342 Special needs teachers

2343 Other language teachers
2344 Other music teachers
2345 Other arts teachers
2346 Information technology trainers
2349 Teaching professionals not elsewhere classified

24 Business and administration professionals
241 Finance professionals
2411 Accountants
2412 Financial and investment advisers
2413 Financial analysts
2414 Procurement and logistics professionals
242 Administration professionals
2421 Management and organization analysts
2422 Policy administration professionals
2423 Personnel and careers professionals
2424 Training and staff development professionals
243 Sales, marketing and public relations professionals
2431 Advertising and marketing professionals
2432 Public relations professionals
2433 Technical and medical sales professionals (excluding ICT)
2434 Information and communications technology sales professionals
244 Regulatory government professionals
2441 Customs and border inspector professionals
2442 Taxation and excise officials' professionals
2449 Regulatory government professionals, not elsewhere classified

25 Information and communications technology professionals
251 Software and applications developers and analysts
2511 Systems analysts
2512 Software developers
2513 Web and multimedia developers
2514 Applications programmers
2519 Software and applications developers and analysts not elsewhere classified
252 Database and network professionals
2521 Database designers and administrators
2522 Computer systems administrators
2523 Computer network professionals
2529 Database and network professionals not elsewhere classified

26 Legal, social and cultural professionals
261 Legal professionals
2611 Lawyers
2612 Judges
2613 Other legal professionals

2619 Legal professionals not elsewhere classified
262 Librarians, archivists and curators
2621 Archivists and curators
2622 Librarians and related information professionals
263 Social and religious professionals
2631 Economists
2632 Sociologists, anthropologists and related professionals
2633 Philosophers, historians and political scientists
2634 Psychologists
2635 Social work and counselling professionals
2636 Religious professionals
2639 Social and religious professionals not else where classified
264 Authors, journalists and linguists
2641 Authors and related writers
2642 Journalists
2643 Translators, interpreters and other linguists
265 Creative and performing artists
2651 Visual artists
2652 Musicians, singers and composers
2653 Dancers and choreographers
2654 Film, stage and producers
2655 Actors and entertainers
2656 Announcers on radio, television and other media
2657 Clowns, magicians, acrobats and related professionals
2658 Animal/ bird keepers and trainers care
2659 Creative and performing artists not elsewhere classified

MAJOR GROUP 3: TECHNICIANS AND ASSOCIATE PROFFESIONALS

31 Science and engineering associate professionals
311 Physical and engineering science technicians
3111 Chemical and physical science technicians
3112 Civil engineering technicians
3113 Electrical engineering technicians
3114 Electronics engineering technicians
3115 Mechanical engineering technicians
3116 Chemical engineering technicians
3117 Mining and metallurgical technicians
3118 Draughtspersons
3119 Physical and engineering science technicians not elsewhere classified
312 Mining, manufacturing and construction supervisors
3121 Mining supervisors
3122 Manufacturing supervisors
3123 Construction supervisors

3124 Food and beverage processing supervisors
3129 Other supervisors not elsewhere classified
313 Process control technicians
3131 Power production plant operators
3132 Incinerator and water treatment plant operators
3133 Chemical processing plant controllers
3134 Petroleum and natural gas refining plant operators
3135 Metal production process controllers
3139 Process control technicians not elsewhere classified
314 Life science technicians and related associate professionals
3141 Life science technicians (excluding medical)
3142 Agricultural technicians
3143 Forestry technicians
3144 Livestock technicians
3145 Fishery technicians
315 Ship and aircraft controllers and technicians
3151 Ships' engineers
3152 Ships' deck officers and pilots
3153 Aircraft pilots and related associate professionals
3154 Air traffic controllers
3155 Air traffic safety electronics technicians
32 Health associate professionals
321 Medical and pharmaceutical technicians
3211 Medical imaging and therapeutic equipment technicians
3212 Medical and pathology laboratory technicians
3213 Pharmaceutical technicians and assistants
3214 Medical and dental prosthetic technicians
322 Nursing and midwifery associate professionals
3221 Nursing associate professionals
3222 Midwifery associate professionals
323 Traditional and complementary medicine associate professionals
3230 Traditional and complementary medicine associate professionals
324 Veterinary technicians and assistants
3240 Veterinary technicians and assistants
325 Other health associate professionals
3251 Dental assistants and therapists
3252 Medical records and health information technicians
3253 Community health workers
3254 Dispensing opticians
3255 Physiotherapy technicians and assistants
3256 Medical assistants
3257 Environmental and occupational health inspectors and associates
3258 Ambulance workers
3259 Health associate professionals not elsewhere classified

33 Business and administration associate professionals
- 331 Financial and mathematical associate professionals
  - 3311 Securities and finance dealers and brokers
  - 3312 Credit and loans officers
  - 3313 Accounting associate professionals
  - 3314 Statistical, mathematical and actuarial associate professionals
  - 3315 Valuers and loss assessors
- 332 Sales and purchasing agents and brokers
  - 3321 Insurance representatives
  - 3322 Commercial sales representatives
  - 3323 Buyers
  - 3324 Trade brokers
- 333 Business services agents
  - 3331 Clearing and forwarding agents
  - 3332 Conference and event planners
  - 3333 Employment agents and contractors
  - 3334 Real estate and property agents
  - 3335 Travel agents
  - 3339 Business services agents not elsewhere classified
- 334 Administrative and specialized secretaries
  - 3341 Office supervisors
  - 3342 Legal secretaries
  - 3343 Administrative and executive secretaries
  - 3344 Medical secretaries
  - 3349 Administrative and executive secretaries, NEC
- 335 Regulatory government associate professionals
  - 3351 Customs and border inspectors
  - 3352 Government tax and excise officials
  - 3353 Government social benefits officials
  - 3354 Government licensing officials
  - 3355 Police inspectors and detectives
  - 3359 Regulatory government associate professionals not elsewhere classified

34 Legal, social, cultural and related associate professionals
- 341 Legal, social and religious associate professionals
  - 3411 Legal and related associate professionals
  - 3412 Social work associate professionals
  - 3413 Religious associate professionals
- 342 Sports and fitness workers
  - 3421 Athletes and sports players
  - 3422 Sports coaches, instructors and officials
  - 3423 Fitness and recreation instructors and program leaders
- 343 Artistic, cultural and culinary associate professionals
  - 3431 Photographers
  - 3432 Interior designers and decorators

3433 Gallery, museum and library technicians
3434 Chefs
3435 Other artistic and cultural associate professionals

35 Information and communications technicians

351 Information and communications technology operations and user support technicians

3511 Information and communications technology operations technicians
3512 Information and communications technology user support technicians
3513 Computer network and systems technicians
3514 Web technicians

352 Telecommunications and broadcasting technicians

3521 Broadcasting and audio-visual technicians
3522 Telecommunications engineering technicians

36 Teaching associate professionals

361 Primary school and early childhood teachers

3611 Primary school teachers
3612 Early childhood educators

362 Secondary education teachers, associate professionals

3620 Secondary education teachers, associate professionals

363 Technical/vocational education teachers, associate professionals

3630 Technical/vocational education teachers, associate professionals

364 Special education teachers, associate professionals

3640 Special education teachers, associate professionals

MAJOR GROUP 4: CLERICAL SUPPORT WORKERS

41 General and keyboard clerks

411 General office clerks

4110 General office clerks

412 Secretaries (general)

4120 Secretaries (general)

413 Keyboard operators

4131 Typists and word processing operators
4132 Data entry clerks
4133 Registry assistant

42 Customer services clerks

421 Tellers, money collectors and related clerks

4211 Bank tellers and related clerks
4212 Bookmakers, croupiers and related gaming workers
4213 Pawnbrokers and money-lenders
4214 Debt-collectors and related workers

422 Client information workers
4221 Travel consultants and clerks
4222 Contact centre information clerks
4223 Telephone switchboard operators
4224 Hotel receptionists
4225 Enquiry clerks
4226 Receptionists (general)
4227 Survey and market research interviewers
4229 Client information workers not elsewhere classified

43 Numerical and material recording clerks
431 Numerical clerks
4311 Accounting and bookkeeping clerks
4312 Statistical, finance and insurance clerks
4313 Payroll clerks
432 Material-recording and transport clerks
4321 Stock clerks
4322 Production clerks
4323 Transport clerks
4329 Material recording and transport clerk, not elsewhere classified

44 Other clerical support workers
441 Other clerical support workers
4411 Library clerks
4412 Mail carriers and sorting clerks
4413 Coding, proof-reading and related clerks
4414 Scribes and related workers
4415 Filing and copying clerks
4416 Personnel clerks
4419 Clerical support workers not elsewhere classified

MAJOR GROUP 5: SERVICE AND SALES WORKERS

51 Personal service workers
511 Travel attendants, conductors and guides
5111 Travel attendants and travel stewards
5112 Transport conductors
5113 Travel guides
512 Cooks
5120 Cooks
513 Waiters and bartenders
5131 Waiters
5132 Bartenders
514 Hairdressers, beauticians and related workers
5141 Hairdressers
5142 Beauticians and related workers
515 Building and housekeeping supervisors

- 5151 Cleaning and housekeeping supervisors in offices, hotels and other establishments
- 5152 Domestic housekeepers
- 5153 Building caretakers

516 Other personal services workers
- 5161 Astrologers, fortune-tellers and related workers
- 5162 Companions and valets
- 5163 Undertakers and embalmers
- 5164 Pet groomers and animal care workers
- 5165 Driving instructors
- 5169 Personal services workers not elsewhere classified

52 Sales workers

521 Street and market salespersons
- 5211 Stall and market salespersons
- 5212 Street food salespersons

522 Shop salespersons
- 5221 Shop keepers
- 5222 Shop supervisors
- 5223 Shop sales assistants

523 Cashiers and ticket clerks
- 5230 Cashiers and ticket clerks

524 Other sales workers
- 5241 Fashion and other models
- 5242 Sales demonstrators
- 5243 Door to door salespersons
- 5244 Contact centre salespersons
- 5245 Service station attendants
- 5246 Food service counters attendants
- 5249 Sales workers not elsewhere classified

53 Personal care workers

531 Child care workers and teachers' aides
- 5311 Child care workers
- 5312 Teachers' aides

532 Personal care workers in health services
- 5321 Health care assistants
- 5322 Home-based personal care workers
- 5329 Personal care workers in health services not elsewhere classified

54 Protective services workers

541 Protective services workers
- 5411 Fire-fighters
- 5412 Police officers
- 5413 Prison guards
- 5414 Security guards
- 5419 Protective services workers not elsewhere classified

MAJOR GROUP 6: SKILLED AGRICULTURAL, FORESTRY AND FISHERY WORKERS

- 61 Market-oriented skilled agricultural workers
  - 611 Market gardeners and crop growers
    - 6111 Field crop and vegetable growers
    - 6112 Tree and shrub crop growers
    - 6113 Gardeners, horticultural and nursery growers
    - 6114 Mixed crop growers
  - 612 Animal producers
    - 6121 Livestock and dairy producers
    - 6122 Poultry producers
    - 6123 Apiarists and sericulturists
    - 6129 Animal producers not elsewhere classified
  - 613 Mixed crop and animal producers
    - 6130 Mixed crop and animal producers
- 62 Market-oriented skilled forestry, fishery and hunting workers
  - 621 Forestry and related workers
    - 6210 Forestry and related workers
  - 622 Fishery workers, hunters and trappers
    - 6221 Aquaculture workers
    - 6222 Inland and coastal waters fishery workers
    - 6223 Deep-sea fishery workers
    - 6224 Hunters and trappers
- 63 Subsistence farmers, fishers, hunters and gatherers
  - 631 Subsistence crop farmers
    - 6310 Subsistence crop farmers
  - 632 Subsistence livestock farmers
    - 6320 Subsistence livestock farmers
  - 633 Subsistence mixed crop and livestock farmers
    - 6330 Subsistence mixed crop and livestock farmers
  - 634 Subsistence fishers, hunters, trappers and gatherers

6340 Subsistence fishers, hunters, trappers and gatherers

MAJOR GROUP 7: CRAFT AND RELATED TRADES WORKERS

- 71 Building and related trades' workers, excluding electricians
  - 711 Building frame and related trades workers
    - 7111 House builders
    - 7112 Bricklayers and related workers
    - 7113 Stonemasons, stone cutters, splitters and carvers
    - 7114 Concrete placers, concrete finishers and related workers
    - 7115 Carpenters and joiners
    - 7119 Building frame and related trades workers not elsewhere classified
  - 712 Building finishers and related trades workers

7121 Roofers
7122 Floor layers and tile setters
7123 Plasterers
7124 Insulation workers
7125 Glaziers
7126 Plumbers and pipe fitters
7127 Air conditioning and refrigeration mechanics
713 Painters, building structure cleaners and related trades workers
7131 Painters and related workers
7132 Spray painters and varnishes
7133 Building structure cleaners
72 Metal, machinery and related trades workers
721 Sheet and structural metal workers, moulders and welders, and related workers
7211 Metal moulders and core makers
7212 Welders and flame cutters
7213 Sheet-metal workers
7214 Structural-metal preparers and erectors
7215 Riggers and cable splicers
722 Blacksmiths, toolmakers and related trades workers
7221 Blacksmiths, hammersmiths and forging press workers
7222 Toolmakers and related workers
7223 Metal working machine tool setters and operators
7224 Metal polishers, wheel grinders and tool sharpeners
723 Machinery mechanics and repairers
7231 Motor vehicle mechanics and repairers
7232 Aircraft engine mechanics and repairers
7233 Agricultural and industrial machinery mechanics and repairers
7234 Bicycle and related repairers
73 Handicraft and printing workers
731 Handicraft workers
7311 Precision-instrument makers and repairers
7312 Musical instrument makers and tuners
7313 Jewellery and precious-metal workers
7314 Potters and related workers
7315 Glass makers, cutters, grinders and finishers
7316 Sign writers, decorative painters, engravers and etchers
7317 Handicraft workers in wood, basketry and related materials
7318 Handicraft workers in textile, leather and related materials
7319 Handicraft workers not elsewhere classified
732 Printing trades workers
7321 Pre-press technicians
7322 Printers
7323 Print finishing and binding workers
74 Electrical and electronic trades workers

741 Electrical equipment installers and repairers
7411 Building and related electricians
7412 Electrical mechanics and fitters
7413 Electrical line installers and repairers
742 Electronics and telecommunications installers and repairers
7421 Electronics mechanics and servicers
7422 Information and communications technology installers and servicers
75 Food processing, wood working, garment and other craft and related trades workers
751 Food processing and related trades workers
7511 Butchers, fishmongers and related food preparers
7512 Bakers, pastry-cooks and confectionery makers
7513 Dairy-products makers
7514 Fruit, vegetable and related preservers
7515 Food and beverage tasters and graders
7516 Tobacco preparers and tobacco products makers
752 Wood treaters, cabinet-makers and related trades workers
7521 Wood treaters
7522 Cabinet-makers and related workers
7523 Woodworking-machine tool setters and operators
753 Garment and related trades workers
7531 Tailors, dressmakers, furriers and hatters
7532 Garment and related pattern-makers and cutters
7533 Sewing, embroidery and related workers
7534 Upholsterers and related workers
7535 Pelt dressers, tanners and fellmongers
7536 Shoemakers and related workers
754 Other craft and related workers
7541 Underwater divers
7542 Shot firers and blasters
7543 Product graders and testers (excluding foods and beverages)
7544 Fumigators and other pest and weed controllers
7549 Craft and related workers not elsewhere classified

MAJOR GROUP 8: PLANT MACHINE OPERATORS AND ASSEMBLERS

81 Stationary plant and machine operators
811 Mining and mineral processing plant operators
8111 Miners and quarriers
8112 Mineral and stone processing plant operators
8113 Well drillers and borers and related workers
8114 Cement, stone and other mineral products machine operators
812 Metal processing and finishing plant operators
8121 Metal processing plant operators
8122 Metal finishing, plating and coating machine operators

813 Chemical and photographic products plant and machine operators
8131 Chemical products plant and machine operators
8132 Photographic products machine operators
814 Rubber, plastic and paper products machine operators
8141 Rubber products machine operators
8142 Plastic products machine operators
8143 Paper products machine operators
815 Textile, fur and leather products machine operators
8151 Fibre preparing, spinning and winding machine operators
8152 Weaving and knitting machine operators
8153 Sewing machine operators
8154 Bleaching, dyeing and fabric cleaning machine operators
8155 Fur and leather preparing machine operators
8156 Shoemaking and related machine operators
8157 Laundry machine operators
8159 Textile, fur and leather products machine operators not elsewhere classified
816 Food and related products machine operators
8160 Food and related products machine operators
817 Wood processing and papermaking plant operators
8171 Pulp and papermaking plant operators
8172 Wood processing plant operators
818 Other stationary plant and machine operators
8181 Glass and ceramics plant operators
8182 Steam engine and boiler operators
8183 Packing, bottling and labelling machine operators
8184 Power-generating plant operator
8185 Other power-generating and related plant operators
8189 Stationary plant and machine operators not elsewhere classified

82 Assemblers
821 Assemblers
8211 Mechanical machinery assemblers
8212 Electrical and electronic equipment assemblers
8219 Assemblers not elsewhere classified

83 Drivers and mobile plant operators
831 Locomotive engine drivers and related workers
8311 Locomotive engine drivers
8312 Railway brake, signal and switch operators
832 Car, van and motorcycle drivers
8321 Motorcycle drivers
8322 Car, taxi and van drivers
833 Heavy truck and bus drivers
8331 Bus and tram drivers
8332 Heavy truck and lorry drivers
834 Mobile plant operators

8341 Mobile farm and forestry plant operators
8342 Earthmoving and related plant operators
8343 Crane, hoist and related plant operators
8344 Lifting truck operators
835 Ships' deck crews and related workers
8350 Ships' deck crews and related workers

MAJOR GROUP 9: ELEMENTARY OCCUPATIONS

91 Cleaners and helpers
911 Domestic, hotel and office cleaners and helpers
9111 Domestic cleaners and helpers
9112 Cleaners and helpers in offices, hotels and other establishments
912 Vehicle, window, laundry and other hand cleaning workers
9121 Hand launderers and pressers
9122 Vehicle cleaners
9123 Window cleaners
9129 Other cleaning workers
92 Agricultural, forestry and fishery labourers
921 Agricultural, forestry and fishery labourers
9211 Crop farm labourers
9212 Livestock farm labourers
9213 Mixed crop and livestock farm labourers
9214 Garden and horticultural labourers
9215 Forestry labourers
9216 Fishery and aquaculture labourers
93 Labourers in mining, construction, manufacturing and transport
931 Mining and construction labourers
9311 Mining and quarrying labourers
9312 Civil engineering labourers
9313 Building construction labourers
932 Manufacturing labourers
9321 Hand packers
9329 Manufacturing labourers not elsewhere classified
933 Transport and storage labourers
9331 Hand and pedal vehicle drivers
9332 Drivers of animal-drawn vehicles and machinery
9333 Freight handlers
9334 Shelf fillers
9335 Storage labourers
94 Food preparation assistants
941 Food preparation assistants
9411 Fast food preparers
9412 Kitchen helpers
95 Street and related sales and service workers
951 Street and related service workers

9510 Street and related service workers
952 Street vendors (excluding food)
9520 Street vendors (excluding food)
953 Shoe cleaning and other street services
9530 Shoe cleaning and other street services
96 Refuse workers and other elementary workers
961 Refuse workers
9611 Garbage and recycling collectors
9612 Refuse sorters
9613 Sweepers and related labourers
962 Other elementary workers
9621 Messengers, package deliverers and luggage porters
9622 Odd job persons
9623 Meter readers and vending-machine collectors
9624 Water and firewood collectors
9629 Elementary workers not elsewhere classified

MAJOR GROUP 0: ARMED FORCES OCCUPATIONS

01 Commissioned armed forces officers
011 Commissioned armed forces officers
0110 Commissioned armed forces officers
02 Non-commissioned armed forces officers
021 Non-commissioned armed forces officers
0210 Non-commissioned armed forces officers
03 Armed forces occupations, other ranks
031 Armed forces occupations, other ranks
0310 Armed forces occupations, other ranks

Appendix F: List of International Standard Industry Codes (ISIC)

In Section E of the Household Questionnaire (HH_SEC_E1.dta), the ISIC codes are used. Respondents were asked to describe their association to different trades and businesses. Based off the respondent's description, the ISIC codes were assigned. Depending on the specificity of trade/business description affects if there is a two, three, or four digit ISIC code. Respondents were asked to be specific as possible but in some cases, their responses did not allow for a three or four digit ISIC code to be assigned. The following list is all of the potential ISIC codes and those used within the survey.

A - *Agriculture, forestry and fishing*
- 01 - Crop and animal production, hunting and related service activities
- 02 - Forestry and logging
- 03 - Fishing and aquaculture

B - *Mining and quarrying*
- 05 - Mining of coal and lignite
- 06 - Extraction of crude petroleum and natural gas
- 07 - Mining of metal ores
- 08 - Other mining and quarrying
- 09 - Mining support service activities

C - *Manufacturing*
- 10 - Manufacture of food products
  - 101 - Processing and preserving of meat
  - 102 - Processing and preserving of fish, crustaceans and mollusks
  - 103 - Processing and preserving of fruit and vegetables
  - 104 - Manufacture of vegetable and animal oils and fats
  - 105 - Manufacture of dairy products
  - 106 - Manufacture of grain mill products, starches and starch products
  - 107 - Manufacture of other food products
  - 108 - Manufacture of prepared animal feeds
- 11 - Manufacture of beverages
- 12 - Manufacture of tobacco products
- 13 - Manufacture of textiles
- 14 - Manufacture of wearing apparel
- 15 - Manufacture of leather and related products
- 16 - Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials
- 17 - Manufacture of paper and paper products
- 18 - Printing and reproduction of recorded media
- 19 - Manufacture of coke and refined petroleum products
- 20 - Manufacture of chemicals and chemical products
- 21 - Manufacture of basic pharmaceutical products and pharmaceutical preparations
- 22 - Manufacture of rubber and plastics products
- 23 - Manufacture of other non-metallic mineral products
- 24 - Manufacture of basic metals

25 - Manufacture of fabricated metal products, except machinery and equipment
26 - Manufacture of computer, electronic and optical products
27 - Manufacture of electrical equipment
28 - Manufacture of machinery and equipment n.e.c.
29 - Manufacture of motor vehicles, trailers and semi-trailers
30 - Manufacture of other transport equipment
31 - Manufacture of furniture
32 - Other manufacturing
33 - Repair and installation of machinery and equipment

D - *Electricity, gas, steam and air conditioning supply*

35 - Electricity, gas, steam and air conditioning supply

E - *Water supply; sewerage, waste management and remediation activities*

36 - Water collection, treatment and supply
37 – Sewerage
38 - Waste collection, treatment and disposal activities; materials recovery
39 - Remediation activities and other waste management services

F - *Construction*

41 - Construction of buildings
42 - Civil engineering
43 - Specialized construction activities

G - *Wholesale and retail trade; repair of motor vehicles and motorcycles*

45 - Wholesale and retail trade and repair of motor vehicles and motorcycles
46 - Wholesale trade, except of motor vehicles and motorcycles
47 - Retail trade, except of motor vehicles and motorcycles
471 - Retail sale in non-specialized stores
472 - Retail sale of food, beverages and tobacco in specialized stores
473 - Retail sale of automotive fuel in specialized stores
474 - Retail sale of information and communications equipment in specialized stores
475 - Retail sale of other household equipment in specialized stores
476 - Retail sale of cultural and recreation goods in specialized stores
477 - Retail sale of other goods in specialized stores
478 - Retail sale via stalls and markets
479 - Retail trade not in stores, stalls or markets

H - *Transportation and storage*

49 - Land transport and transport via pipelines
491 - Transport via railways
492 - Other land transport
4921 - Urban and suburban passenger land transport
4922 - Other passenger land transport
4923 - Freight transport by road
493 - Transport via pipeline
50 - Water transport
51 - Air transport
52 - Warehousing and support activities for transportation
53 - Postal and courier activities

I - *Accommodation and food service activities*

55 – Accommodation
56 - Food and beverage service activities
561 - Restaurants and mobile food service activities
562 - Event catering and other food service activities
563 - Beverage serving activities

J - *Information and communication*
58 - Publishing activities
59 - Motion picture, video and television programme production, sound recording and music publishing activities
60 - Programming and broadcasting activities
61 – Telecommunications
62 - Computer programming, consultancy and related activities
63 - Information service activities

K - *Financial and insurance activities*
64 - Financial service activities, except insurance and pension funding
65 - Insurance, reinsurance and pension funding, except compulsory social security
66 - Activities auxiliary to financial service and insurance activities

L - *Real estate activities*
68 - Real estate activities

M - *Professional, scientific and technical activities*
69 - Legal and accounting activities
70 - Activities of head offices; management consultancy activities
71 - Architectural and engineering activities; technical testing and analysis
72 - Scientific research and development
73 - Advertising and market research
74 - Other professional, scientific and technical activities
75 - Veterinary activities

N - *Administrative and support service activities*
77 - Rental and leasing activities
78 - Employment activities
79 - Travel agency, tour operator, reservation service and related activities
80 - Security and investigation activities
81 - Services to buildings and landscape activities
82 - Office administrative, office support and other business support activities

O - *Public administration and defense; compulsory social security*
84 - Public administration and defense; compulsory social security

P – *Education*
85 - Education

Q - *Human health and social work activities*
86 - Human health activities
87 - Residential care activities
88 - Social work activities without accommodation

R - *Arts, entertainment and recreation*
90 - Creative, arts and entertainment activities
91 - Libraries, archives, museums and other cultural activities
92 - Gambling and betting activities

93 - Sports activities and amusement and recreation activities

S - *Other service activities*

94 - Activities of membership organizations

95 - Repair of computers and personal and household goods

96 - Other personal service activities

T - *Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use*

97 - Activities of households as employers of domestic personnel

98 - Undifferentiated goods- and services-producing activities of private households for own use

U - *Activities of extraterritorial organizations and bodies*

99 - Activities of extraterritorial organizations and bodies

Appendix G: Main Animal Diseases & Key Terms

DISEASE NAME	DEFINITION	HOSTS AFFECTS	CLINICAL SIGNS / SYMPTONS
FMD Foot and Mouth Disease	A highly contagious viral disease	Cattle, sheep, goats, pigs and water buffalos	Fever, ulcerations in the mouth, nose, muzzle, feet, teats. Drooling of saliva can also be observed
CBPP (Contagious Bovine Pleuro-pneumonia)	A highly infectious acute/sub-acute/chronic disease affecting lungs and occasionally joints	Cattle	Fever, cough, increased respiratory rate (in sequence)
CCPP (Contagious Caprine Pleuro-pneumonia)	An acute highly contagious disease	Goats	Fatal weeping pneumonia, labored breathing, death within 7-10 days after onset of clinical signs
LSD Lumpy Skin Disease	It is a viral disease transmitted by insects	Cattle	Persistent fever, generalized skin nodules, edema of corresponding lymph nodes, hair loss
Anthrax	It is infectious but non-contagious	All domestic animals and buffalo	Sudden death, fever, severe diarrhea with blood, bloody nasal discharge
Blackleg	Blackleg is a highly fatal disease of young cattle	Young cattle	Lameness, loss of appetite, rapid breathing and the animal is usually depressed and has a high fever. Characteristic swellings develop in the hip, shoulder, chest, back, neck or elsewhere
PPR (Peste des petits ruminants)	An acute viral Rinderpest like disease of sheep and goats	Goats and sheep	Similar to those of rinderpest in cattle: Respiratory difficulty with cough, high fever, weakness, dull coat, nasal discharge, severe bloody diarrhea
Sheep and goat pox	Contagious viral disease of small ruminants, There is a high mortality rate in susceptible populations	Small ruminants (sheep and goats)	Onset of fever followed by erythematous macules that develop into papules. Lesions may also develop on the mucous membrane and on internal organs, causing systemic signs (respiratory signs, diarrhea, depression, emaciation, abortion and sometimes death)
Newcastle disease	A highly contagious zoonotic poultry disease	Chicken and other domestics fowls and wild avian species	High mortality above 30%, characterized by respiratory and nervous symptoms
Fowl pox	Fowlpox is a relatively low-spreading viral disease of poultry	Chicken and other domestics fowls and wild avian species	Skin lesions and/or plaques in the pharynx
Avian flu	Avian influenza is a highly-contagious flu infection in birds. It is a zoonotic disease	Chicken and other domestics fowls and wild avian species	Decrease in bird's activities, decline in egg production, swelling of the face with blue colored combs and wattles, breathing problems, diarrhoea, muscle paralysis and sudden death
IBD (Infectious Bursal Disease) / Gumboro	A highly contagious disease of young chicks	Chicken	Mortality at 3-6 weeks and can rise up to 75%. Birds are depressed, watery diarrhea

KEY TERMS	
Contagious	Spread by means of contact
Infectious	Spread by means of virus, bacteria or a parasite
Fatal	Leads to death
Zoonotic disease	An infectious disease that can be transmitted to humans and vice versa
Curative	Care provided to improve a situation (especially medical procedures or applications that are intended to relieve illness or injury)

Appendix H: Animal Breeding Types

CODE	BREEDING TYPE	DESCRIPTION
1	**None**	No particular strategy is adopted for the breeding of animals. Animals are allowed to mate
2	**Natural mating, sire selected from within herd**	Animals mate naturally, but specific male(s) are selected from within the herds for mating/reproduction.
3	**Natural mating, sire purchased**	Animals mate naturally, but specific male(s) are purchased specifically for mating/reproduction.
4	**Natural mating, sire exchanged**	Animals mate naturally, but male(s) are exchanged with other households specifically for mating/reproduction.
5	**Artificial insemination**	The technique of placing semen from the male into the reproductive tract of the female by means other than natural service.
6	**Dam purchased**	Animals mate naturally, but specific female(s) are purchased specifically for mating/reproduction.
7	**Dam exchanged**	Animals mate naturally, but specific female(s) are exchanged with other households specifically for mating/reproduction.
8	**Non-breeding males castrated**	Males that are not preferred for breeding/reproduction are castrated to prevent them from mating.

Appendix I: List of Definitions

GULIO (local market)
Gulio is a local market mostly at the Village level which can occur at any frequency – once a week, twice etc. One can get any sort of goods at a Gulio and it mostly operates in rural areas. Farmers get together at a certain place once a week and sell their produce. Note that these sellers travel across the country – it could be that every Monday they operate in Village A, every Tuesday in Village B, every Wednesday in Village C and so on.

SOKO KUU (main market)
Soko kuu is the main market that people go to that usually operates daily and they operate at the village or ward level. Soko Kuu is the primary market for all goods. The difference between Soko Kuu and Gulio is that Soko Kuu operates daily while Gulio operates on a fixed schedule basis, which is why they are referred to as the Main Market.

SOKO (market)
Soko is a small market located at street corner that sells few items like vegetables, and some other household goods. Soko's are frequently found in Dar for example. Usually these markets do not carry high value goods and have fewer items than a Soko Kuu.

MNADA (auction)
Mnanda is an English auction for a very specific good – like a cow for example. A minimum price is set, beyond which the bidders can go up to any amount. The highest bidder wins. It is an open auction in the sense that all the bidders find out how much each is bidding and can competitively bid more.

M/BIASHARA BINAFSI (private business person)
This is usually a vendor or a hawker who sells goods by walking door to door – like cigarette/water sellers or vegetable vendors in Dar.

DUKANI/MCHUUZI (grocery local merchant)
This is a small shop owner around street corners that sell very specific items. This could be a shop for stationery, shop for buying everyday essentials like bread, toothpaste, etc… or even a medical shop. Dukani/Mchuuzi sell very specific goods and are fixed shops owned by merchants unlike the Biasharas who are travellers.

Appendix J: Anthropometry

One approach to studying nutrition is to assess nutritional status on the basis of anthropometric indicators. By comparing these indicators with the distribution of the same indicator for a "healthy" reference group, and identifying "extreme" or "abnormal" departures from this distribution, it is possible to assess the adequacy of growth and diet, in particular in infants and children.

While a number of anthropometric indicators exist based on various physical body measurements, the assessment of nutritional status is presented in the NPS 2020/21 in terms of height and weight, as these are widely applied measurements and allow highly-specific and broadly-accepted interpretations. In the NPS 2020/21, the height and weight of respondents are collected in "HH_SEC_V" for all household members under the age of 15 and women of child-bearing age (15-49 years), while upper arm circumference is additionally collected for children under 5 years of age.

Using height and weight, objective measurements of body dimensions and composition are generated as proxy indicators of nutritional status. The supplemental dataset "NPSY5.CHILD.ANTHRO" presents the most frequently used anthropometric indices used to assess nutritional status in infants and children — height-for-age, weight-for-age, and weight-for-height — for all household members ages 0-59 months. While the latter indicator requires only physical measurements, the former two rely on the accurate and credible reporting of age.

Height-for-age

Height-for-age is a measure of cumulative linear growth, and deficits typically indicate chronic malnutrition and/or chronic or frequent illness. Moderate to severe cases of low height-for-age (<-2 SDs) are referred to as "stunting". The height-for-age index is primarily used as a population indicator rather than for individual growth monitoring.

Weight-for-height

Weight-for-height reflects an individual's body weight relative to height, and has the advantage of not requiring sensitive and often unreliable age data. Weight-for-height is a useful index for assessing *current* nutritional status, and since it is particularly sensitive to acute growth disturbances can be useful for screening children at risk and for measuring short-term changes in nutritional status. Moderate to severe cases of low weight-for-height (<-2 SDs) are referred to as "wasting". Wasting may be the consequence of starvation or severe disease (i.e. diarrhea), but can also be due to chronic conditions.

Weight-for-age

Weight-for-age measures body mass relative to age and, in effect, is a composite measure of height-for-age and weight-for-height. Moderate to severe low weight-for-age (<-2 SDs), relative

to a child of the same sex and age in the reference population, is referred to as “underweight”. Weight-for-age is commonly used for monitoring growth and to assess changes in the magnitude of undernutrition over time; however, the weight-for-age index confounds the effects of short- and long-term health and nutrition problems.

Anthropometric indices for the NPS 2020/21 were constructed by comparing relevant measures with those of comparable individuals (in regard to age and sex) in the reference population of “healthy” infants and children. This was done for infants and children ages 0-59 months. The preferred and most common way of expressing anthropometric indices is in the form of Z-scores, a dimensionless quantity derived by dividing the difference between the value for an individual and the mean value of the reference population for the same sex and age (or height) by the standard deviation of the reference population.

The general malnutrition classification proposed by the WHO distinguishes between mild (Z-score <-1 SDs), moderate (Z-score <-2 SDs), and severe malnutrition (Z-score <-3 SDs). Assessments of nutritional status are also commonly characterized by each indices’ respective binary indicator, defined by the severity of the anthropometric deficit as measured by Z-scores. Using this method on a normalized distribution curve, the most appropriate cut-off point to define abnormal anthropometry is -2 SDs, regardless of the indicator. For example, a child with a height-for-age Z-score <-2 SDs would be considered stunted, while a child with a height-for-age Z-score <-1 SDs would not be considered stunted.

The supplemental dataset “NPSY5.CHILD.ANTHRO” provides Z-scores for each of the anthropometric indices identified above, as well as a binary variable indicating severity of the anthropometric deficit (<-2 SDs or not). For many purposes, anthropometric data should be presented according age and sex groups. As such, the age in months of each respondent has also been provided in the supplemental dataset “NPSY5.CHILD.ANTHRO”, as well as “y5_hhid” and “indidy5”, which together serve as the unique identifier for the supplemental dataset as well as other individual-level household data file.