

# **Reference Manual**

Africa Region Poverty Program for Harmonized  
Household Surveys (SHIP)

The World Bank  
AFR  
1818 H Street, N.W.  
Washington, D.C.  
20433, USA.

November 2011

## Table of Contents

<b>0.</b>	<b>Chapter 0: Harmonized Household Surveys</b>	<b>4</b>
<b>1.</b>	<b>Chapter 1: Module I – Harmonized Individual-Level File</b>	<b>5</b>
1.1.	<i>Sample, Geography and Basic Household Identifier</i>	5
1.2.	<i>Basic Demographic Characteristics</i>	5
1.3.	<i>Literacy and Education</i>	7
1.4.	<i>Health</i>	9
1.5.	<i>Child Survival and Anthropometry for Children Aged 0-59 Months</i>	9
<b>2.</b>	<b>Chapter 2: Module II – Harmonized Labor Variables</b>	<b>11</b>
2.1.	<i>Identification, Individual, Educational and Training Variables Relevant to Labor</i>	13
2.2.	<i>Labor-force: vocational training</i>	15
2.3.	<i>Screening questions and primary employment</i>	16
2.4.	<i>Secondary Employment</i>	27
<b>3.</b>	<b>Chapter 3: Module III – Standardized Household-Level File</b>	<b>31</b>
3.1.	<i>Sample, Geography and Basic Household Identifier</i>	31
3.2.	<i>Household Head Characteristics</i>	32
3.3.	<i>Housing and Utilities</i>	33
3.4.	<i>Ownership of Durable and Non-Durable Goods</i>	36
3.5.	<i>Household Expenditure</i>	38
3.6.	<i>Household Remittances</i>	38
3.7.	<i>Household Income Variables</i>	39
<b>4.</b>	<b>Chapter 4: Module IV – Harmonized Expenditure File</b>	<b>40</b>
4.1.	<i>Table 0: Household Composition and Weights Summary Table</i>	42
4.2.	<i>Table 1: Food Expenditures</i>	42
4.3.	<i>Table 2: Alcoholic Beverages, Tobacco and Narcotics</i>	43
4.4.	<i>Table 3: Clothing and Footwear</i>	43
4.5.	<i>Table 4: Housing, Water, Electricity, Gas and Other Fuels</i>	44
4.6.	<i>Table 5: Furnishings, Household Equipment and Routine Household Maintenance</i>	44
4.7.	<i>Table 6: Health Care Expenditures</i>	45
4.8.	<i>Table 7: Transport Expenditures</i>	45
4.9.	<i>Table 8: Communication Expenditures</i>	46
4.10.	<i>Table 9: Recreation and Culture Expenditures</i>	46
4.11.	<i>Table 10: Education Expenditures</i>	46
4.12.	<i>Table 11: Restaurants and Hotels</i>	48
4.13.	<i>Table 12: Miscellaneous Goods and Services</i>	48
4.14.	<i>Table 13: Regional Price Deflators</i>	49

4.14.1.	Sub-File	49
4.14.2.	Final File	49
4.15.	<i>Table 14: Total Annual Household Expenditures</i>	50
5.	Chapter 5: Documentation and Computer Programs	52
6.	Appendices	53
6.1	<i>Appendix 1: Naming Convention</i>	53
6.2	<i>Appendix 2: Directory Organization</i>	53
6.3	<i>Appendix 3: Standard Excel Spreadsheet</i>	54
6.4	<i>Appendix 4: Issues In Household Expenditure Aggregation</i>	56
7	Annexure	59
7.1	<i>Annexure 1: Country Codes</i>	59
7.2	<i>Annexure 2: International Standard Industrial Classification (ISIC)</i>	60
	References	71

For any further inquiries please email [xye@worldbank.org](mailto:xye@worldbank.org)

## 0. Chapter 0: Harmonized Household Surveys

Based on Monitoring and Evaluation (M&E) needs, each year many countries conduct new household surveys in Africa. To ensure the quality and transparency of the data, it is critical to document the procedures of compiling consumption aggregation and other indicators so that the results can be duplicated with ease. This process enables consistency and continuity that make temporal and cross-country comparisons consistent and more reliable.

This manual presents detailed guidelines for compiling household survey data into a set of most commonly available variables from various types of household surveys. However, users also need to consult other literature on household consumption aggregation.<sup>1</sup>

Four harmonized data files will be prepared for each survey to generate a set of harmonized variables that have same variable names, definitions and the procedures of compiling these variables will be extensively documented. Invariably, in each survey, questions will be asked in a slightly different way, which poses challenges on consistent definition of harmonized variables. The harmonized household survey data present the best available variables with harmonized definition, but not identical variables. The three harmonized data files

- a) **Module I: Individual level file** (Labor force indicators in a separate file): This file has information on basic characteristics of individuals such as age and sex, literacy, education, health, anthropometry and child survival.
- b) **Module II: Labor force file**: This file has information on labor force including employment/unemployment, earnings, sectors of employment, etc.
- c) **Module III: Household level file**: This file has information on household expenditure, household head characteristics (age and sex, level of education, employment), housing amenities, assets, and access to infrastructure and services.
- d) **Module IV: Household Expenditure file**: This file has consumption/expenditure aggregates by consumption groups according to Purpose (COICOP) of Household Consumption of the UN.

All files should be named according to this manual (Refer Appendix 1: Naming Convention). All data editing checks (cleaning and assumptions) should be documented in the computer programs as comments, and the computer programs should be attached in the appropriate sequence. In addition, five sub-directories should be created under the main directory named after each survey (refer Appendix 2: Organization of Directories).

---

<sup>1</sup> Deaton, Angus and Salman Zaidi (2002), Guidelines for Constructing Consumption Aggregates for Welfare Analysis, Living Standards Measurements Study Working Paper Number 135, The World Bank: Washington DC.

## 1. Chapter 1: Module I – Harmonized Individual-Level File

This module consists of information on individuals in the household and covers approximately forty quantitative indicators. The information is organized into 5 sections that provide basic demographic and other socio-economic data such as age, sex and marital status of individuals, occupation, child malnutrition etc. among others..

### 1.1. Sample, Geography and Basic Household Identifier

No	NAME	LABEL AND CODES	COMMENTS, QUALITY CONTROL CHECKS AND CORRECTIONS
1	COUNTRY	<b>Country code</b> <i>Refer Annexure 1 for country codes.</i>	String variable, of 3 character length.  To be merged from Expenditure-level file.
2	REGION	<b>Region code</b>	Codes vary by country. To be merged from Expenditure-level file.
2b	PROVINCE	<b>Province code</b>	Codes vary by country. To be merged from Expenditure-level file.
2c	DISTRICT	<b>District code</b>	Codes vary by country. To be merged from Expenditure-level file.
3	HID	<b>Household unique identification</b>	String variable, length varies by country. To be merged from Expenditure-level file.
4	SURVEYR	<b>Year of survey</b>	To be merged from Expenditure-level file.
5	RURURB	<b>Area of residence</b> <i>1 = Rural</i> <i>2 = Urban</i>	To be merged from Expenditure-level file.
6	HHSIZE	<b>Household members</b>	Total # of residents (regular members) in the household, excluding maids and servants
6b	HHSIZE_S	<b>Household members including servants</b>	Total # of residents (regular members) in the household, including maids and servants
7	WTA_HH	<b>Individual weighting coefficient</b>	Country-specific. Weighting co-efficient to derive individual level indicators. To be merged from Expenditure-level file.
8	WTA_POP	<b>Population weighting coefficient</b> $WTA\_POP = WTA\_HH * HH\ SIZE$	Weighting co-efficient to derive population-level indicators.

### 1.2. Basic Demographic Characteristics

No	NAME	LABEL AND CODES	COMMENTS, QUALITY CONTROL CHECKS AND CORRECTIONS
11	INDID	<b>Individual identification</b>	This variable uniquely identifies the members of a household. Sequentially numbered from 1 for the household head to <i>N</i> (household size).

12	<b>SEX</b>	<b>Sex</b> <i>1 = Male</i> <i>0 = Female</i>	Sex of the individual
13	<b>RELAT</b>	<b>Relationship to household Head</b> <i>1 = Head</i> <i>2 = Spouse</i> <i>3 = Child</i> <i>4 = Father/Mother</i> <i>5 = Grandchild</i> <i>6 = Son-in-law/daughter-in-law</i> <i>7 = Other relative</i> <i>8 = Domestic help/paying boarder</i> <i>9 = None relative</i>	This refers to the relationship of each household member to the household HEAD. Each household must have only <b>one</b> head.
14	<b>HEAD</b>	<b>Household Head</b> <i>0 = No</i> <i>1 = Yes</i>	Household head is assigned based on reported household head status.  The head of household is defined by the household itself.  Check to ascertain that a household has only one unique household head
15	<b>MARSTAT</b>	<b>Marital status</b> <i>1 = Never married</i> <i>2 = Married monogamous</i> <i>3 = Married polygamous</i> <i>4 = Common law, union coutumiere, union libre, living together</i> <i>5 = Divorced/Separated</i> <i>6 = Widowed</i>	If polygamous status is not reported, check to find polygamous status in a household by counting the number of spouses in a household
16	<b>SP_PRE</b>	<b>Spouse present</b> <i>1 = Yes</i> <i>0 = No</i>	Code on the basis of a question that asks Whether or not the spouse lives in the household.  Some surveys instead ask for length of absence. If absence is more than 6 months code as 0 (No).
17	<b>AGEY</b>	<b>Age in completed years</b>	Ages >= 99 must be coded as 98. <i>Missing values</i> should be left as missing. For children <= 60 months, check consistency with age in months to get correct age in completed years.
18	<b>ETHNICITY</b>	<b>Ethnicity</b>	Codes vary by country.
19	<b>LANGUAGE</b>	<b>Language usually speak</b>	Codes vary by country.
110	<b>RELIGION</b>	<b>Religion</b>	Codes vary by country.

### 1.3. Literacy and Education

No	NAME	LABEL AND CODES	COMMENTS, QUALITY CONTROL CHECKS AND CORRECTIONS
I11	<b>LITERACY</b>	<b>Literacy status</b> 1 = <i>Can read and write</i> 2 = <i>Cannot read or write</i> 3 = <i>Cannot be determined</i>	For people aged 5 and over only. Value must be missing for all others.
I12	<b>EVERATTD</b>	<b>Ever attended school</b> 1 = Yes 0 = No	Country-specific. All persons aged 5 and above. Value must be missing for others.
I13	<b>ATSCHOOL</b>	<b>Currently enrolled in school</b> 1 = Yes 0 = No	. Code using the question that explicitly asks for enrollment over the past 12 months. Do not use the question that asks for current attendance. Use current attendance only if enrollment question is missing. In such surveys record this in the comments.  Code as 0 if EVERATTD is 0. Code as system missing if EVERATTD is missing.
I14	<b>EDLEVEL_AR</b>	<b>Highest level of education completed</b> 0 = <i>No education</i> 1 = <i>Pre-school/ Primary not completed</i> 2 = <i>Completed primary, but less than completed lower secondary</i> 3 = <i>Completed lower secondary (or post-primary vocational education) but less than completed upper secondary</i> 4 = <i>Completed upper secondary (or extended vocational/technical education)</i> 5 = <i>Post secondary technical</i> 6 = <i>University and higher</i> 7 = <i>Formal adult education or literacy program</i> 9 = <i>Other</i>	Code for all individuals who were administered the Education section.  If a person is currently in school and in the highest year of a level, then his/her highest level of education completed should be pushed down by one level. For example, if a person is currently enrolled in P6, then his/her level should be coded as 1(Pre-school/ Primary, not completed).  Individuals enrolled in higher/ University level are coded as 6 (University and higher) regardless of whether it is completed or not.  If there are more than 1 percent of missing values please see supervisor
I15	<b>EDYEARS</b>	<b>Years of completed education</b> 0 = pre-school 1 = grade 1 2 = grade 2 .	Code for all individuals who were administered the Education section.  For individuals who are currently enrolled in school, their years of education completed correspond to the class <i>currently attending</i>

		<p>.</p> <p>.</p> <p>99 = unspecified/incomparable</p>	<p><i>minus one.</i></p> <p>For individuals who are not currently enrolled in school, the years of completed education corresponds to the highest level of education completed.</p> <p>This is a continuous variable of the number of years of formal schooling completed. It is constructed only if the survey asked number of year of education or highest grade level completed; otherwise, the values are constructed as missing. The years of education that each grade corresponds to varies by country, for example - some countries may have 5 or 6 years of primary school, 3 years of lower-secondary school, while other countries may have 4 years of primary school and 4 years of lower-secondary school,.</p> <p>For higher education, the grades/years may not have been asked explicitly in the survey questionnaire. In such cases The variable should be constructed based on the following assumptions.</p> <ol style="list-style-type: none"> <li>1) If the individual has completed the tertiary education specified, add to years of completed education - 4 years for BA/BSc, 6 years for MA/MSc, and 8 Years for PhD after the completion of secondary education.</li> <li>2) If the individual has not completed tertiary education or completion cannot be ascertained, add to years of completed education – 2 years for BA/BSc , 5 years for MA/MSc , and 7 years for PhD.</li> </ol> <p>The variable does not take into account the actual number of years required to reach this grade level, in other words, first grade repeated three times only counts as 1 year of schooling.</p>
116	<b>EDYEARS_AT</b>	<p><b>Years of education for currently enrolled</b></p> <p>0 = pre-school</p> <p>1 = grade 1</p> <p>2 = grade 2</p> <p>.</p> <p>.</p>	<p>Code this variable only for individuals currently enrolled in school. Otherwise code as missing.</p> <p>Code years of education corresponding to the current grade being attended.</p>



		. 99 = unspecified/incomparable	
117	<b>SCHLTYP</b>	<b>Type of school attending</b> <i>1 = Public</i> <i>2 = Private</i> <i>9 = Other</i>	Code only for individuals currently attending school (ATSCHOOL=1).  Value must be missing for others.

#### 1.4. Health

No	NAME	LABEL and CODES	COMMENTS, QUALITY CONTROL CHECKS AND CORRECTIONS
118	<b>MORB_TR</b>	<b>Time reference for variable MORBID (weeks)</b>	Number of weeks. By convention, 1 month = 4 weeks
119	<b>MORBID</b>	<b>Morbidity last MORB_TR</b> <i>1 = Yes</i> <i>0 = No</i>	Has the person been sick or injured in the past X weeks? See time reference (MORB_TR) in 19a.
120	<b>RTREATM</b>	<b>Sought medical attention</b> <i>1 = Yes</i> <i>0 = No</i>	Only for sick individuals (MORBID=1).
121	<b>HLTHFAC</b>	<b>Health provider visited</b> <i>1 = Hospital</i> <i>2 = Clinics</i> <i>3 = Dispensary, health center, health post</i> <i>4 = Outreach</i> <i>5 = Pharmacy</i> <i>6 = Traditional healer</i> <i>9 = Other</i>	Only for individuals with RTREATM=1  Code for all individuals who reported it in the survey irrespective of whether they sought medical attention.
122	<b>OWHPROVD</b>	<b>Ownership of health provider visited</b> <i>1 = Public</i> <i>2 = Private - modern medicine</i> <i>3 = Private - traditional healers</i> <i>4 = Missionary/NGO</i> <i>9 = Other</i>	Code only for all individuals who visited a health care facility.

#### 1.5. Child Survival and Anthropometry for Children Aged 0-60 Months

No	NAME	LABEL AND CODES	COMMENTS, QUALITY CONTROL CHECKS AND CORRECTIONS
123	<b>AGEM</b>	<b>Age of child in completed</b>	<i>Only for children aged 0 - 60 months; otherwise code as missing</i>

			<i>Missing values</i> should be left as missing.
I24	<b>WHDELIV</b>	<b>Who helped assist Childbirth</b> 1 = <i>Trained Birth Attendant</i> 2 = <i>Traditional Birth Attendant</i> 3 = <i>Relative/friend</i> 9 = <i>Other</i>	For children aged 0-60 months.
I25	<b>VCARD</b>	<b>Child has vaccine card</b> 1 = <i>Yes</i> 2 = <i>No, but child has been vaccinated</i> 3 = <i>No, child never vaccinated</i>	
I26	<b>BCG</b>	<b>Child immunized against BCG</b> 1 = <i>Yes</i> 0 = <i>No</i> 9 = <i>Don't know</i>	For children aged 0-60 months.  BCG is the vaccine against tuberculosis.
I26	<b>POLIO_DO</b>	<b>Number of doses of Polio received</b> 0 = <i>No dose</i> 1 = <i>One dose</i> 2 = <i>Two doses</i> 3 = <i>Three doses</i> 4 = <i>More than three doses</i> 9 = <i>Unspecified number of doses</i>	For children aged 0-60 months.  Also referred to as OPV.  If Polio given but number of doses unknown or not specified, code 9.
I27	<b>DPT_DOSE</b>	<b>Number of doses of DPT received</b> 0 = <i>No dose</i> 1 = <i>One dose</i> 2 = <i>Two doses</i> 3 = <i>Three doses</i> 4 = <i>More than three doses</i> 9 = <i>Unspecified number of doses</i>	For children aged 0-60 months.  Three doses refer to full immunization against Diphtheria, Pertussis and Tetanus.
I28	<b>MEASLES</b>	<b>Child immunized against measles</b> 1 = <i>Yes</i> 0 = <i>No</i> 9 = <i>Don't know</i>	For children aged 0-60 months.
I29	<b>VACCINE</b>	<b>Immunization status of Child</b> 1 = <i>Fully immunized</i> 2 = <i>Partially immunized</i> 3 = <i>Immunized but no</i>	For children aged 0-60 months. <i>Fully immunized</i> refers to:- (a) BCG (1 or more doses). (b) Three or more doses of Polio (1-2-3). 3 or more doses may include Polio at birth

		<i>information on completeness</i> <i>4 = Not immunized</i>	(Polio-0). (c) Three or more doses of DPT (1-2-3) and (d) Measles (1 or more doses).
I30	<b>WEIGHT</b>	<b>Weight of Child (kg)</b>	For children aged 0-60 months.
I31	<b>HEIGHT</b>	<b>Height of Child (cms)</b>	For children aged 0-60 months. By convention 1 cm = 10 mms 1 inch = 2.54 cms.

**Note:**

- a) For any variable not collected in a country, variable should be created and left as system missing (.) in the final harmonized aggregated expenditure file.
- b) Variables in the data files must follow the sequence in which they appear in the manual..

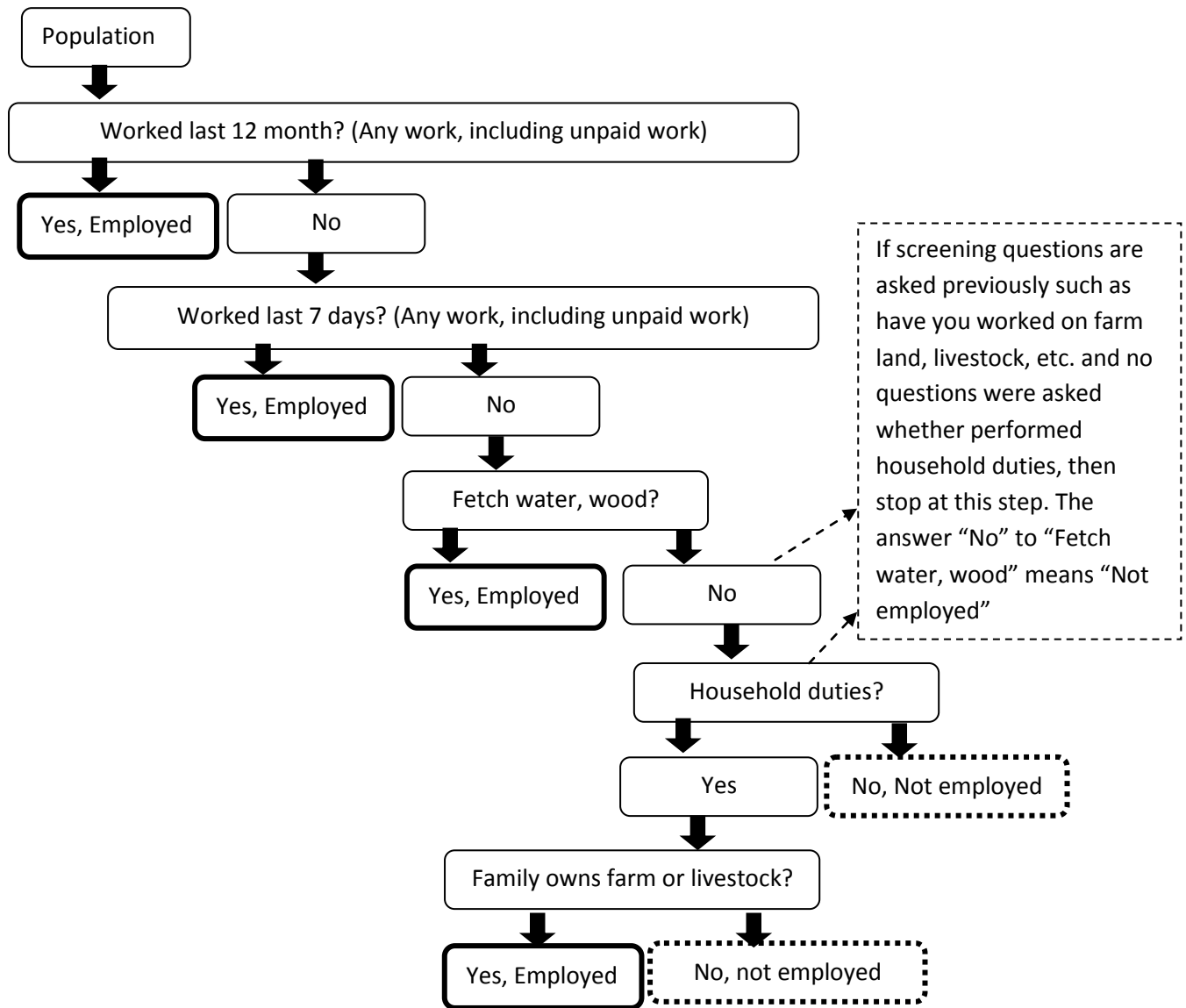
## 1. Chapter 2: Module II – Harmonized Labor Variables

**Note:** Informal labor participation is a norm among African countries and evidences have shown that the informal workers often under report their economic activities, especially for individuals who report domestic duties as their main activity. These individuals inevitably undertake some unpaid economic activity such as cultivating, raising livestock, fetching water, collecting wood, etc. (Preparing meals for the family and caring for own children are not classified as economic activities by ILO standards.)

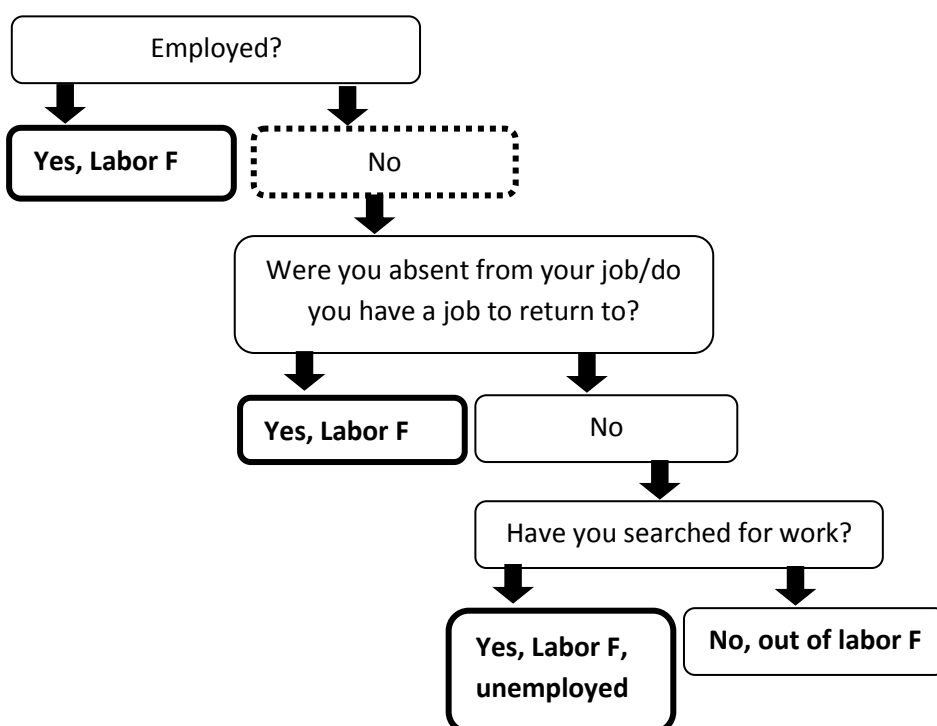
This manual tries to reclassify the employment status of these individuals who claim household duties as their main activity . Because labor force questionnaires are significantly different from one another, it is not possible to provide a set of very specific steps that one can follow to classify employment status in Africa. The diagram below illustrates the logic used to classify unpaid economic activities. This logic is based on a research on underreporting labor participation in Tanzania. It found that many unpaid family workers, especially women, under report their economic activities due to poor questionnaire design. For details refer to Bardasi, Beegle, Dillon and Serneels "Do Labor Statistics Depend on How and to Whom the Questions Are Asked", World Bank Policy Research Working Paper 5192.

Additionally, because informal economic activities are so common, there are often missing values for the sectors of employment. To remedy this situation we designed a variable called "EMPFRM\_1" and "EMPFRM\_2" classifying primary and secondary employment respectively into two categories "Farm" and "Nonfarm" activities. Everyone who had a job in last year or last 7 days, must be classified into either Farm or Nonfarm activities using any information available in the survey that might help in this classification. (less than one percent of missing values are allowed).

## Definition of Employment in Africa Region



## Definition of Unemployment and Labor Force in Africa Region



*(Variables with sequence number starting with L are labor force variables; other variables are imported from other Individual/ household level SHIP files. Generate for all individuals)*

### 2.1. Identification, Individual, Educational and Training Variables Relevant to Labor

No	NAME	LABEL AND CODES	COMMENTS, QUALITY CONTROL CHECKS AND CORRECTIONS
1	<b>COUNTRY</b>	<b>Country code</b>	To be merged from Expenditure-level file.
2	<b>REGION</b>	<b>Region code</b>	To be merged from Expenditure-level file.
2b	<b>PROVINCE</b>	<b>Province code</b>	To be merged from Expenditure-level file.
2c	<b>DISTRICT</b>	<b>District code</b>	To be merged from Expenditure-level file.
3	<b>HID</b>	<b>Household unique identification</b>	To be merged from Expenditure-level file.
4	<b>SURVEYR</b>	<b>Year of survey</b>	To be merged from Expenditure-level file.
5	<b>RURURB</b>	<b>Area of residence</b> 1 = Rural 2 = Urban	To be merged from Expenditure-level file.
7	<b>WTA_HH</b>	<b>Individual weighting coefficient</b>	To be merged from Expenditure-level file.  To obtain population estimates, this is the weight to be used in all calculations referring to individuals.
11	<b>INDID</b>	<b>Individual</b>	To be merged from Individual level file

		<b>identification</b>	
I2	<b>SEX</b>	<b>Sex</b> 1 = <i>Male</i> 0 = <i>Female</i>	To be merged from Individual level file
I3	<b>RELAT</b>	<b>Relationship to household Head</b> 1 = <i>Head</i> 2 = <i>Spouse</i> 3 = <i>Child</i> 4 = <i>Father/Mother</i> 5 = <i>Grandchild</i> 6 = <i>Son/daughter-in-law</i> 7 = <i>Other relative</i> 8 = <i>Domestic help/paying boarder</i> 9 = <i>None relative</i>	To be merged from Individual level file.
I4	<b>HEAD</b>	<b>Household Head</b> 0 = <i>No</i> 1 = <i>Yes</i>	To be merged from Individual level file.
I5	<b>MARSTAT</b>	<b>Marital status</b> 1 = <i>Never married</i> 2 = <i>Married monogamous</i> 3 = <i>Married polygamous</i> 4 = <i>Common law, union coutumière, union libre, living together</i> 5 = <i>Divorced/Separated</i> 6 = <i>Widowed</i>	To be merged from Individual-level file.
I7	<b>AGEY</b>	<b>Age in completed years</b>	To be merged from Individual-level file.
I11	<b>LITERACY</b>	<b>Literacy status</b> 1 = <i>Can read and write</i> 2 = <i>Cannot read or write</i> 3 = <i>Cannot be determined</i>	To be merged from Individual level file
I12	<b>EVERATTD</b>	<b>Ever attended school</b> 0 = <i>No</i> 1 = <i>Yes</i>	To be merged from Individual level file.
I14	<b>EDLEVEL_AR</b>	<b>Education level</b> 0 = <i>No education</i> 1 = <i>Pre-school Primary, incomplete</i> 2 = <i>Primary</i>	To be merged from Individual level file

		<i>completed but less than lower secondary completed</i> 3 = Lower secondary completed (or post-primary vocational education) but less than upper secondary completed 4 = Completed upper secondary 5 = University and higher 6 = Any post secondary technical education 7 = Adult education or literacy program 9 = Other	
I15	<b>EDYEARS</b>	<b>Years of education</b> 0 = pre-school 1 = grade 1 2 = grade 2. . 99 = unspecified / incomparable	To be merged from Individual level file.
E65	<b>FDPINDEX</b>	<b>Regional food price deflators</b>	To be merged from expenditure level file
E70	<b>CPI2005_DEF</b>	<b>CPI Deflator</b>	To be merged from expenditure level file
E71	<b>PPP2005</b>	<b>2005 PPP exchange rate</b>	To be merged from expenditure level file

## 2.2. Labor-force: vocational training

No	NAME	LABEL AND CODES	COMMENTS, QUALITY CONTROL CHECKS AND CORRECTIONS
L1	<b>EMPLEDU</b>	<b>Post school vocational training</b> 0 = No 1 = Yes	Code as 1 (YES) if the individual has ever attended or is currently attending post school vocational training, otherwise code as 0 (NO).
L2	<b>SVYAGE</b>	<b>Lower age cut off for employment svy</b>	The lower age cutoff for individuals who were administered the employment section.
L3	<b>EMPLEDU_TU</b>	<b>Time Unit of course duration</b> 1=Days 2=Weeks 3=Months 4=Years	Categorical variables to capture the unit of time for the duration of the course.
L4	<b>EMPLEDU_TIME</b>	<b>Number of EMPLEDU_TU</b>	Continuous variable to capture the duration of the course.
L5	<b>EMPLEDU_CER</b>	<b>Highest certificate</b>	The name of the certificate earned. Codes

		<b>obtained</b>	will vary with the country
L6	<b>EMPLEDU_SCHLTYP</b>	<b>Type of Training Institute</b> 1=Public 2=Private 9=Other	Code only if EMPLEDU=1.  Leave as missing if EMPLEDU is not equal to 1.
L7	<b>APPREN</b>	<b>Apprenticeship</b> 0 = No 1 = Yes	Code as 1 (YES), if the individual <b>HAD</b> participated in an apprenticeship.  <i>Note: If currently working as an apprentice, <b>DO NOT</b> code as 1 here (code as 0). It will be included in variable EMPTYTYPE_WB_1 and/or _2</i>

### 2.3. Screening questions and primary employment

No	NAME	LABEL AND CODES	COMMENTS, QUALITY CONTROL CHECKS AND CORRECTIONS
L8	<b>WORKED_Y</b>	<b>Labor force screening <u>last year</u></b> 0 = No 1 = Yes 99=The question is not in the survey	Use 12 month employment section only.  Code as 1 (YES), if the individual performed an economic activity for cash or for in-kind payment or for barter or for self-consumption for any length of time during the past 12 month, otherwise code as 0 (NO).
L9	<b>WORKED _7</b>	<b>Labor force screening <u>last 7 days</u></b> 0 = No 1 = Yes 99= The question is not in the survey	Use 7 day employment section only.  Code as 1 (YES), if the individual performed an economic activity for cash or for in-kind payment or for barter or for self-consumption for any length of time during last 7 days, otherwise code as 0 (NO).  Code as 0 (No), if a person has a job but was absent for last 7 days or more.  Do not re-code missing values.
L10	<b>ABSENT</b>	<b>Absent for 7 days or more</b> 0 = No 1 = Yes 99 = The question is not in the survey	Use 7 day employment section only.  <b>If WORKED_7 = 1 then code as missing value</b> (skip the variable)  If WORKED_7 = 0, then code accordingly  Code as 1 (YES), if the person has a job but was absent from it for any reason (sick, other leave, temp. layoff, is about to start work etc.), otherwise code as 0 (NO).  Thus all individuals who did not work in the



			past 7 days and do not have a job to go back to are coded as 0(NO).
L11	<b>LOOKJOB_TU</b>	<b>Up to xx weeks</b>	<p>If both a 7 day and 12 month employment section is present in a survey use only the 7 day employment section to code this variable</p> <p>Record the reference time period (in weeks) for which the question is asked.</p> <p>If the question gives a range of reference period, take the longest time period up to 4 weeks.</p> <p>4 weeks is the time period defined by the ILO</p> <p><i>Note: (7 days = 1 week)</i></p>
L12	<b>LOOKJOB</b>	<b>Looked for job in last "LOOKJOB_TU" weeks</b>  <i>0 = No 1 = Yes 99 = The question is not in the survey</i>	<p>If both a 7 day and 12 month employment section is present in a survey use only the 7 day employment section to code this variable.</p> <p>Code as 1 (YES) if <b>BOTH</b> the following conditions are satisfied, otherwise code as 0 (NO).</p> <ol style="list-style-type: none"> <li>1. The individual did not perform an economic activity for cash or in-kind payment in the last 7 days.</li> <li>2. The individual looked for a job in the LOOKJOB_TU weeks.</li> </ol>
L13	<b>FETCHWOOD</b>	<b>Fetches wood for the household</b> <i>0= No 1= Yes 99 = The question is not in the survey</i>	<p>Code as 1 (YES) if the individual fetched wood for his/her own household, otherwise code 0 (NO).</p> <p>Based on UN definition of SSN, Fetching wood is an economic activity.</p>
L14	<b>FETCHWATER</b>	<b>Fetches water for the household</b> <i>0= No 1= Yes 99 = The question is not in the survey</i>	<p>Code as 1 (YES) if the individual fetched water for his/her own household, otherwise code 0 (NO).</p> <p>Based on UN definition of SSN, Fetching water is an economic activity.</p>
L15	<b>NOTLAB_WB</b>	<b>Reason not in labor force</b> <i>0 = In labor force 1 = Household duties 2 = Retired 3 = Student 4 = Dependent (too old or too young to</i>	<p>Use information from all sections (12 months or 7 days) that ask this question.</p> <p>First use information from the 12 month section. For all missing values and observations coded as 9, use the 7 day section.</p>

		<p>work)  5 = <i>Disabled/ill</i>  6 = <i>Discouraged worker</i>  9 = <i>Other</i></p>	<p>Reason for not being in the labor force:</p> <p>If LABFORCE_WB=1 then code as 0, otherwise if LABFORCE_WB=0 then code according to responses.</p> <p>Code as 4 (Dependent), if the response is 'dependent' or 'too old' or 'too young' or if the individual is <b>below 5 years of age</b>.</p>
L16	EMPLHOME	<p><b>Employment in household duties in own house</b></p> <p>0 = No  1 = Yes  99= The question is not in the survey</p>	<p>This variable is created to capture those individuals who were incorrectly classified as out of the labor force because in the questionnaire design, they had an option of reporting 'domestic duties' as their activity.</p> <p>Construct this variable in 2 stages:-</p> <p><b>Stage1: Identifying individuals engaged in household duties</b></p> <p>An individual is classified as engaged in household duties if <b>ALL</b> the following conditions are met:-</p> <ol style="list-style-type: none"> <li>1. Did not work in the past 1 year (WORKED_Y != 1 )</li> <li>2. Did not work in the past 7 days (WORKED_7 != 1 )</li> <li>3. Was not absent from work (ABSENT != 1 )</li> <li>4. Performed household duties (notlab_wb=1).</li> </ol> <p><i>Note: Performed household duties should be checked from the responses used to code NOTLAB_WB. If no question can be found to code NOTLAB_WB, use information from the section on household chores.</i></p> <p><b>Stage 2: Checking for relevant conditions to classify individuals identified above as EMLHOME</b></p> <p>If such an individual belongs to a household that carries out agricultural activities (farming/ sharecropping/ hunting/ fishing/ logging/forestry/raising livestock/bee-keeping).</p> <p>Then, such an individual should be considered as employed and coded as 1, otherwise code as zero.</p>

L17	EMPTYTYPE_WB_1	<p><b>Type of employment, primary job</b></p> <p>1 = Wage &amp; salaried worker  2 = Self-employed with employees  3 =Self-employed w/o employees  4= Employer  5 = Domestic employees (work for compensation for private house)  6 = Family worker  7= Apprentice  9 = volunteer  99=Other/ Not known/ Missing.</p>	<p>To generate this variable look for all available information in the following sequence:-</p> <ol style="list-style-type: none"> <li>1. First use the longest recall period (e.g. 12 month recall period) to code this variable.</li> <li>2. Replace the value of missing (coded 99) and volunteer/other (coded 9) based on information provided in sections with lesser recall periods (e.g. 7 day recall period)</li> <li>3. Replace the value of the missing (coded 99) and volunteer/other (coded 9) based on information provided in other sections of the survey that identify an individual with an activity in the survey sections on 'agriculture' and 'enterprise'. Look explicitly for individual level information (i.e. the person responsible or working at the farm/enterprise). Match individual IDs from enterprise/farm sections with those recorded in the individual level file.</li> <li>4. Replace the value of the missing (coded 99) and volunteer/other (coded 9) based on information provided at the household or farm or enterprise level. If the household is involved in any agricultural then code the household head as 3 (self-employed without employees). If the household owns a household enterprise, then code the household head as 2 (self-employed with employees or 3 (self-employed without employees) based on number of employees. If no information on the number of employees is available then code as 3 (self-employed without employees).</li> <li>5. Replace the value of the missing (coded 99) and volunteer/other (coded 9) based on EMPLHOME/ FETCHWOOD/ FETCHWATER. If EMPLHOME=1 OR FETCHWOOD=1 OR FETCHWATER=1 then code such individuals as 6 (Family worker).</li> </ol>
-----	----------------	--	--

			<p>In questionnaires that do not explicitly ask for primary and secondary jobs (eg. Malawi), ascertain the primary and secondary job by comparing and ranking the jobs in descending order of the time spent in each kind of job. The comparison should be made from all sections (labor, enterprise, agriculture etc.), that ask questions over the past 12 months. Thus the job which takes most work hours is the primary job.</p> <p>However, even while following this methodology the above-mentioned sequence must be maintained, i. e. 12 month questions must be given precedence over 7 day questions and so on.</p> <p><i>Category definitions</i></p> <ol style="list-style-type: none"> <li>1) <i>Wage and salaried worker</i> is an individual employed by others outside the household including workers who are paid on a daily or hourly basis or based on task.</li> <li>2) <i>Self-employed with employees</i> is an individual who runs a (farm or non-farm) household enterprise or farm that employs at least one non-household member for pay during the reference period.</li> <li>3) A <i>self-employed without employee</i> is an individual who works alone with no paid employees. May have other family members categorized as family or family unpaid workers. May have an apprentice. Sometimes coded as "own account" worker in questionnaires.</li> <li>4) An <i>employer</i> owns an incorporated business separate from the household, with paid employees. It could also be a large commercial farm. This category is coded only if reported as employer in the questionnaire.</li> <li>5) <i>Apprentice</i> may or may not be paid</li> <li>6) <i>Family worker</i> can be farm or non-</li> </ol>
--	--	--	---

			<p>farm. Note that family workers may be paid or unpaid.</p> <p>7) The term <i>household enterprise</i> refers to a non-farm business or home production activity that is run by household members. It need not be physically located at the household dwelling unit. It refers to both :–</p> <p>a) Businesses engaged in market based transactions and</p> <p>b) Activities designed to produce home consumption.</p>
L18	<b>EMP_IND_SVY_1</b>	<b>Industry classification for primary employment</b>	<p>Use information from all sections (12 months or 7 days) that ask this question.</p> <p>First use information from the 12 month section. For all missing values and observations coded as 9, use the 7 day section.</p> <p>Keep the original classification of the survey. Most countries follow the ISIC classification.</p>
L19	<b>INDUSTRY_1</b>	<p><b>1 digit Industry classification for primary employment</b></p> <p><i>1 = Agriculture and Fishing</i>  <i>2 = Mining</i>  <i>3 = Manufacturing</i>  <i>4 = Electricity and Utilities</i>  <i>5 = Construction</i>  <i>6 = Commerce</i>  <i>7 = Transportation, Storage and Communication</i>  <i>8 = Financial, Insurance and Real Estate</i>  <i>9 = Services: Public Administration</i>  <i>10 = Other Services</i>  <i>11 = Unspecified</i></p>	<p>Use information from all sections (12 months/7 days/any other section) that may provide information.</p> <p>First use information from the 12 month section. For all missing values and observations coded as 9, use the 7 day section.</p> <p>If data allows generation of EMP_IND_SVY_1, recode it to generate INDUSTRY_1 according to SHIP requirements (refer Annexure-2).</p> <p>If data does not follow ISIC classification system, make appropriate assumptions to classify employment industry. Carefully document assumptions.</p> <p>Individuals identified in any section on agricultural activity are coded as 1 (Agricultural and fishing).</p> <p>Individuals identified in a household enterprise section and where the industry of the enterprise is available, code them accordingly. In many household enterprise sections, the industry of the enterprise is not available. In such cases do not classify the</p>

			<p>individual in any industry.</p> <p>Heads of households engaged in agricultural activities are coded as 1 (Agriculture and Fishing)</p> <p>Heads of households owning household enterprises are coded according to the industry of the household enterprise, if this information is available; or left as uncoded if the industry of the enterprise is not available.</p> <p><b>Individuals identified as EMPLHOME=1 OR FETCHWOOD=1 OR FETCHWATER=1 are coded as 1.</b></p>
L20	<b>EMPFRM_1</b>	<p><b>Farm employment in primary job</b>  <i>0 = No</i>  <i>1 = Yes</i></p>	<p>Farming activities include cultivating, sharecropping, forestry, raising livestock, bee-keeping, fishing and hunting.</p> <p><b>If INDUSTRY_1=1 then code as 1 (YES).  If INDUSTRY_1=2 to 10 then code as 0 (No)</b></p> <p>If INDUSTRY_1 variable is missing or 11 (unspecified) , code EMPFRM_1 on the basis of all information available in any section of the survey.</p> <p>Individuals identified in household non-agricultural enterprises should be coded as 0(No). These individuals may have missing values for INDUSTRY_1, if the industry information of the enterprise was not available. Similarly heads of households owning non-agricultural enterprises should be coded as 0 (NO).</p> <p>As far as possible, EMPFRM_1 should be coded for all individuals coded as 1 to 9 in EMPTYPE_WB_1. Less than one percent of missing values are allowed for EMPFRM_1 if EMPTYPE_WB_1 has a meaningful value (i.e. NOT 99(missing)).</p>
L21	<b>EMPSEC_1</b>	<p><b>Employment sector of primary job</b>  <i>1 = Public Government, Army</i>  <i>2 = State Owned</i>  <i>3 = Private or NGO</i></p>	<p><b>Code only for EMPTYPE_WB_1=1, 4, 5, 7, 9 (wage employees), otherwise code as missing.</b></p> <p>Classifies the main job's sector of activity of any individual with a wage job.</p> <p>Use information from all sections (12 months or 7 days) that ask this question.</p>

			<p>First use information from the 12 month section. For all missing values and observations coded as 9, use the 7 day section.</p> <p><i>Armed forces and international institutions are included in Public sector.</i></p> <p><i>State owned</i> includes para-statal firms and all other enterprises in which the government has control.</p> <p><b>Note: if no such question leave as missing. This will mean that in the variable EMPCAT_2 wage jobs will not be further disaggregated. Do not code it on the basis of ISCO or ISIC codes.</b></p>
L22	CONTRACT	<p><b>Employment contract</b>  0 = No  1 = Yes</p>	<p><b>Code only for EMPTYTYPE_WB_1=1, 4, 5, 7, 9 (wage employees), otherwise code as missing.</b></p> <p>Indicates if a person has a signed (formal) contract, regardless of the duration.</p> <p>This variable is constructed only if there is an explicit question about contracts otherwise code as missing.</p> <p>Use information from all sections (12 months or 7 days) that ask this question. First use information from the 12 month section. For all missing values and observations coded as 9, use the 7 day section.</p>
L23	EMP_CAT_1	<p><b>Employment categories for primary job</b></p> <p><i>Wage worker</i>  11=wage public  12=wage private non-agriculture  13=wage private agriculture  <i>Self employment/ nonwage</i>  21=family enterprise non-agri  22=family farmer  <b>31=Not known</b></p>	<p><b>This is a derived variable from the EMPTYTYPE_WB_1, EMPSEC_1 and EMPFRM_1.</b></p> <p><b>If EMPTYTYPE_WB_1=1, 4, 5, 7, 9 then code as 1</b></p> <p>If EMP_CAT_1=1 and EMPSEC_1=1, 2 then code 11  If EMP_CAT_1=1 and EMPSEC_1= 3 and EMPFRM_1!=1 then code 12  If EMP_CAT_1=1 and EMPSEC_1= 3 and EMPFRM_1=1 then code 13</p> <p><b>If EMPTYTYPE_WB_1=2, 3, 6 then code as 2</b>  If EMP_CAT_1=2 and EMPFRM_1!=1 then code 21</p>

			<p>If EMP_CAT_1=2 and EMPFRM_1=1 then code 22</p> <p><b>If EMPTYTYPE_WB_1=10 then code as 31</b></p>
L24	<b>WAGE_1</b>	<p><b>Last wage payment in nominal local currency for primary employment</b></p>	<p><b>For EMPTYTYPE_WB_1=1, 4, 5, 7, 9 only</b> (wage employees).</p> <p>Use information from all sections (12 months or 7 days) that ask this question.</p> <p>First use information from the 12 month section. For all missing values and observations coded as 9, use the 7 day section.</p> <p>Wage from main job (job to which the person dedicated most time in the week preceding the survey).</p> <p>Note: This <b>includes tips, bonuses, and other cash payments. Do not include the monetary value of in-kind payments or perks</b> associated with the job</p> <p>By definition non-paid employees should have wage=0.</p>
L25	<b>UNITWAGE_1</b>	<p><b>Time unit for wage in primary employment</b></p> <p>1 = Hourly  2 = Daily  3 = Weekly  4 = Every two weeks  5 = Twice per month  6 = Monthly  7 = Quarterly  8 = Biannual  9 = Annually  10 = Other</p>	<p><b>Code only for EMPTYTYPE_WB_1=1, 4, 5, 7, 9 (wage employees), otherwise code as missing.</b></p> <p>Unit of time for the wage variable of the <b>main</b> job.</p> <p>Use information from all sections (12 months or 7 days) that ask this question.</p> <p>First use information from the 12 month section. For all missing values and observations coded as 9, use the 7 day section.</p>
L26	<b>WHOURS_1</b>	<p><b>Hours worked per week in primary employment</b></p>	<p>This is the usual number of hours worked per week in the <b>main</b> job.</p> <p>Use information from all sections (12 months or 7 days) that ask this question.</p> <p>First use information from the 12 month section. For all missing values and observations coded as 9, use the 7 day section.</p> <p>If only information on how many hours</p>



			<p>individuals worked per day is available but no information on number of days worked a week is available, multiply the hours by 5 days.</p> <p>Similarly for hours worked per month divide by 4.2 to get weekly hours.</p>
L27	<b>SOU_AGRI_1</b>	<b>Individual income primary employment self-employed Agricultural job</b> 0 = No 1 = Yes	<p>Code as 1(YES) if the source of income is from self-employed agriculture activities, otherwise code as 0 (NO).</p> <p><b>If EMP_CAT_1! =. Then code as 0</b>  <b>Replace code = 1 If EMP_CAT_1=13 or 22</b></p>
L28	<b>SOU_NFARM_1</b>	<b>Individual income primary employment self-employed non-agricultural job</b> 0 = No 1 = Yes	<p>Code as 1(YES) if the source of income is from self-employed, non-farming activities, otherwise code as 0 (NO).</p> <p><b>If EMP_CAT_1! =. Then code as 0</b>  <b>Replace code = 1 If EMP_CAT_1=21</b></p>
L29	<b>SOU_WAGE_1</b>	<b>Individual income primary employment non-agricultural wage job</b> 0 = No 1 = Yes	<p>Code as 1( YES) if the source of income is from non-agricultural wage job, otherwise code as 0 (NO).</p> <p><b>If EMP_CAT_1! =. Then code as 0</b>  <b>Replace code = 1 If EMP_CAT_1=11 or 12</b></p>
L30	<b>UNEMPLOYED_ILO</b>	<b>Unemployment ILO</b> 0 = No 1 = Yes	<p>Individuals are considered as UNEMPLOYED if they meet <b>ALL</b> the following conditions:-</p> <ol style="list-style-type: none"> <li>1. Did not work in the last 7 days (WORKED_7!=1) <b>and</b></li> <li>2. Were not absent from their job (ABSENT !=1) <b>and</b></li> <li>3. Looked for a job (LOOKJOB=1)</li> </ol> <p>Code as 0 (NO) if WORKED_7=1 <b>OR</b> ABSENT=1</p> <p>To exclude 'employed individuals who are engaged in household duties from unemployed please replace</p> <p><b>UNEMPLOYED_ILO=0 if EMPLHOME==1</b></p>
L31	<b>EMPLOYED_WB</b>	<b>Employment WB</b> 0 = No 1 = Yes	<p>This is a derived employment status that has the most inclusive definition.</p> <p>If LABFORCE_WB=1 then code 0;</p> <p>Replace code with 1 (YES) if WORKED_Y=1</p>

			<p><b>or</b> WORKED_7=1 <b>or</b> EMPLHOME=1 <b>or</b> ABSENT = 1.</p> <p>Replace code with 0 (NO) if EMPTYPE_WB_1 = 99.</p> <p>(If an individual worked in the last year <b>or</b> last 7 days <b>or</b> had a job to go back to <b>or</b> is an individual who fetched water/firewood/ <b>or</b> worked on family farm or raised livestock; then such an individual is considered employed.)</p> <p><b>Note: Do not miss-assign 0 to missing values. 0 should be coded if ANY of the variables use to derive this variable have non missing values. If ALL the values of the deriving variables are missing, then this variable should also be missing.</b></p>
L32	UNEMPLOYED_WB	<b>Unemployment WB</b> 0 = No 1 = Yes	<p>This is a derived unemployment status that has the most exclusive definition.</p> <p>If LABFORCE_WB =1 then code 0;</p> <p>Replace code with 1 (YES) if WORKED_Y! =1 and UNEMPLOYED_ILO = 1.</p> <p>Replace code with 0 (NO) if EMPTYPE_WB_1 != 99.</p> <p>(If an individual did not work in past one year <b>and</b> did not work in the past seven days <b>and</b> did not work on fetching water, firewood, etc. <b>and</b> looked for a job; then such an individual is considered unemployed).</p> <p><b>Note: Do not miss-assign 0 to missing values. 0 should be coded if ANY of the variables use to derive this variable have non missing values. If ALL the values of the deriving variables are missing, then this variable should also be missing.</b></p>
L33	LABFORCE_WB	<b>Labor-force Participation WB</b> 0 = No 1 = Yes	<p>This is a derived employment status that has the most inclusive definition.</p> <p>First Code 0 (NO) if WORKED_Y ==0 <b>or</b> EMPLOYED_7==0 <b>or</b> EMPLHOME==0 <b>or</b> ABSENT ==0 <b>or</b> UNEMPLOYED_ILO==0</p> <p>Replace with 1(YES) if WORKED_Y=1 <b>or</b> WORKED_7=1 <b>or</b> EMPLHOME=1 <b>or</b> ABSENT = 1 <b>or</b> UNEMPLOYED_ILO=1 <b>or</b></p>

			<p>EMPTYTYPE_WB_1 !=99 .</p> <p>(If an individual worked for last year <b>or</b> last 7 days <b>or</b> had a job to go back to <b>or</b> an individual fetched water/firewood/ <b>or</b> worked on family farm or raised livestock, <b>or</b> is looking for a job; then such an individual is considered to be in the labor force).</p> <p><b>Note: Do not miss-assign 0 to missing values. 0 should be coded if ANY of the variables used to derive this variable have non missing values. If ALL the values of the deriving variables are missing, then this variable should also be missing.</b></p>
L34	HOUR_NECON	Hours spent per week on non-economic activities.	<p>One month equals 4-weeks.</p> <p>These include activities such as preparing food, and care for children.</p> <p>Fetching wood and water are considered economic activities and should <b>not</b> be included here.</p>

#### 2.4. Secondary Employment

No	NAME	LABEL AND CODES	COMMENTS, QUALITY CONTROL CHECKS AND CORRECTIONS
L35	EMPTYTYPE_WB_2	<p><b>Type of employment</b>  <b>Secondary job</b>  1 = Wage &amp; salaried worker  2 = Self-employed with employees  3 = Self-employed w/o employees  4 = Employer  5 = Domestic employees (work for compensation for private house)  6 = Family worker  7 = Apprentice  9 = volunteer  99 = Other /Not known/ Missing</p>	<p>To generate this variable only look for information in the following sequence:-</p> <ol style="list-style-type: none"> <li>1. First use the longest recall period (e.g. 12 month recall period) to code this variable</li> <li>2. Replace the value of missing (coded 99) and volunteer/other (coded 9) based on information provided in sections with lesser recall periods (e.g. 7 day recall period)</li> </ol> <p>Do not look for additional information beyond the secondary job section.</p> <p>Do not code household heads or EMPLHOME or FETCHWOOD or FETCHWOOD as done in EMPTYTYPE_WB_1.</p>

L36	<b>EMP_IND_SVY_2</b>	<b>Industry classification secondary employment</b>	<p>Use information from all sections (12 months or 7 days) that ask this question.</p> <p>First use information from the 12 month section. For all missing values and observations coded as 9, use the 7 day section.</p> <p>Keep the original classification of the survey. Most countries follow the ISIC classification.</p>
L37	<b>INDUSTRY_2</b>	<b>1 digit Industry classification secondary employment</b>  <i>Refer Annexure 2 for details</i>  1 = Agriculture and Fishing 2 = Mining 3 = Manufacturing 4 = Electricity and Utilities 5 = Construction 6 = Commerce 7 = Transportation, Storage and Communication 8 = Financial, Insurance and Real Estate 9 = Services, Public Administration 10 = Other Services 11 = Unspecified	<p>Classifies the <b>second job</b> of an individual <b>with a second job</b>, missing otherwise.</p> <p>If data allows generation of EMP_IND_SVY_2, recode INDUSTRY_2 according to SHIP requirements (refer Annexure-2)</p> <p>If data does not follow ISIC classification system, make appropriate assumptions to classify employment industry.</p> <p>Carefully document assumptions.</p> <p>Do not code household heads or EMPLHOME or FETCHWOOD or FETCHWOOD as done in INDUSTRY_1.</p>
L38	<b>EMPFRM_2</b>	<b>Farm employment in primary job</b> 0 = No 1= Yes	<p>Farming activities include cultivating, sharecropping, forestry, raising livestock, bee-keeping, fishing and hunting.</p> <p><b>If INDUSTRY_2=1 then code as 1 (YES). If INDUSTRY_2=2 to 10 then code as 0 (No)</b></p> <p>Do not code household heads or EMPLHOME or FETCHWOOD or FETCHWOOD as done in EMPFRM_1.</p>
L39	<b>EMPSEC_2</b>	<b>Employment sector of primary job</b>	<b>Code only for EMPLOYEE_WB_2=1, 4, 5, 7, 9 (wage employees), otherwise code as</b>

		1 = Public Government, Army 2 = State Owned 3 = Private, NGO	<p><b>missing.</b></p> <p>Classifies the secondary job's sector of activity for individuals with wage jobs.</p> <p>Use information from all sections (12 months or 7 days) that ask this question.</p> <p>First use information from the 12 month section. For all missing values and observations coded as 9, use the 7 day section.</p> <p><i>Armed forces and international institutions are included in Public sector.</i></p> <p><i>State owned</i> includes para-statal firms and all other enterprises in which the government has control.</p> <p><b>Note: if no such question leave as missing. This will mean that the in the variable EMPCAT_2 wage jobs will not be further disaggregated. Do not code it on the basis of ISCO or ISIC codes.</b></p>
L40	EMP_CAT_2	<p><b>Employment categories for primary job</b></p> <p>1= Wage worker  11=wage public  12=wage private non-agriculture  13=wage private agriculture  2 = Self employment/nonwage  21=family enterprise non-agri  22=family farmer</p>	<p><b>This is a derived variable from the EMPTYTYPE_WB_2, EMPSEC_2 and EMPFRM_2.</b></p> <p><b>If EMPTYTYPE_WB_2=1, 4, 5, 7, 9 then code as 1</b></p> <p>If EMP_CAT_2=1 and EMPSEC_2=1, 2 then code 11  If EMP_CAT_2=1 and EMPSEC_2= 3 and EMPFRM_2=0 then code 12  If EMP_CAT_2=1 and EMPSEC_2= 3 and EMPFRM_1=1 then code 13</p> <p><b>If EMPTYTYPE_WB_2=2, 3, 6 then code as 2</b>  If EMP_CAT_2=2 and EMPFRM_2=0 then code 21  If EMP_CAT_2=2 and EMPFRM_2=1 then code 22</p>
L41	WAGE_2	<p><b>Last wage payment in nominal local currency for secondary employment</b></p>	<p><b>For EMPTYTYPE_WB_1=1, 4, 5, 7, 9 only (wage employees).</b></p> <p>Use information from all sections (12 months or 7 days) that ask this question.</p>

			<p>First use information from the 12 month section. For all missing values and observations coded as 9, use the 7 day section.</p> <p>Wage from secondary job. This includes tips, bonuses, and other wage payments.</p> <p>By definition non-paid employees should have wage=0.</p>
L42	<b>UNITWAGE_2</b>	<p><b>Time unit for wage in secondary employment</b></p> <p>1 = Hourly 2 = Daily 3 = Weekly 4 = Every two weeks 5 = Twice per month 6 = Monthly 7 = Quarterly 8 = Biannual 9 = Annually 10 = Other</p>	<p><b>Code only for EMPTYTYPE_WB_1=1, 4, 5, 7, 9 (wage employees), otherwise code as missing.</b></p> <p>Unit of time for the wage variable of the secondary job.</p> <p>Use information from all sections (12 months or 7 days) that ask this question.</p> <p>First use information from the 12 month section. For all missing values and observations coded as 9, use the 7 day section.</p>
L43	<b>WHOURS_2</b>	<p><b>Hours worked in last week in secondary employment</b></p>	<p>This is the usual number of hours worked per week in the <b>secondary</b> job.</p> <p>Use information from all sections (12 months or 7 days) that ask this question.</p> <p>First use information from the 12 month section. For all missing values and observations coded as 9, use the 7 day section.</p> <p>If only information on how many hours individuals worked per day is available but no information on number of days worked a week is available, multiply the hours by 5 days.</p> <p>Similarly for hours worked per month divide by 4.2 to get weekly hours.</p>
L44	<b>SOU_AGRI_2</b>	<p><b>Individual income secondary employment self-employed Agricultural job</b></p> <p>0 = No 1 = Yes</p>	<p>Code as 1(YES) if the source of income is from self-employed agriculture activities, otherwise code as 0 (NO).</p> <p>If EMP_CAT_2! =. Or EMP_CAT_2! =. Then code as 0 Replace code = 1 If EMP_CAT_2=13 or 22</p>

L45	<b>SOU_NFARM_2</b>	<b>Individual income secondary employment self employed non-agricultural job</b> <i>0 = No</i> <i>1 = Yes</i>	Code as 1(YES) if the source of income is from self-employed agriculture activities, otherwise code as 0 (NO).  If EMP_CAT_2! =. Then code as 0 Replace code = 1 If EMP_CAT_2=21
L46	<b>SOU_WAGE_2</b>	<b>Individual income from secondary employment in a non-agricultural wage job</b> <i>0 = No</i> <i>1 = Yes</i>	Code as 1(YES) if the source of income is from self-employed agriculture activities, otherwise code as 0 (NO).  If EMP_CAT_2! =. Then code as 0 Replace code = 1 If EMP_CAT_2=11 or 12

## 2. Chapter 3: Module III – Standardized Household-Level File

This module records household-level information and includes household head characteristics and some selected information from individual-level module. Other information include housing characteristics and utilities, access to various amenities measured in terms of distances/time and ownership of durable goods among others. All variables are numeric unless otherwise stated.

### 3.1. Sample, Geography and Basic Household Identifier

No	NAME	LABEL AND CODES	COMMENTS, QUALITY CONTROL CHECKS AND CORRECTIONS
1	<b>COUNTRY</b>	<b>Country code</b>	String variable, length of 3. To be merged from Expenditure-level file.
2	<b>REGION</b>	Region code	Codes vary by country. To be merged from Expenditure-level file.
2b	<b>PROVINCE</b>	Province code	Codes vary by country. To be merged from Expenditure-level file.
2c	<b>DISTRICT</b>	District code	Codes vary by country. To be merged from Expenditure-level file.
3	<b>HID</b>	<b>Household unique identification</b>	String variable, length varies by country. To be merged from Expenditure-level file.
4	<b>SURVEYR</b>	<b>Year of survey</b>	To be merged from Individual-level file.
5	<b>RURURB</b>	<b>Area of residence</b> <i>1 = Rural</i> <i>2 = Urban</i>	To be merged from Expenditure-level file.
6	<b>HHSIZE</b>	<b>Household members</b>	Total # of residents (regular members) in the household, excluding maids and servants
6b	<b>HHSIZE_S</b>	<b>Household members including servants</b>	Total # of residents (regular members) in the household, including maids and servants
7	<b>WTA_HH</b>	<b>Household weighting coefficient</b>	To be merged from Expenditure-level file.
8	<b>WTA_POP</b>	<b>Population weighting coefficient</b>	To be merged from Expenditure-level file.

--	--	--	--

### 3.2. Household Head Characteristics

No	NAME	LABEL and CODES	COMMENTS, QUALITY CONTROL CHECKS AND CORRECTIONS
H1	HHSEX	<b>Sex of household Head</b> <i>1 = Male</i> <i>0 = Female</i>	To be merged from Individual-level file "SEX".
H2	HHAGEY	<b>Age of household Head</b>	To be merged from Individual-level file "AGEY".
H3	HHMARST	<b>Marital status of household Head</b> <i>1 = Never married</i> <i>2 = Married monogamous</i> <i>3 = Married polygamous</i> <i>4 = Common law, union coutumiere, union libre, living together</i> <i>5 = Divorced/Separated</i> <i>6 = Widowed</i>	To be merged from Individual-level file "MARSTAT".
H4	HHEDLEV	<b>Education level of household Head</b> <i>0 = No education</i> <i>1 = Pre-school</i> <i>2 = Primary, not completed</i> <i>3 = Completed primary, but less than completed lower secondary</i> <i>4 = Completed lower secondary (or post-primary vocational education) but less than completed upper secondary</i> <i>5 = Completed upper secondary (or extended vocational/technical education)</i> <i>6 = Post secondary technical</i> <i>7 = University and higher</i> <i>8 = Formal adult education or literacy program</i> <i>9 = Other</i>	To be merged from individual-level file "EDLEVEL_AR.
H5	HHEMSTAT	<b>Employment status of household Head (usual)</b>	To be merged from Labor variables.



		<i>0 = Unemployed</i> <i>1 = Employed</i>	Use variable EMPLOYED_WB and UNEMPLOYED_WB from Labor module.
H6	<b>HHEMTYPE</b>	<b>Employment type of household Head</b> <i>1 = Wage &amp; salaried worker</i> <i>2 = Self-employed with employees</i> <i>3 = Self-employed w/o employees</i> <i>4 = Employer</i> <i>5 = Domestic employees (work for compensation for private house)</i> <i>6 = Family worker</i> <i>7 = Apprentice</i> <i>9 = volunteer/Other</i>	To be merged from Labor variables.  For primary job of household head only.  Use variable EMPTYPE_WB_1 from Labor module
H7	<b>HHINDUSTRY</b>	<b>Sector of activity of household Head</b> <i>1 = Agriculture and fishing</i> <i>2 = Mining</i> <i>3 = Manufacturing</i> <i>4 = Electricity and utilities</i> <i>5 = Construction</i> <i>6 = Commerce</i> <i>7 = Transport, storage and communication</i> <i>8 = Financial, insurance and real estate</i> <i>9 = Services; Public administration</i> <i>10 = Other services</i> <i>11 = Unspecified</i>	To be merged from Labor variables.  For the primary job of household head only.  Use variable INDUSTRY_1 from Labor module

### 3.3. Housing and Utilities

No	NAME	LABEL and CODES	COMMENTS, QUALITY CONTROL CHECKS and CORRECTIONS
H8	<b>OWNHOUSE</b>	<b>Ownership of dwelling unit</b> <i>1 = Yes</i> <i>0 = No</i>	YES refers to: - (a) Actual ownership or (b) If household is in the process of purchasing the dwelling unit.
H9	<b>ROOMS</b>	<b>Number of habitable rooms</b>	This refers to number of rooms in the dwelling unit and may consist of one or more structures. Must be $\geq 1$ . Includes all rooms used for living, sleeping and eating. Excludes store, bathrooms and kitchens.
H 10	<b>ROOF</b>	<b>Main material used for roof</b>	

		1 = Concrete/cement/ brick/stone 2 = Wood 3 = Bamboo/thatch 4 = Tiles/shingles 5 = Tin/metal sheets 6 = Makeshift 9 = Other	
H 11	<b>WALLS</b>	<b>Main material used for external walls</b> 1 = Concrete/cement/ brick/stone 2 = Wood 3 = Bamboo/thatch 4 = Iron/metal sheets 5 = Clay/mud 6 = Makeshift 9 = Other	
H 12	<b>FLOOR</b>	<b>Main material used for floor</b> 1 = Cement/tiles/marble 2 = Wood/Bamboo 3 = Vinyl 4 = Earth/clay/mud 9 = Other	
H 13	<b>WATER</b>	<b>Main source of water</b> 1 = Pipe (own tap) 2 = Public standpipe 3 = Borehole 4 = Wells (protected) 5 = Wells (unprotected) 6 = Surface water 7 = Rain water 8 = Vendor/truck 9 = Other	Drinking water if specified.
H 14	<b>ADISWAT</b>	<b>Actual distance to main water point (kms)</b>	This refers to actual distance to water point ( <b>one way</b> ) used by household in kms. 1km = 1000 m; 1 km = 5/8 mile. If within dwelling, zero
H 15	<b>ATIMWAT</b>	<b>Actual time taken to main water point (mins)</b>	This refers to actual time taken to water point used by household. If roundtrip provided, divide by 2.
H 16	<b>ELECTCON</b>	<b>Connection of electricity in dwelling</b> 1 = Central/local 2 = Solar/wind 3 = Generator 9 = Other 0 = None	<i>Electricity</i> refers to mains, generator and solar energy. <i>Public or quasi</i> refers to electricity from a public or semi-public firm. <i>Private</i> refers to electricity from private company, generator or solar. If electricity source connection cannot be determined code 9 ( <i>Yes, but source unstated</i> ).

H 17	<b>FUELCOOK</b>	<b>Main cooking fuel</b> 1 = Firewood 2 = Kerosene 3 = Charcoal 4 = Electricity 5 = Gas 9 = Other	<i>Other</i> includes fuel derived from coffee waste, saw dust, crop residue, cow dung among others.
H 18	<b>FUELLIGH</b>	<b>Main lighting fuel</b> 1 = Electricity 2 = Kerosene 3 = Candles 4 = Gas 9 = Other	
H 19	<b>TOILET</b>	<b>Main toilet facility</b> 1 = Flush toilet 2 = Improved pit latrine 3 = Pit latrine 4 = No facility 9 = Other	
H 20	<b>GARBDISP</b>	<b>Garbage and trash disposal</b> 1 = Collected 2 = Buried/burned 3 = Discarded in empty lots, street, rivers 9 = Other	Refers to only garbage or trash generated by household.
H 21	<b>DISPSCH</b>	<b>Distance to nearest elementary/primary school (kms)</b>	This refers to one-way distance to nearest primary school in kms.
H 22	<b>DISPSCH1</b>	<b>Distance to nearest elementary/primary school</b> 1 = Outside dwelling and up to 1 km 2 = >1-2 km 3 = >2-5 km 4 = >5-10 km 5 = Over 10 km 9 = Undefined	One-way distance to school.
H 23	<b>TIMPSCH</b>	<b>Time taken to nearest elementary/primary school (minutes)</b>	One way
H 24	<b>TIMPSCH1</b>	<b>Time taken to nearest elementary/primary school</b> 1 = Within ½ hour 2 = >½ to 1 hour away 3 = >1 to 6 hours 4 = >6 to 12 hrs 5 = Over 12 hrs 9 = Undefined	One way

H 25	DISHEAL	Distance to nearest health facility (kms)	One way
H 26	DISHEAL1	Distance to nearest health facility 1 = Outside dwelling and up to 1 km 2 = >1-2 km 3 = >2-5 km 4 = >5-10 km 5 = Over 10 km 9 = Undefined	One way
H 27	TIMHEAL	Time taken to nearest health facility (minutes)	One way
H 28	TIMHEAL1	Time taken to nearest health facility 1 = Within ½ hour 2 = >½ to 1 hour away 3 = >1 to 6 hours 4 = >6 to 12 hrs 5 = Over 12 hrs 9 = Undefined	One way

### 3.4. Ownership of Durable and Non-Durable Goods

No	NAME	LABEL AND CODES	COMMENTS, QUALITY CONTROL CHECKS AND CORRECTIONS
H 29	AGLAND	Ownership of agricultural land 1 = Yes 2 = Cannot be determined 3 = No	Country specific on definition of land ownership. This only refers to <b>agricultural land</b> . Some countries inquire about land ownership irrespective whether it is agricultural or non-agricultural purposes. If the survey is not explicit on type of land code 2 and skip to LANDSIZE.
H 30	AGLNDOW	Agricultural land owned by household (ha)	Refers only to <b>agricultural</b> land size in hectares.  Check, if AGLAND==3 , code as 0  By convention 1 ha = 2.471 acres.  <b>Note: This would exclude rental farmers but include farm land owners who are not farmers.</b>
H 31	OTHLNDOW	Non Agricultural land owned by household (ha)	Refers only to <b>non-agricultural</b> land size in hectares. By convention 1 ha = 2.471 acres.
H 32	LANDSIZE	Land size owned by household (ha)	Area of all land owned by household regardless of use, that is, both agricultural and non-agricultural. By convention 1 ha = 2.471 acres.

H 33	<b>AREA_IRRIG</b>	<b>Area of irrigated agricultural land (ha)</b>	
H 34	<b>AREA_NONIR RIG</b>	<b>Area of non-irrigated agricultural land (ha)</b>	
H 35	<b>AREA_COLL AT</b>	<b>Total plot areas can be used as collateral for loan (ha)</b>	
H 36	<b>LANDCON</b>	<b>Conflict about the land</b> 1 = Yes 0 = No	
H 37	<b>RADIO</b>	<b>Ownership of radio</b> 1 = Yes 0 = No	Functioning radio includes a radio, radio cassette, and 3-in-1-radio cassette.
H 38	<b>TV</b>	<b>Ownership of television</b> 1 = Yes 0 = No	Presence of a functioning television in house.
H 39	<b>PHONE</b>	<b>Ownership of landline (fixed) phone</b> 1 = Yes 0 = No	Presence of a functioning fixed land line telephone in house.
H 40	<b>CPHONE</b>	<b>Ownership of at least one cellular phone</b> 1 = Yes 0 = No	Presence of a functioning cellular in house.
H 41	<b>INTERNET</b>	<b>Access to internet</b> 1 = Yes 0 = No	Whether inside or outside the house.
H 42	<b>FRIDGE</b>	<b>Ownership of frig</b> 1 = Yes 0 = No	Presence of a functioning refrigerator in house.
H 43	<b>SEWMACH</b>	<b>Ownership of sewing machine</b> 1 = Yes 0 = No	Presence of a functioning sewing machine in house.
H 44	<b>COMPUTER</b>	<b>Ownership of computer</b> 1 = Yes 0 = No	Presence of a functioning computer in house.
H 45	<b>STOVE</b>	<b>Ownership of stove</b> 1 = Yes 0 = No	Presence of a functioning stove or cooker in house.
H 46	<b>OX CART</b>	<b>Ownership of animal cart</b> 1 = Yes 0 = No	Presence of a functioning animal cart, which is used as a means of transport or a farm tool.
H 47	<b>BCYCLE</b>	<b>Ownership of bicycle</b> 1 = Yes 0 = No	Presence of a functioning bicycle.
H 48	<b>BOAT</b>	<b>Ownership of boat</b> 1 = Yes 0 = No	Presence of a functioning boat or canoe.
H 49	<b>MCYCLE</b>	<b>Ownership of</b>	Presence of a functioning motorcycle is

		<b>motorcycle</b> 1 = Yes 0 = No	important.
H 50	<b>CAR</b>	<b>Ownership of private car</b> 1 = Yes 0 = No	Presence of a functioning car is important. This refers to car for household use and NOT a commercial vehicle.
H 51	<b>LLIVESK</b>	<b>Number of large-sized livestock owned</b>	This question asks the actual number of large-sized livestock. These include cattle, camels, donkeys and horse.
H 52	<b>MLIVESK</b>	<b>Number of medium-sized livestock owned</b>	This question asks the actual number of medium-sized livestock. These include sheep, goats and pigs.
H 53	<b>POULTRY</b>	<b>Number of poultry owned</b>	This question asks the actual number of poultry/birds. These include all forms of birds such as chicken, geese, and doves.

### 3.5. Household Expenditure

H 54	<b>PCEXP</b>	<b>Per capita annual food and non-food consumption expenditure in national currency, not regionally deflated</b>	To be merged from expenditure-level file.
H 55	<b>PCEXPDR</b>	<b>Per capita annual food and non-food consumption expenditure in national currency, adjusted by regional food price indices, and <i>cleaned</i>.</b>	Per capital annual food and non-food consumption is first adjusted by regional food price indices. Then outliers (separately for food and non-food expenditure) greater than three standard deviations are replaced with the medians of their correspondent per capita food and non-food consumption by region and rural/urban.
H 56	<b>PCEXPDR_PPP</b>	<b>Per capita annual food and non-food consumption expenditure in real terms and 2005 \$PPP</b>	Calculated using PCEXPDR, CPI of the survey year, and 2005 \$PPP conversion factor. When the regional food price deflators are not available, remove the outliers in PCEXP following the same procedures as removing the outliers in PCEXPDR, and calculate PCEXP_PPP using PCEXP, CPI of the survey year, and 2005 \$PPP conversion factor.
H 57	<b>QUINTILE_N</b>	<b>Income quintiles at the national level</b>	Calculated based on PCEXP_PPP, and weighted by population weights.
H 58	<b>QUINTILE_RU</b>	<b>Income quintiles by rural/urban</b>	Calculated based on PCEXP_PPP, and weighted by population weights.

### 3.6. Household Remittances

H 59a	<b>SEX_RMT_1</b>	<b>Sex of the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> remittance sending member</b> <i>1 = Male</i> <i>0 = Female</i>	The order of the sending members is in decreasing order of amount of remittance (remittance includes cash, gifts and food).
H 59b	<b>SEX_RMT_2</b>		
H 59c	<b>SEX_RMT_3</b>		
H 60a	<b>RELAT_RMT_1</b>	<b>Relationship to the household head of the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> remittance sending member</b> <i>2 = Spouse</i> <i>3 = Son/daughter</i> <i>4 = Parents/grandparents</i> <i>5 = Grandchild</i> <i>6 = Son/daughter-in-law</i> <i>7 = Siblings and other relative</i>	The order of the sending members is in decreasing order of amount of remittance (remittance includes cash, gifts and food).
H 60b	<b>RELAT_RMT_2</b>		
H 60c	<b>RELAT_RMT_3</b>		
H 61a	<b>DES_MIG_1</b>	<b>Destination of migration of the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> remittance sending member</b> <i>1 = Capital</i> <i>2 = Within the country (but not capital)</i> <i>3 = Abroad</i>	The order of the sending members is in decreasing order of amount of remittance (remittance includes cash, gifts and food).
H 61b	<b>DES_MIG_2</b>		
H 61c	<b>DES_MIG_3</b>		
H 62	<b>ORIGIN_RMT</b>	<b>Origin of the remittances</b> <i>1 = Domestic</i> <i>2 = Abroad</i> <i>3 = Both</i>	Location category from where the household has received remittances.
H 63a	<b>AMT_RMT_1</b>	<b>Amount of annual remittance by the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> remittance sending member</b>	The order of the sending members is in decreasing order of amount of remittance (remittance includes cash, gifts and food).
H 63b	<b>AMT_RMT_2</b>		
H 63c	<b>AMT_RMT_3</b>		
H 64	<b>INC_RMT</b>	<b>Total amount of annual remittances received from remittance sending members (annual)</b>	The total includes the remittances from the most important three migrant members and others.

### 3.7. Household Income Variables

H 65	<b>INC_WAGE</b>	<b>Wage and salaries (annual)</b>	Aggregated from individual data.
H 66	<b>INC_SOCA</b>	<b>Bonus and social allowances derived from wage jobs (annual)</b>	Aggregated from individual data.
H 67	<b>INC_FARM_G</b>	<b>Gross income from household farm businesses (annual)</b>	Gross receipt/revenue. Income from leasing agricultural land is counted in INC_RENT.
H 68	<b>INC_NFARM_G</b>	<b>Gross income from household nonfarm</b>	Gross receipt/revenue. Gross receipt/revenue minus operating cost

		<b>businesses (annual)</b>	(as reported).
H 69	<b>INC_RENT</b>	<b>Rental income (annual)</b>	Income from selling or leasing land, houses, and other properties. Most survey questionnaires have asked about both rent from leasing land and houses in the housing module, and rent from leasing land in the agricultural module. Count both, assuming there's no overlap.
H 70	<b>INC_FIN</b>	<b>Financial income (annual)</b>	Income from savings, dividends, loans, tax refunds, and maturity payment on insurance
H 71	<b>INC_PEN</b>	<b>Pensions/ social assistance (annual)</b>	Income from pensions and other social assistance. A similar variable is in the labor module. However, that is at an individual level whereas this is at a household level.
H 72	<b>INC_STUDY</b>	<b>Study support (annual)</b>	Income from scholarships, stipends, or other study support
H 73	<b>INC_OTH</b>	<b>Other income (annual)</b>	
H 74	<b>INC_TOT_G</b>	<b>Total gross household income, (annual)</b>	In national currency. Total of labor income and non-labor incomes including financial income, remittances, pensions and other social assistance, study support and other.  <b>Sum of all variables starting with INC_. (Include INC_RMT from section 3.6)</b>

**Note:**

- a) Any variable, for which information is not collected by the country, should be created and left as system missing in the final harmonized aggregated expenditure file.
- b) The order of variables should follow as shown above for ease in reference use.

**4. Chapter 4: Module IV – Harmonized Expenditure File**

One of the most important information from the household survey data is household expenditure aggregates, upon which poverty and inequality measures are based.

The following guidelines should be kept in mind:-

- a) Before starting the expenditure aggregation, household size must be first calculated. This is because using expenditure per capita one can determine outliers.
- b) If a particular sub-category does not exist in the survey data, list all values as missing.
- c) All expenditures should be annualized.
- d) Although the summary tables are produced with great care to make the sub-aggregates as well defined and exhaustive as possible, the users will still need to make circumstantial (important) decisions. Thus all computer programs should be attached so the aggregation can be appropriately documented.

The following 15 summary tables must be generated, which serve both as guidelines and record the process of expenditure aggregation.



<b>Contents</b>	<b>Sequence</b>	<b>Table</b>	<b>Data File</b>
Household composition	File 0	Table 0	00_BASICINFO.xxx
Food and non-alcoholic beverages	File 1	Table 1	01_EXPFOOD.xxx
Alcoholic beverages, tobacco and narcotics	File 2	Table 2	02_EXPALCH.xxx
Clothing and footwear	File 3	Table 3	03_EXPCLTH.xxx
Housing, water, electricity, gas and other fuels	File 4	Table 4	04_EXPHOUS.xxx
Furnishings, household equipment, etc.	File 5	Table 5	05_EXPFURN.xxx
Health	File 6	Table 6	06_EXPHLTH.xxx
Transport	File 7	Table 7	07_EXPTRSP.xxx
Communication	File 8	Table 8	08_EXPRCRE.xxx
Recreation and culture	File 9	Table 9	09_EXPCMNU.xxx
Education	File 10	Table 10	10_EXPEDUC.xxx
Restaurants and hotels	File 11	Table 11	11_EXPHOTL.xxx
Miscellaneous goods and services	File 12	Table 12	12_EXPMISC.xxx
Regional food price deflators	File 13	Table 13	13_FDPINDEX.xxx
Total household expenditures	File 14	Table 14	14_HHEXP.xxx

**4.0. Table 0: Household Composition and Weights Summary Table**

	<b>Variable name</b>	<b>Variable definition</b>	<b>Aggregated from variables in X files</b>	<b>Section</b>	<b>Syntax and data files</b>
1	COUNTRY	Country name (3-letters based on ISO)			00_BASIC INFO.xxx
2	REGION	Geographical code This code will be the country code for regions			00_BASIC INFO.xxx
3	HID	Household unique ID			00_BASIC INFO.xxx
4	SURVEYR	Survey year			00_BASIC INFO.xxx
5	RURURB	Area of residence, rural = 1, Urban = 2			00_BASIC INFO.xxx
6	HHSIZE	Total # of residents (regular members) in the household, excluding maids and servants			00_BASIC INFO.xxx
6b	HHSIZE_S	Total # of residents (regular members) in the household, including maids and servants			00_BASIC INFO.xxx
7	WTA_HH	Household weight. This is the weighting co-efficient to derive household-level estimates of indicators.			00_BASIC INFO.xxx
8	WTA_POP	Population weight This is the weighting co-efficient to derive population-level estimates of indicators.	<b>WTA_HH*HHSIZE</b>		00_BASIC INFO.xxx

**4.1. Table 1: Food Expenditures**

All variables are annualized and in current local nominal prices

	<b>Variable name</b>	<b>Variable definition</b>	<b>Aggregated from variables in X files</b>	<b>Section</b>	<b>Syntax and data files</b>
3	HID	Household unique ID			01_EXPFOOD.xxx
E1	FDRECALL	Food purchase recall period, in weeks. By convention 1 week = 7 days.			01_EXPFOOD.xxx
E2	FD_B	Total expenditure on purchased food. Beverages, tobacco, catering and restaurant foods are excluded			01_EXPFOOD.xxx
E3	FDNONALC_B	Total expenditure on purchased non-alcoholic			01_EXPFOOD.xxx

		beverages			
E4	FD_P	Total value of food other than purchased, such as self-produced, in-kind receipts, etc., beverages and tobacco excluded			01_EXPFOOD.xxx
E5	FDNONALC_P	Total value of non-alcoholic beverages other than purchased, such as self-produced, in-kind receipts, etc.			01_EXPFOOD.xxx
E6	TOTFOOD	Total food expenditure, sum of variables with prefix FD			01_EXPFOOD.xxx

#### 4.2. Table 2: Alcoholic Beverages, Tobacco and Narcotics

All variables are annualized and in current local nominal prices

3	HID	Household unique ID			02_EXPALCH.xxx
E7	FDALC_B	Total expenditure on purchased alcoholic beverages			02_EXPALCH.xxx
E8	FDALC_P	Total value of alcoholic beverages other than purchased, such as self-produced, gifts received, etc.			02_EXPALCH.xxx
E9	TOBACCO	Sum of total expenditure on purchased tobacco and tobacco products, and total value of tobacco received in-kind or self-produced			02_EXPALCH.xxx
E10	NARCOTICS	Sum of total expenditure on purchased narcotics and narcotics products, and total value of narcotics received in-kind or self-produced			02_EXPALCH.xxx
E11	TOTALCH	Total expenditures on alcoholic beverages, tobacco and narcotics, sum of all above variables			02_EXPALCH.xxx

#### 4.3. Table 3: Clothing and Footwear

All variables are annualized and in current local nominal prices

	<b>Variable name</b>	<b>Variable definition</b>	<b>Aggregated from variables in X files</b>	<b>Section</b>	<b>Syntax and data files</b>
3	HID	Household unique ID			03_EXPCLTH.xxx
E12	HSCLOTH	Total expenditure on clothing, including self produced and in-kind receipts			03_EXPCLTH.xxx
E13	HSFOOTW	Total expenditure on footwear, including self-produced and in-kind			03_EXPCLTH.xxx

		receipts			
E14	TOTCLTH	Total expenditure on clothing and footwear, including self-produced and in-kind receipts, sum of the above two variables			03_EXPCLTH.xxx

**4.4. Table 4: Housing, Water, Electricity, Gas and Other Fuels**  
All variables are annualized and in current local nominal prices

3	HID	Household unique ID			04_EXPHOUS.xx x
E15	RENT_ACT	Actual rentals for housing			04_EXPHOUS.xx x
E16	RENT_IMP	Imputed housing for rental			04_EXPHOUS.xx x
E17	HSREPAIR	Maintenance and repair of dwelling			04_EXPHOUS.xx x
E18	HSWATER	Water supply and miscellaneous services relating to the dwelling such as garbage collection fees			04_EXPHOUS.xx x
E19	HSFUEL	Firewood, and charcoal			04_EXPHOUS.xx x
E20	GAS	Gas for cooking			04_EXPHOUS.xx x
E21	ELEC	Electricity			04_EXPHOUS.xx x
E22	HSKEROSE NE	Kerosene (Paraffin)			04_EXPHOUS.xx x
E23	HSDIESEL	Diesel for all purposes except transportation. Note: If in the questionnaire does not specify the purpose of diesel use, include it in TRFUEL in Transport Expenditures in Table 4.7.			04_EXPHOUS.xx x
E24	HSUTILITY	Electricity and gas (row total of GAS and ELEC)			04_EXPHOUS.xx x
E25	TOTHOUS	Total expenditures on housing, sum of variables with the prefix HS			04_EXPHOUS.xx x

**4.5. Table 5: Furnishings, Household Equipment and Routine Household Maintenance**  
All variables are annualized and in current local nominal prices

3	HID	Household unique ID			05_EXPFURN.xxx
E26	HSFURNI	Furniture and furnishings, carpets and other floor coverings.			05_EXPFURN.xxx
E27	HSTEXTIL	Household textiles such as beddings, and drapes.			05_EXPFURN.xxx
E28	HSAPPLIAN	Household appliances			05_EXPFURN.xxx
E29	HSUTENSIL	Glassware, tableware,			05_EXPFURN.xxx

		kitchenware, and household utensils.			
E30	HSTOOLS	Tools and equipment for house and garden			05_EXPFURN.xxx
E31	HSSERV	Goods and services for routine household maintenance, e.g., cleaning products, domestic services.			05_EXPFURN.xxx
E32	TOTFURN	Total furnishing etc. expenditure, sum of all variables with the prefix HS			05_EXPFURN.xxx

#### 4.6. Table 6: Health Care Expenditures

All variables are annualized and in current local nominal prices

	<b>Variable name</b>	<b>Variable definition</b>	<b>Aggregated from variables in X files</b>	<b>Section</b>	<b>Syntax and data files</b>
3	HID	Household unique ID			06_EXPHLTH.xxx
E33	HLMEDI	Medical products, appliances and equipment			06_EXPHLTH.xxx
E34	HLOUTP	Outpatient services			06_EXPHLTH.xxx
E35	HOSP	Hospitalization expenditure			06_EXPHLTH.xxx
E36	TOTHLTH	Total health expenditures, sum of all variables with the prefix HL (Do <b>NOT</b> include <b>HOSP</b> )			06_EXPHLTH.xxx

#### 4.7. Table 7: Transport Expenditures

All variables are annualized and in current local nominal prices

	<b>Variable name</b>	<b>Variable definition</b>	<b>Aggregated from variables in X files</b>	<b>Section</b>	<b>Syntax and data files</b>
3	HID	Household unique ID			07_EXPTRSP.xxx
E37	VEHICLES	Purchase of vehicles			07_EXPTRSP.xxx
E38	TRFUEL	Fuels for transportation, such as petrol, diesel, etc... Note: If in the questionnaire does not specify the purpose of diesel/petrol use, include it in TRFUEL.			07_EXPTRSP.xxx
E39	TRPERSON	Operation of personal transport equipment, such as repairs and hire of chauffeurs, but excluding fuel.			07_EXPTRSP.xxx
E40	TRSERVE	Transport services, such as public transportation, purchase of airplane tickets, etc.			07_EXPTRSP.xxx

E41	TOTTRSP	Total transport expenditures, sum of all variables with the prefix TR (Do <b>NOT</b> include <b>Vehicles</b> )			07_EXPTRSP.xxx
-----	---------	--	--	--	----------------

#### 4.8. Table 8: Communication Expenditures

All variables are annualized and in current local nominal prices

	<b>Variable name</b>	<b>Variable definition</b>	<b>Aggregated from variables in X files</b>	<b>Section</b>	<b>Syntax and data files</b>
3	HID	Household unique ID			08_EXPCMNQ.xxx
E42	COMPOST	Postal services			08_EXPCMNQ.xxx
E43	COMTEL	Telephone and fax equipment			08_EXPCMNQ.xxx
E44	COMSERVE	Telephone, fax, and internet services			08_EXPCMNQ.xxx
E45	TOTCMNQ	Total communication expenditures, sum of all variables with the prefix COM			08_EXPCMNQ.xxx

#### 4.9. Table 9: Recreation and Culture Expenditures

All variables are annualized and in current local nominal prices

	<b>Variable name</b>	<b>Variable definition</b>	<b>Aggregated from variables in X files</b>	<b>Section</b>	<b>Syntax and data files</b>
4	HID	Household unique ID			09_EXPRCRE.xxx
E46	RCAUDI	Audio-visual, photographic and information processing equipment			09_EXPRCRE.xxx
E47	DURRC	Other major durables for recreation and culture			09_EXPRCRE.xxx
E48	RCPETS	Other recreational items and equipment, gardens and pets			09_EXPRCRE.xxx
E49	RCSERVE	Recreational and cultural services, such as movies, etc.			09_EXPRCRE.xxx
E50	RCNEWS	Newspapers, books and stationary of non-educational purpose			09_EXPRCRE.xxx
E51	RCHOLID	Package holidays			09_EXPRCRE.xxx
E52	TOTRCRE	Total recreation expenditures, sum of all variables with the prefix RC (Do <b>NOT</b> include <b>DURRC</b> )			09_EXPRCRE.xxx

#### 4.10. Table 10: Education Expenditures

All variables are annualized and in current local nominal prices

	<b>Variable name</b>	<b>Variable definition</b>	<b>Aggregated from variables in X files</b>	<b>Section</b>	<b>Syntax and data files</b>
3	HID	Household unique ID			10_EXPEDUC.xxx
E53	EDPRIMAR	Pre-primary and primary			10_EXPEDUC.xxx
E54	EDSECOND	Secondary education			10_EXPEDUC.xxx
E55	EDPOSTSEC	Post-secondary but non-tertiary education			10_EXPEDUC.xxx
E56	EDTERTIAR	Tertiary education			10_EXPEDUC.xxx
E57	EDUNDEFIN	Education cannot be defined			10_EXPEDUC.xxx
E58	TOTEDUC	Total expenditure on education, sum of all variables with ED prefix			10_EXPEDUC.xxx

#### 4.11. Table 11: Restaurants and Hotels

All variables are annualized and in current local nominal prices

	<i>Variable name</i>	<i>Variable definition</i>	<i>Aggregated from variables in X files</i>	<i>Section</i>	<i>Syntax and data files</i>
3	HID	Household unique ID			11_EXPHOTL.xxx
E59	HOTCAT	Catering services, including eating out			11_EXPHOTL.xxx
E60	HOTACC	Accommodation services			11_EXPHOTL.xxx
E61	TOTHOTL	Total Restaurants and hotel expenditures, sum of the variables with HOT			11_EXPHOTL.xxx

#### 4.12. Table 12: Miscellaneous Goods and Services

All variables are annualized and in current local nominal prices

	<i>Variable name</i>	<i>Variable definition</i>	<i>Aggregated from variables in X files</i>	<i>Section</i>	<i>Syntax and data files</i>
3	HID	Household unique ID			12_EXPMISC.xxx
E62	MISPERSN	Personal care, e.g., hairdressing, electronic appliances for personal care.			12_EXPMISC.xxx
E63	MISPROST	Spending on prostitution			12_EXPMISC.xxx
E64	MISPEREFF	Personal effects not mentioned elsewhere			12_EXPMISC.xxx
E65	MISSPROT	Contribution to social protection			12_EXPMISC.xxx
E66	MISINSUR	Premium paid for insurances, including housing, health, etc.			12_EXPMISC.xxx
E67	MISFINAN	Financial services			12_EXPMISC.xxx
E68	MISOTHER	Other services not mentioned elsewhere. <b>Include Ceremonies such as weddings/funerals etc. if they are not found to be lumpy expenditures i.e more than 5% of the total households incurred it.</b>			12_EXPMISC.xxx
E69	TOTMISC	Total miscellaneous spending, sum of all variables with the prefix MIS			12_EXPMISC.xxx



#### 4.13. Table 13: Regional Price Deflators

These price variables will be used to take away regional price variation and the inflation effect (see PowerPoint presentation for details). Food and non-food indices may be computed separately.

##### 4.13.1. Sub-File

	<b>Variable name</b>	<b>Variable definition</b>	<b>Aggregated from variables in X files</b>	<b>Section</b>	<b>Syntax and data files</b>
A	REGION	ID for regions, the lowest level to calculate price index			13a_Pindex1.xxx
B	FDID	ID for items, select 20 items that have the largest food share			13a_Pindex1.xxx
C	FDRPRICE	Regional food prices for the 20 selected items			13a_Pindex1.xxx
D	FDSHR	Normalize the 20 item food shares to 100 percent			13a_Pindex1.xxx
E	FDINDEX	Regional food price deflator.			13a_Pindex1.xxx

**Note:**

- The food price index and CPI will then be AGGREGATED and MERGED into the HH-level file based on the lowest level to calculate price index (region).
- Variables FDINDEX, and NPCPI are the ones that will be merged into the overall HH-level file. Therefore households in the same region will have identical values for FDINDEX.
- FDINDEX is calculated as  $FDINDEX = \frac{\sum_{i=1}^{20} p^i \times q^i}{\sum_{i=1}^{20} \bar{p}^i \times \bar{q}^i}$  where  $p^i$  is the regional price,  $\bar{p}^i$  is national average price and  $q^i$  is the national average food share.

##### 4.13.2. Final File

	<b>Variable name</b>	<b>Variable definition</b>	<b>Aggregated from variables in X files</b>	<b>Section</b>	<b>Syntax and data files</b>
2	REGION	ID for regions, the lowest level to calculate price index			13_FDPINDEX.xxx
E70	FDPINDEX	Regional food price deflators.			13_FDPINDEX.xxx

#### 4.14. Table 14: Total Annual Household Expenditures

Merge Tables 1 – 13. Code is the country code (Ref Annex 1)

	<b>Variable name</b>	<b>Variable definition</b>	Comments if any	<b>Syntax and data files</b>
3	HID	Household unique ID		Ccode_YYYY_E.xxx
2	REGION	Country specific region codes		Ccode_YYYY_E.xxx
5	RURURB	Rural = 1 Urban =2		Ccode_YYYY_E.xxx
6	HHSIZE	Total # of residents (regular members) in the household, excluding maids and servants		Ccode_YYYY_E.xxx
6b	HHSIZE_S	Total # of residents (regular members) in the household, including maids and servants		Ccode_YYYY_E.xxx
E6	TOTFOOD	Food and non-alcoholic beverages		Ccode_YYYY_E.xxx
E11	TOTALCH	Alcoholic beverages, tobacco and narcotics		Ccode_YYYY_E.xxx
E66	TOTFDAL	Sum of TOTFOOD(E6) and TOTALCH(E11)		Ccode_YYYY_E.xxx
E14	TOTCLTH	Clothing and footwear		Ccode_YYYY_E.xxx
E24	HSUTILITY	Electricity and gas		Ccode_YYYY_E.xxx
E19	HSFUEL	Firewood, and charcoal		Ccode_YYYY_E.xxx
E20	GAS	Gas for cooking		Ccode_YYYY_E.xxx
E21	ELEC	Electricity		Ccode_YYYY_E.xxx
E22	HSKEROSENE	Kerosene (Paraffin)		Ccode_YYYY_E.xxx
E23	HSDIESEL	Diesel for all purposes except transportation		Ccode_YYYY_E.xxx
E25	TOTHOUS	Housing, water, electricity, gas and other fuels		Ccode_YYYY_E.xxx
E32	TOTFURN	Furnishings, household equipment, and routine household maintenance		Ccode_YYYY_E.xxx
E36	TOTHLTH	Health		Ccode_YYYY_E.xxx
E38	TRFUEL	Fuel for transportation, such as petrol, diesel, etc.....		Ccode_YYYY_E.xxx
E 39	TRPERSON	Operation of personal transport equipment, such as gas, repairs and hire of chauffeurs		Ccode_YYYY_E.xxx
E 40	TRSERVE	Transport services, such as public transportation, purchase of airplane tickets, etc.		Ccode_YYYY_E.xxx
E41	TOTTRSP	Transportation		Ccode_YYYY_E.xxx
E45	TOTCMNQ	Communication		Ccode_YYYY_E.xxx
E52	TOTRCRE	Recreations and culture		Ccode_YYYY_E.xxx
E58	TOTEDUC	Education		Ccode_YYYY_E.xxx
E61	TOTHOTL	Restaurants and hotels		Ccode_YYYY_E.xxx
E69	TOTMISC	Miscellaneous goods and services		Ccode_YYYY_E.xxx
E71	TOTNFD	Sum of all non-food expenditure		Ccode_YYYY_E.xxx

		variables (exclude TOTFOOD and TOTALCH)		
E70	FDPINDEX	Regional food price deflators		Ccode_YYYY_E.xxx
E72	HHEXP_N	Total annual nominal household expenditures		Ccode_YYYY_E.xxx
E73	HHEXP_R	Total annual real household expenditures deflated by FDPINDEX		Ccode_YYYY_E.xxx
E74	CPI2005_DEF	CPI to deflate/inflate the HHEXP_R to 2005 local prices		Ccode_YYYY_E.xxx
E75	PPP2005	2005 PPP exchange rate to 2005 local currency		Ccode_YYYY_E.xxx

**Note: All missing values for expenditure aggregates must be recoded to 0 in the Total Annual Expenditure file (E file)**

## **5. Chapter 5: Documentation and Computer Programs**

For each standardized data file, a short document must be written, containing at least the following information: -

- a) Date of production of the harmonized data files.
- b) Person responsible for production.
- c) Extrapolated population, by Urban/Rural area.
- d) Definition of the variable HID (this variable is usually the concatenation of several variables in the original dataset; these variables must be clearly identified, in order to facilitate the link with the original dataset).
- e) All data editing checks (cleaning and assumptions) should be documented in the computer programs as comments.
- f) Imputations and assumptions made (number of cases and method) should be recorded meticulously.
- g) Computer programs should be attached in sequence.
- h) Comments and warnings to the users (limits in survey coverage, reliability of data, etc).
- i) Any other information pertaining to the data that might be useful to data users.

## 6. Appendices

### 6.1 Appendix 1: Naming Convention

Naming of the data files will start with the country code (3 letters), followed by the start year of survey (2 digits) and one letter (“H” for household files; “I” for individual level files and “E” for expenditure files). The character “\_” will separate these three components. The extension of the files for SPSS-format files will be (\*.SAV), SAS-format files (\*.SSD) and STATA-format files (\*.STA). For example, the standardized data sets produced for the 2009 survey for Indonesia will be called IND\_09\_H.XX for the household-level file; IND\_09\_I.XX for the individual-level file and IND\_09\_E.XX for expenditure file.

It is strongly recommended that the variable and file names proposed in the manual, be followed closely.

The file names for the computer programs and their resulting data files should be same, only differing in file types (extension of the file name).

### 6.2 Appendix 2: Directory Organization

Six sub-directories should be created under the main directory named after the survey, as shown below.

Main directory:	Storage Purposes
“\COUNTRYCODE_SURVEYNAME_YEAR\”	Parent directory of all sub-directories
<b>Sub-Directory</b>	
“\RAWDATA”	All survey data sets as keyed in (no editing has been done on this data). This can further be divided into various format, STATA, SPSS, SAS directories
“\CLEANDATA”	All clean survey data sets before aggregation This can further be divided into various format, STATA, SPSS, SAS directories
“\PROGRAMS”	All program files that derive the above indicators (Sec 2.7)
“\AGGDATA”	All aggregated data files COUNTRY_YEAROFSURVEY_I ( <i>one file</i> Sec 2.7.1) COUNTRY_YEAROFSURVEY_H ( <i>one file</i> Sec 2.7.2) COUNTRY_YEAROFSURVEY_E ( <i>8 sub-files - 2.7.3.1-2.7.3.8 - which are merged to form one expenditure file: 2.7.3.9). The expenditure file will follow Annex X templates.</i>
“\DOC”	All survey and aggregation documentation, and survey reports PDF documents are better than WORD or EXCEL documents.

### 6.3 Appendix 3: Standard Excel Spreadsheet

Country Survey Name Year: Selected Social & MDG Indicators by Rural Urban and Consumption Level																				
Page 1 of 2		Expenditure Quintile																		
		National						Rural						Urban						
	Unit of measure	Poorest 20%	II	III	IV	Richest 20%	All	Poorest 20%	II	III	IV	Richest 20%	All	Poorest 20%	II	III	IV	Richest 20%	All	
<b>Indicators</b>																				
<b>Demographic indicators</b>																				
Sample size (households)	Number																			
Total population (weighted)	Million																			
Age dependency ratio (age <15 or >64)	Percent																			
Average household size	Number																			
<b>Head of household characteristics</b>																				
<b>Marital status of head</b>																				
Married/common law male	Percent																			
Single/ divorced/widowed male	Percent																			
Married/common law female	Percent																			
Single/ divorced/widowed female	Percent																			
<b>Education level of head</b>																				
No level	Percent																			
Preschool/Primary not completed	Percent																			
Primary completed	Percent																			
Lower secondary completed	Percent																			
Senior secondary completed	Percent																			
Post secondary technical																				
Tertiary	Percent																			
Other	Percent																			
Missing	Percent																			
<b>Labor force</b>																				
<b>Employment (15 years and above)</b>																				
Proportion of total population aged 15 yrs or more	Percent																			
Employed	Percent																			
of which females	Percent																			
<b>Employment Categories (15 and above) Primary employment</b>																				
<b>Wage workers</b>																				
Wage public	Percent																			
of which females	Percent																			
Wage private non-agriculture	Percent																			
of which females	Percent																			
Wage private agriculture	Percent																			
of which females	Percent																			
<b>Self employed/ non-wage</b>																				
Family enterprise	Percent																			
of which females	Percent																			
Family farmer	Percent																			
of which females	Percent																			
<b>Household Sources of Income (Primary &amp; Secondary employment)</b>																				
Agriculture	Percent																			
Non-Agriculture	Percent																			
Wage	Percent																			

Country Survey Name Year: Selected Social & MDG Indicators by Rural Urban and Consumption Level																			
Page 2 of 2																			
		National						Rural						Urban					
	Unit of measure	Poorest 20%	II	III	IV	Richest 20%	All	Poorest 20%	II	III	IV	Richest 20%	All	Poorest 20%	II	III	IV	Richest 20%	All
<b>Indicators</b>																			
<b>MDG1: Eradicate extreme poverty and hunger</b>																			
Mean monthly per capita expenditure	PPP2005																		
Poverty headcount under \$PPP1.25/day	Percent																		
Gini coefficient, consumption	Ratio																		
<b>Mean share of household expenditures</b>																			
Mean share of food expenditure	Percent																		
Mean share of health expenditure	Percent																		
Mean share of education expenditure	Percent																		
Public transportation	Percent																		
Private transportation (excludes fuel)	Percent																		
Fuel (for transportation)	Percent																		
Electricity	Percent																		
Gas	Percent																		
Transport and utility	Percent																		
Kerosene	Percent																		
Diesel (all purposes except transportation)	Percent																		
<b>Gross primary enrollment rate</b>																			
Female	Percent																		
Male	Percent																		
<b>Gross lower secondary enrollment ratio</b>																			
Female	Percent																		
Male	Percent																		
<b>Gross senior secondary enrollment ratio</b>																			
Female	Percent																		
Male	Percent																		
<b>Net primary enrollment rate</b>																			
Female	Percent																		
Male	Percent																		
<b>Primary completion rate on time</b>																			
Female	Percent																		
Male	Percent																		
<b>Net lower secondary enrollment ratio</b>																			
Female	Percent																		
Male	Percent																		
<b>Net senior secondary enrollment ratio</b>																			
Female	Percent																		
Male	Percent																		
<b>Tertiary enrollment over age 18-22*</b>																			
of which female	Percent																		
<b>Youth literacy rate ages 15-24</b>																			
Female	Percent																		
Male	Percent																		
<b>MDG4: Reduce child mortality; MDG5: Improve maternal health</b>																			
Distance to health centers < 5 km	Percent																		
Percent claiming to be sick last 2 wks	Percent																		
of which percent sought medical care	Percent																		
Birth assisted by trained staff	Percent																		
1-year-olds total immunization	Percent																		
1-year-olds immunization, measles	Percent																		
<b>MDG7: Ensure environment sustainability</b>																			
Owner occupancy rate	Percent																		
Access to flush/improved toilets	Percent																		
of which improved pit latrine	Percent																		
Access to garbage collection	Percent																		
Connected to electricity	Percent																		
of which solar/wind electricity	Percent																		
<b>Fuel use</b>																			
Firewood/charcoal	Percent																		
Kerosene/electricity/gas	Percent																		
Other	Percent																		
Missing	Percent																		
<b>Access to water</b>																			
Improved	Percent																		
Unimproved	Percent																		
Other	Percent																		
Missing	Percent																		

## 6.4 Appendix 4: Issues In Household Expenditure Aggregation

### Data cleaning

**Eliminating duplicates:** Before the aggregation, one should make sure that there is no duplicated household or individual id (This problem should have been eliminated at the data entering stage, this is just a quick check). There are two types of duplicate household id. One is that one household was entered twice, in which case all other variables should be also the same, such as physical features of the housing, rent paid, etc. In such case, one of the duplicated observations should be eliminated (STATA) can perform this task easily). Second type of duplicate is two or more households have identical households ID. This type of duplicate is hard to discover because one often takes two households as one household. However, unusual large household size, such as over 30, should warrant a check of the household to make sure that they are not two households. When there are two or more same ids signed to household members, if the two or more members have exact same information, such as age, sex, status enrollment, etc., a double entry of the same member is usually the case, and one should eliminate one of the observations. However, if they are two different individuals, one should reassign a new id to one of the members.

**Cleaning expenditure per capita variable:** Per capital annual food and non-food consumption is first adjusted by regional food price indices. Outliers (separately for food and non-food expenditure), greater than three standard deviations will then be replaced with the median values of the corresponding per capita food and non-food consumption, by region and by rural/urban areas. Total number of observations cleaned should be less than 1 percent of total observations. If it is greater than 1 percent of total observations, aggregation procedures should be rechecked to ascertain accuracy.

**Do not assign missing values:** A common mistake is to assign missing values with some meaningful values when aggregating a variable. To avoid incorrectly assigning missing values, avoid open ended grouping while coding. For example, do not use “if AGE<5 then CHDUNDR5=1”, instead, use “if 0<=AGE<=5 then CHDUNDR5=1”. In addition, the number of missing values must be checked before and after the aggregation to see if the numbers of missing values are consistent.

### Data Aggregation

**Purchased Food expenditure aggregation** in general is straight forward. For purchased food, one just adds up all values of relevant items according to the sub-categories. If the recall period is 2 weeks, the annualized expenditure will be derived by multiplying the reported values by 365/14, and if the recall period is one month, then by 365/30.

**Auto food expenditure aggregation** is more complicated. Food auto-consumption comes in two varieties: that which is produced by the household and the gifts in kind which the household consumes. Both must be valued, although the first may be the only one to immediately spring to the reader's mind.

If the total value of auto-consumption is not asked, in general, the valuation of auto-consumption is a two-step process. First, the prices are imputed for the goods that are auto-consumed, and then the aggregation is done as for purchased expenditure. If the survey had a community price questionnaire, this is usually the source of prices for auto-consumption. If such information is missing but prices can be inferred from household survey on auto consumption, then this is the source for prices. If neither method is possible, external sources of price data are used. An example is administrative records of market prices. In all these cases, there is some degree of



bias, since the market price for a good is not necessarily farm gate prices. There may be an intermediary who purchases at less than the market price, or there are travel and time costs associated with selling the good at market. However, it is better to have the information and (potentially) over-estimate the value than to ignore it entirely.

Generally, surveys collect information on food expenditure using multiple units of measure. A common unit of measure is necessary in order to compare prices; the simplest are often metric units like kilos and liters. Note that prices will not be the same for all units of measure: buying rice by the gram is likely to be more expensive than buying it in 50-kg bags (volume discount).

Once the price data is found, the level of disaggregation at which the prices are useful must be determined. The tradeoff is one of number of observations versus the household's reality. As the number of observations increases, the confidence in the mean (or median) price increases too; hence the imputed price is better. However, all enumeration areas in a district need not have the same market and growing conditions; the less fine the geography, the greater the risk of "noise" due to dissimilarities. This is a fine line, but generally 30 observations in for each product in each region are a minimum.

Once the prices are determined, the remainder of the work is identical to that of the purchased food consumption, except that the imputed prices are used to value the consumption. For purposes of validation and analysis, it is recommended that the same categories as for purchased expenditure be used.

**Education expenditure aggregation** should be annualized based on the time of school year. For example, monthly transportation fee should only be multiplied by the number of the months when children go to school, normally 9, to get the annual transportation expenditure.

**Health care expenditure aggregation** is more difficult. Health care expenditures are often asked for the last three months or so. Although if one was sick in the last three months, it does not necessarily mean that one was sick every three months. In practice, however, the three-month expenditure is often multiplied by 4 to get the annualized expenditure. This is more or less fine with expenditures on consultation fees and medication. For hospitalization fees, however, if it is lumpy, it should be treated as a onetime annual expenditure.

**Housing expenditure** is often only available for households who rent. For owner occupied houses, if mortgage or the market value of the rent is not asked code it as missing.<sup>2</sup>

---

<sup>2</sup> The most commonly used methodology is to use a hedonic regression to estimate the housing value for owner occupied houses. First, a regression model is fitted to households who paid rent, with dependent variable being the reported monthly rent, and independent variables being the physical features of the households, including latrines, number of rooms, materials of floor, wall and roof, the presence of a kitchen, the source of drinking water, the source of lighting power, and the cooking fuel. Other variables, known as "controls" are added to the regression to adjust for regional or rural-urban differences.

Once the model is done, a useful check of the quality is to use it to predict the rent for households that paid rent. If there is a systematic difference, something remains to be exploited. Otherwise, the desired result is small deviations between the actual and predicted rents. The parameters estimated from the hedonic regression are then used to estimate the rent for owner occupied houses. For the final rent that will be included in the total household expenditure, there are two choices. One is to use rent predicted by the regression for all the households, including those who paid rent. The reason of doing so is that regression often predicts rent somewhat lower than the actual rent paid. Using regression estimated rent for all the households puts everyone on equal footing. Another one is to use regression estimated rent for the owner occupied households only, the actual rent paid is used for renters.

**Durable goods and Lumpy expenditures are excluded:** This rule needs to be implemented with caution: in some countries one may find that many households report weddings, baptisms, and funerals in a survey. In the case of the 1994-95 Enquête Permanente de Conjoncture Economique et Sociale in Niger, over 60 percent (weighted) of all regions' households reported expenditure related to a marriage or a baptism. In this case, it is no longer an infrequent, uncommon expenditure item. The general rule is to omit lumpy expenditure, the ones that are not frequently or commonly spent, especially if they are large relative to total household expenditure, from the total household expenditure aggregate.

**Adjustment of cost of living** is to take into account that same money can buy different amount of goods in different regions. For example, a pound of beef would be much more expensive in a capital city than in a rural town where the cattle are raised nearby. To calculate regional price deflators, national average price and food share for the largest 20 items will be used as base price index. The regional price deflator is then calculated as  $P_r^k$  where  $P_r^k$  is price deflator for region  $r$ ,  $P^k$  is the national share of commodity "k" for the national food basket,  $P_r^k$  if the regional price for commodity "k", and  $P^k$  is the national average price for commodity "k".

**Common Sense:** Expenditure aggregation is easier said than done, many judgments must be made on individual cases. The most important rule perhaps is to use on the ground knowledge/ common sense while using statistical tools. The second important rule is to document in detail the steps used for aggregation. Finally, some basic statistics should be checked against reality.

## 7 Annexure

## 7.1 Annexure 1: Country Codes

AGO	Angola
BEN	Benin
BWA	Botswana
BFA	Burkina Faso
BDI	Burundi
CMR	Cameroon
CPV	Cape Verde
CAF	Central African Republic
TCD	Chad
COM	Comoros
ZAR	Congo, Dem. Rep.
COG	Congo, Rep.
CIV	Cote d'Ivoire
GNQ	Equatorial Guinea
ERI	Eritrea
ETH	Ethiopia
GAB	Gabon
GMB	Gambia, The
GHA	Ghana
GIN	Guinea
GNB	Guinea-Bissau
KEN	Kenya
LSO	Lesotho
LBR	Liberia
MDG	Madagascar
MWI	Malawi
MLI	Mali
MRT	Mauritania
MUS	Mauritius
MOZ	Mozambique
NAM	Namibia
NER	Niger
NGA	Nigeria
RWA	Rwanda
STP	Sao Tome and Principe
SEN	Senegal
SYC	Seychelles
SLE	Sierra Leone
SOM	Somalia
ZAF	South Africa
SDN	Sudan
SWZ	Swaziland
TZA	Tanzania
TGO	Togo
UGA	Uganda
ZMB	Zambia
ZWE	Zimbabwe

## 7.2 Annexure 2: International Standard Industrial Classification (ISIC)

The following is an extract from the International Standard Industrial Classification of Economic Activities<sup>3</sup> (ISIC) Rev. 3.1. It will be used to generate variables INDUSTRY\_1 and INDUSTRY\_2.

### I. Broad structure

The individual categories of ISIC have been aggregated into the following 17 sections:

Section	Divisions	Description
A	01, 02	Agriculture, hunting and forestry
B	05	Fishing
C	10-14	Mining and quarrying
D	15-37	Manufacturing
E	40, 41	Electricity, gas and water supply
F	45	Construction
G	50-52	Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods
H	55	Hotels and restaurants
I	60-64	Transport, storage and communications
J	65-67	Financial intermediation
K	70-74	Real estate, renting and business activities
L	75	Public administration and defense; compulsory social security
M	80	Education
N	85	Health and social work
O	90-93	Other community, social and personal service activities
P	95-97	Activities of private households as employers and undifferentiated production activities of private households
Q	99	Extraterritorial organizations and bodies

### II. Detailed structure

Section/Division Group Class Description

#### Section A Agriculture, hunting and forestry

##### Division 01 Agriculture, hunting and related service activities

011		Growing of crops; market gardening; horticulture
	0111	Growing of cereals and other crops n.e.c.
	0112	Growing of vegetables, horticultural specialties and nursery products
	0113	Growing of fruit, nuts, beverage and spice crops
012		Farming of animals
	0121	Farming of cattle, sheep, goats, horses, asses, mules and hinnies; dairy farming
	0122	Other animal farming; production of animal products n.e.c.
013	0130	Growing of crops combined with farming of animals (mixed farming)
014	0140	Agricultural and animal husbandry service activities, except veterinary activities
015	0150	Hunting, trapping and game propagation including related service activities

##### Division 02 Forestry, logging and related service activities

<sup>3</sup> <http://esa.un.org/unsd/cr/registry/regcst.asp?Cl=17&Lg=1>

020      0200      Forestry, logging and related service activities

## **Section B Fishing**

### **Division 05 Fishing, aquaculture and service activities incidental to fishing**

050      Fishing, aquaculture and service activities incidental to fishing  
          0501      Fishing  
          0502      Aquaculture

## **Section C Mining and quarrying**

### **Division 10 Mining of coal and lignite; extraction of peat**

101      1010      Mining and agglomeration of hard coal  
102      1020      Mining and agglomeration of lignite  
103      1030      Extraction and agglomeration of peat

### **Division 11 Extraction of crude petroleum and natural gas; service activities incidental to oil and gas extraction, excluding surveying**

111      1110      Extraction of crude petroleum and natural gas  
112      1120      Service activities incidental to oil and gas extraction excluding surveying

### **Division 12 Mining of uranium and thorium ores**

120      1200      Mining of uranium and thorium ores

### **Division 13 Mining of metal ores**

131      1310      Mining of iron ores  
132      1320      Mining of non-ferrous metal ores, except uranium and thorium ores

### **Division 14 Other mining and quarrying**

141      1410      Quarrying of stone, sand and clay  
142      Mining and quarrying n.e.c.  
          1421      Mining of chemical and fertilizer minerals  
          1422      Extraction of salt  
          1429      Other mining and quarrying n.e.c.

## **Section D Manufacturing**

### **Division 15 Manufacture of food products and beverages**

151      Production, processing and preservation of meat, fish, fruit, vegetables, oils and fats  
          1511      Production, processing and preserving of meat and meat products  
          1512      Processing and preserving of fish and fish products  
          1513      Processing and preserving of fruit and vegetables  
          1514      Manufacture of vegetable and animal oils and fats  
152      1520      Manufacture of dairy products  
153      Manufacture of grain mill products, starches and starch products, and prepared animal feeds  
          1531      Manufacture of grain mill products  
          1532      Manufacture of starches and starch products  
          1533      Manufacture of prepared animal feeds  
154      Manufacture of other food products  
          1541      Manufacture of bakery products  
          1542      Manufacture of sugar  
          1543      Manufacture of cocoa, chocolate and sugar confectionery  
          1544      Manufacture of macaroni, noodles, couscous and similar farinaceous products  
          1549      Manufacture of other food products n.e.c.  
155      Manufacture of beverages  
          1551      Distilling, rectifying and blending of spirits; ethyl alcohol production from fermented materials

- 1552 Manufacture of wines
- 1553 Manufacture of malt liquors and malt
- 1554 Manufacture of soft drinks; production of mineral waters

#### **Division 16 Manufacture of tobacco products**

- 160 1600 Manufacture of tobacco products

#### **Division 17 Manufacture of textiles**

- 171 Spinning, weaving and finishing of textiles
  - 1711 Preparation and spinning of textile fibres; weaving of textiles
  - 1712 Finishing of textiles
- 172 Manufacture of other textiles
  - 1721 Manufacture of made-up textile articles, except apparel
  - 1722 Manufacture of carpets and rugs
  - 1723 Manufacture of cordage, rope, twine and netting
  - 1729 Manufacture of other textiles n.e.c.
- 173 1730 Manufacture of knitted and crocheted fabrics and articles

#### **Division 18 Manufacture of wearing apparel; dressing and dyeing of fur**

- 181 1810 Manufacture of wearing apparel, except fur apparel
- 182 1820 Dressing and dyeing of fur; manufacture of articles of fur

#### **Division 19 Tanning and dressing of leather; manufacture of luggage, handbags, saddlery, harness and footwear**

- 191 Tanning and dressing of leather; manufacture of luggage, handbags, saddlery and harness
  - 1911 Tanning and dressing of leather
  - 1912 Manufacture of luggage, handbags and the like, saddlery and harness
- 192 1920 Manufacture of footwear

#### **Division 20 Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials**

- 201 2010 Sawmilling and planing of wood
- 202 Manufacture of products of wood, cork, straw and plaiting materials
  - 2021 Manufacture of veneer sheets; manufacture of plywood, laminboard, particle board and other panels and boards
  - 2022 Manufacture of builders' carpentry and joinery
  - 2023 Manufacture of wooden containers
  - 2029 Manufacture of other products of wood; manufacture of articles of cork, straw and plaiting materials

#### **Division 21 Manufacture of paper and paper products**

- 210 Manufacture of paper and paper products
  - 2101 Manufacture of pulp, paper and paperboard
  - 2102 Manufacture of corrugated paper and paperboard and of containers of paper and paperboard
  - 2109 Manufacture of other articles of paper and paperboard

#### **Division 22 Publishing, printing and reproduction of recorded media**

- 221 Publishing
  - 2211 Publishing of books, brochures and other publications
  - 2212 Publishing of newspapers, journals and periodicals
  - 2213 Publishing of music
  - 2219 Other publishing
- 222 Printing and service activities related to printing
  - 2221 Printing

	2222	Service activities related to printing
223	2230	Reproduction of recorded media

### **Division 23 Manufacture of coke, refined petroleum products and nuclear fuel**

231	2310	Manufacture of coke oven products
232	2320	Manufacture of refined petroleum products
233	2330	Processing of nuclear fuel

### **Division 24 Manufacture of chemicals and chemical products**

241		Manufacture of basic chemicals
	2411	Manufacture of basic chemicals, except fertilizers and nitrogen compounds
	2412	Manufacture of fertilizers and nitrogen compounds
	2413	Manufacture of plastics in primary forms and of synthetic rubber
242		Manufacture of other chemical products
	2421	Manufacture of pesticides and other agrochemical products
	2422	Manufacture of paints, varnishes and similar coatings, printing ink and mastics
	2423	Manufacture of pharmaceuticals, medicinal chemicals and botanical products
	2424	Manufacture of soap and detergents, cleaning and polishing preparations, perfumes and toilet preparations
	2429	Manufacture of other chemical products n.e.c.
243	2430	Manufacture of man-made fibres

### **Division 25 Manufacture of rubber and plastics products**

251		Manufacture of rubber products
	2511	Manufacture of rubber tyres and tubes; retreading and rebuilding of rubber tyres
	2519	Manufacture of other rubber products
252	2520	Manufacture of plastics products

### **Division 26 Manufacture of other non-metallic mineral products**

261	2610	Manufacture of glass and glass products
269		Manufacture of non-metallic mineral products n.e.c.
	2691	Manufacture of non-structural non-refractory ceramic ware
	2692	Manufacture of refractory ceramic products
	2693	Manufacture of structural non-refractory clay and ceramic products
	2694	Manufacture of cement, lime and plaster
	2695	Manufacture of articles of concrete, cement and plaster
	2696	Cutting, shaping and finishing of stone
	2699	Manufacture of other non-metallic mineral products n.e.c.

### **Division 27 Manufacture of basic metals**

271	2710	Manufacture of basic iron and steel
272	2720	Manufacture of basic precious and non-ferrous metals
273		Casting of metals
	2731	Casting of iron and steel
	2732	Casting of non-ferrous metals

### **Division 28 Manufacture of fabricated metal products, except machinery and equipment**

281		Manufacture of structural metal products, tanks, reservoirs and steam generators
	2811	Manufacture of structural metal products
	2812	Manufacture of tanks, reservoirs and containers of metal
	2813	Manufacture of steam generators, except central heating hot water boilers
289		Manufacture of other fabricated metal products; metalworking service activities
	2891	Forging, pressing, stamping and roll-forming of metal; powder metallurgy
	2892	Treatment and coating of metals; general mechanical engineering on a fee or contract basis
	2893	Manufacture of cutlery, hand tools and general hardware
	2899	Manufacture of other fabricated metal products n.e.c.

**Division 29 Manufacture of machinery and equipment n.e.c.**

291		Manufacture of general-purpose machinery
	2911	Manufacture of engines and turbines, except aircraft, vehicle and cycle engines
	2912	Manufacture of pumps, compressors, taps and valves
	2913	Manufacture of bearings, gears, gearing and driving elements
	2914	Manufacture of ovens, furnaces and furnace burners
	2915	Manufacture of lifting and handling equipment
	2919	Manufacture of other general-purpose machinery
292		Manufacture of special-purpose machinery
	2921	Manufacture of agricultural and forestry machinery
	2922	Manufacture of machine tools
	2923	Manufacture of machinery for metallurgy
	2924	Manufacture of machinery for mining, quarrying and construction
	2925	Manufacture of machinery for food, beverage and tobacco processing
	2926	Manufacture of machinery for textile, apparel and leather production
	2927	Manufacture of weapons and ammunition
	2929	Manufacture of other special-purpose machinery
293	2930	Manufacture of domestic appliances n.e.c.

**Division 30 Manufacture of office, accounting and computing machinery**

300	3000	Manufacture of office, accounting and computing machinery
-----	------	---

**Division 31 Manufacture of electrical machinery and apparatus n.e.c.**

311	3110	Manufacture of electric motors, generators and transformers
312	3120	Manufacture of electricity distribution and control apparatus
313	3130	Manufacture of insulated wire and cable
314	3140	Manufacture of accumulators, primary cells and primary batteries
315	3150	Manufacture of electric lamps and lighting equipment
319	3190	Manufacture of other electrical equipment n.e.c.

**Division 32 Manufacture of radio, television and communication equipment and apparatus**

321	3210	Manufacture of electronic valves and tubes and other electronic components
322	3220	Manufacture of television and radio transmitters and apparatus for line telephony and line telegraphy
323	3230	Manufacture of television and radio receivers, sound or video recording or reproducing apparatus, and associated goods

**Division 33 Manufacture of medical, precision and optical instruments, watches and clocks**

331		Manufacture of medical appliances and instruments and appliances for measuring, checking, testing, navigating and other purposes, except optical instruments
	3311	Manufacture of medical and surgical equipment and orthopaedic appliances
	3312	Manufacture of instruments and appliances for measuring, checking, testing, navigating and other purposes, except industrial process control equipment
	3313	Manufacture of industrial process control equipment
332	3320	Manufacture of optical instruments and photographic equipment
333	3330	Manufacture of watches and clocks

**Division 34 Manufacture of motor vehicles, trailers and semi-trailers**

341	3410	Manufacture of motor vehicles
342	3420	Manufacture of bodies (coachwork) for motor vehicles; manufacture of trailers and semi-trailers
343	3430	Manufacture of parts and accessories for motor vehicles and their engines



**Division 35 Manufacture of other transport equipment**

351		Building and repairing of ships and boats
	3511	Building and repairing of ships
	3512	Building and repairing of pleasure and sporting boats
352	3520	Manufacture of railway and tramway locomotives and rolling stock
353	3530	Manufacture of aircraft and spacecraft
359		Manufacture of transport equipment n.e.c.
	3591	Manufacture of motorcycles
	3592	Manufacture of bicycles and invalid carriages
	3599	Manufacture of other transport equipment n.e.c.

**Division 36 Manufacture of furniture; manufacturing n.e.c.**

361	3610	Manufacture of furniture
369		Manufacturing n.e.c.
	3691	Manufacture of jewellery and related articles
	3692	Manufacture of musical instruments
	3693	Manufacture of sports goods
	3694	Manufacture of games and toys
	3699	Other manufacturing n.e.c.

**Division 37 Recycling**

371	3710	Recycling of metal waste and scrap
372	3720	Recycling of non-metal waste and scrap

**Section E Electricity, gas and water supply****Division 40 Electricity, gas, steam and hot water supply**

401	4010	Production, transmission and distribution of electricity
402	4020	Manufacture of gas; distribution of gaseous fuels through mains
403	4030	Steam and hot water supply

**Division 41 Collection, purification and distribution of water**

410	4100	Collection, purification and distribution of water
-----	------	--

**Section F Construction****Division 45 Construction**

451	4510	Site preparation
452	4520	Building of complete constructions or parts thereof; civil engineering
453	4530	Building installation
454	4540	Building completion
455	4550	Renting of construction or demolition equipment with operator

**Section G Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods****Division 50 Sale, maintenance and repair of motor vehicles and motorcycles; retail sale of automotive fuel**

501	5010	Sale of motor vehicles
502	5020	Maintenance and repair of motor vehicles
503	5030	Sale of motor vehicle parts and accessories
504	5040	Sale, maintenance and repair of motorcycles and related parts and accessories
505	5050	Retail sale of automotive fuel

**Division 51 Wholesale trade and commission trade, except of motor vehicles and motorcycles**

511	5110	Wholesale on a fee or contract basis
512		Wholesale of agricultural raw materials, live animals, food, beverages and tobacco
	5121	Wholesale of agricultural raw materials and live animals

	5122	Wholesale of food, beverages and tobacco
513		Wholesale of household goods
	5131	Wholesale of textiles, clothing and footwear
	5139	Wholesale of other household goods
514		Wholesale of non-agricultural intermediate products, waste and scrap
	5141	Wholesale of solid, liquid and gaseous fuels and related products
	5142	Wholesale of metals and metal ores
	5143	Wholesale of construction materials, hardware, plumbing and heating equipment and supplies
	5149	Wholesale of other intermediate products, waste and scrap
515		Wholesale of machinery, equipment and supplies
	5151	Wholesale of computers, computer peripheral equipment and software
	5152	Wholesale of electronic and telecommunications parts and equipment
	5159	Wholesale of other machinery, equipment and supplies
519	5190	Other wholesale

## **Division 52 Retail trade, except of motor vehicles and motorcycles; repair of personal and household goods**

521		Non-specialized retail trade in stores
	5211	Retail sale in non-specialized stores with food, beverages or tobacco predominating
	5219	Other retail sale in non-specialized stores
522	5220	Retail sale of food, beverages and tobacco in specialized stores
523		Other retail trade of new goods in specialized stores
	5231	Retail sale of pharmaceutical and medical goods, cosmetic and toilet articles
	5232	Retail sale of textiles, clothing, footwear and leather goods
	5233	Retail sale of household appliances, articles and equipment
	5234	Retail sale of hardware, paints and glass
	5239	Other retail sale in specialized stores
524	5240	Retail sale of second-hand goods in stores
525		Retail trade not in stores
	5251	Retail sale via mail order houses
	5252	Retail sale via stalls and markets
	5259	Other non-store retail sale
526	5260	Repair of personal and household goods

## **Section H Hotels and restaurants**

### **Division 55 Hotels and restaurants**

551	5510	Hotels; camping sites and other provision of short-stay accommodation
552	5520	Restaurants, bars and canteens

## **Section I Transport, storage and communications**

### **Division 60 Land transport; transport via pipelines**

601	6010	Transport via railways
602		Other land transport
	6021	Other scheduled passenger land transport
	6022	Other non-scheduled passenger land transport
	6023	Freight transport by road
603	6030	Transport via pipelines

### **Division 61 Water transport**

611	6110	Sea and coastal water transport
612	6120	Inland water transport

### **Division 62 Air transport**

621	6210	Scheduled air transport
622	6220	Non-scheduled air transport

**Division 63 Supporting and auxiliary transport activities; activities of travel agencies**

- 630 Supporting and auxiliary transport activities; activities of travel agencies
  - 6301 Cargo handling
  - 6302 Storage and warehousing
  - 6303 Other supporting transport activities
  - 6304 Activities of travel agencies and tour operators; tourist assistance activities n.e.c.
  - 6309 Activities of other transport agencies

**Division 64 Post and telecommunications**

- 641 Post and courier activities
  - 6411 National post activities
  - 6412 Courier activities other than national post activities
- 642 6420 Telecommunications

**Section J Financial intermediation****Division 65 Financial intermediation, except insurance and pension funding**

- 651 Monetary intermediation
  - 6511 Central banking
  - 6519 Other monetary intermediation
- 659 Other financial intermediation
  - 6591 Financial leasing
  - 6592 Other credit granting
  - 6599 Other financial intermediation n.e.c.

**Division 66 Insurance and pension funding, except compulsory social security**

- 660 Insurance and pension funding, except compulsory social security
  - 6601 Life insurance
  - 6602 Pension funding
  - 6603 Non-life insurance

**Division 67 Activities auxiliary to financial intermediation**

- 671 Activities auxiliary to financial intermediation, except insurance and pension funding
  - 6711 Administration of financial markets
  - 6712 Security dealing activities
  - 6719 Activities auxiliary to financial intermediation n.e.c.
- 672 6720 Activities auxiliary to insurance and pension funding

**Section K Real estate, renting and business activities****Division 70 Real estate activities**

- 701 7010 Real estate activities with own or leased property
- 702 7020 Real estate activities on a fee or contract basis

**Division 71 Renting of machinery and equipment without operator and of personal and household goods**

- 711 Renting of transport equipment
  - 7111 Renting of land transport equipment
  - 7112 Renting of water transport equipment
  - 7113 Renting of air transport equipment
- 712 Renting of other machinery and equipment
  - 7121 Renting of agricultural machinery and equipment
  - 7122 Renting of construction and civil engineering machinery and equipment
  - 7123 Renting of office machinery and equipment (including computers)
  - 7129 Renting of other machinery and equipment n.e.c.
- 713 7130 Renting of personal and household goods n.e.c.

**Division 72 Computer and related activities**

721	7210	Hardware consultancy
722		Software publishing, consultancy and supply
	7221	Software publishing
	7229	Other software consultancy and supply
723	7230	Data processing
724	7240	Database activities and online distribution of electronic content
725	7250	Maintenance and repair of office, accounting and computing machinery
729	7290	Other computer-related activities

**Division 73 Research and development**

731	7310	Research and experimental development on natural sciences and engineering (NSE)
732	7320	Research and experimental development on social sciences and humanities (SSH)

**Division 74 Other business activities**

741		Legal, accounting, bookkeeping and auditing activities; tax consultancy; market research and public opinion polling; business and management consultancy
	7411	Legal activities
	7412	Accounting, bookkeeping and auditing activities; tax consultancy
	7413	Market research and public opinion polling
	7414	Business and management consultancy activities
742		Architectural, engineering and other technical activities
	7421	Architectural and engineering activities and related technical consultancy
	7422	Technical testing and analysis
743	7430	Advertising
749		Business activities n.e.c.
	7491	Labour recruitment and provision of personnel
	7492	Investigation and security activities
	7493	Building-cleaning and industrial-cleaning activities
	7494	Photographic activities
	7495	Packaging activities
	7499	Other business activities n.e.c.

**Section L Public administration and defence; compulsory social security****Division 75 Public administration and defence; compulsory social security**

751		Administration of the State and the economic and social policy of the community
	7511	General (overall) public service activities
	7512	Regulation of the activities of agencies that provide health care, education, cultural services and other social services, excluding social security
	7513	Regulation of and contribution to more efficient operation of business
	7514	Supporting service activities for the government as a whole
752		Provision of services to the community as a whole
	7521	Foreign affairs
	7522	Defence activities
	7523	Public order and safety activities
753	7530	Compulsory social security activities

**Section M Education****Division 80 Education**

801	8010	Primary education
802		Secondary education
	8021	General secondary education
	8022	Technical and vocational secondary education
803	8030	Higher education
809	8090	Other education

**Section N Health and social work****Division 85 Health and social work**

- 851 Human health activities
  - 8511 Hospital activities
  - 8512 Medical and dental practice activities
  - 8519 Other human health activities
- 852 8520 Veterinary activities
- 853 Social work activities
  - 8531 Social work activities with accommodation
  - 8532 Social work activities without accommodation

**Section O Other community, social and personal service activities****Division 90 Sewage and refuse disposal, sanitation and similar activities**

- 900 9000 Sewage and refuse disposal, sanitation and similar activities

**Division 91 Activities of membership organizations n.e.c.**

- 911 Activities of business, employers and professional organizations
  - 9111 Activities of business and employers organizations
  - 9112 Activities of professional organizations
- 912 9120 Activities of trade unions
- 919 Activities of other membership organizations
  - 9191 Activities of religious organizations
  - 9192 Activities of political organizations
  - 9199 Activities of other membership organizations n.e.c.

**Division 92 Recreational, cultural and sporting activities**

- 921 Motion picture, radio, television and other entertainment activities
  - 9211 Motion picture and video production and distribution
  - 9212 Motion picture projection
  - 9213 Radio and television activities
  - 9214 Dramatic arts, music and other arts activities
  - 9219 Other entertainment activities n.e.c.
- 922 9220 News agency activities
- 923 Library, archives, museums and other cultural activities
  - 9231 Library and archives activities
  - 9232 Museums activities and preservation of historic sites and buildings
  - 9233 Botanical and zoological gardens and nature reserves activities
- 924 Sporting and other recreational activities
  - 9241 Sporting activities
  - 9249 Other recreational activities

**Division 93 Other service activities**

- 930 Other service activities
  - 9301 Washing and (dry-)cleaning of textile and fur products
  - 9302 Hairdressing and other beauty treatment
  - 9303 Funeral and related activities
  - 9309 Other service activities n.e.c.

**Section P Activities of private households as employers and undifferentiated production activities of private households****Division 95 Activities of private households as employers of domestic staff**

- 950 9500 Activities of private households as employers of domestic staff

**Division 96 Undifferentiated goods-producing activities of private households for own use**

- 960 9600 Undifferentiated goods-producing activities of private households for own use

**Division 97 Undifferentiated service-producing activities of private households for own use**

970 9700 Undifferentiated service-producing activities of private households for own use

**Section Q Extraterritorial organizations and bodies**

**Division 99 Extraterritorial organizations and bodies**

990 9900 Extraterritorial organizations and bodies

The Classifications registry keeps updated information on Statistical Classifications maintained by the United Nations Statistics Division (UNSD). For the SHIP we will use ISIC rev. 3.1

For generating EMP\_IND, the classification will be as below.

For standardized file, this classification is used:-

A + B	=	Agriculture and fishing
C	=	Mining and quarrying
D	=	Manufacturing
E	=	Electricity, gas and water supply
F	=	Construction
G + H	=	Commerce
I	=	Transport, storage and communication
J + K	=	Financial, insurance and real estate
L	=	Public administration
M, N	=	Education, health and social work
O + P + Q	=	Other services

## REFERENCES

- Ahmad S., (1980), *Approaches to purchasing power parity and real product comparisons using shortcuts and reduced information*, World Bank Staff working paper No. SWP 418 1980/10/01
- Berghahl M., Japac L., Madaleno M., Signore M., Tzougas I., *Tools for Standardising and Improving the Quality of Statistics Production Processes and other Operations*, Paper presented during the Stockholm 2001 conference, May 14-15, 2001.
- Carson C. S., (2001), *Toward a Framework for Assessing Data Quality*, Paper presented during the Stockholm 2001 conference, May 14-15, 2001.
- Deaton A. and Zaidi S., (2002), *Guidelines for Constructing Consumption Aggregates for Welfare Analysis*, World Bank *LSMS Working Paper No. 135*
- Dupriez O., (2005), *Poverty PPPs: Building a Database on Household Consumption Profiles by ICP Basic Heading*.
- Eurostat, *Assessment of the Quality in Statistics*, Paper presented during the Stockholm 2001 conference, May 14-15, 2001.
- Haworth M., Bergall M., Booleman M., Jones T. and Madaleno M., (2001), *Leadership Expert Group on Quality; Quality Framework Chapter*, Paper presented during the Stockholm 2001 conference, May 14-15, 2001.
- <http://www.cdc.gov>
- <http://esa.un.org/unsd/cr/registry/regcst.asp?Cl=17&Lq=1>
- <http://unescostat.unesco.org/en/stats/stats0.htm>
- <http://www.ilo.org/public/english/bureau/stat/class/icse.htm>
- <http://www.ilo.org/public/english/bureau/stat/res/ecacpop.htm>
- <http://www.ilo.org/public/english/employment/skills/informal/index.htm>
- <http://www.who.int/nutgrowthdb/reference/en/>
- Husmanns R., Mehran F., Verma V., (1990), *Surveys of economically active population, employment, unemployment and underemployment: An ILO manual on concepts and methods*.
- Irving Kravis B., (1986), *The three faces of the International Comparison Project*, *Research Observer* 1 No. 1 January 1986.
- Isenman P., (1979), *Inter-country comparison of "Real" (PPP) incomes : revised estimates and unresolved questions*, World Bank Staff working paper No. SWP 358 1979/09/01.
- Japac L., *Survey of Quality Practices in National Statistical Institutes*, Paper presented during the Stockholm 2001 conference, May 14-15, 2001.
- Kravis I. B., (1986), *The Three faces of the International Comparison Project*, *Research Observer* 1, no. 1.

- Lanjouw P., Prennushi G., Zaidi S. (1996), Building Blocks for a Consumption-Based Analysis of Poverty in Nepal, Washington DC, World Bank.
- Litchfield J. A., (1999), Inequality: methods and Tools, World Bank.
- Minhas, B. S., (1988), "Validation of large-scale sample survey data: case of NSS household consumption expenditure," Sankhya, *Series B 50, Supplement*, 1-63.
- Mungai R. and Marchant T., (2003), Establishing Comparability of Household Data across Countries and Time, Topic IPM 35 Designing Surveys to Provide Cross-National Comparable Data, *Bulletin of the International Statistical Institute 54<sup>th</sup> Session*.
- Ravallion M., (1998), Poverty Lines in Theory and Practice, World Bank *LSMS Working Paper No. 88*
- Ravallion M., (1992), Poverty Comparisons: A Guide to Concepts and Methods, World Bank *LSMS Working Paper No. 133*
- Statistics Canada, (1998), Third Ed., *Statistics Canada Quality Guidelines*,
- World Bank, (2007), *World Development Indicators 2007*, Washington DC, World Bank.
- World Bank, (2001), *World Development Indicators 2001*, Washington DC, World Bank.
- World Bank, (1999), Poverty Reduction and the World Bank: Progress in Fiscal 1998, Washington DC, World Bank.S