

# Building a household consumption database for the calculation of poverty PPPs

## Technical note

DRAFT 1.0

Olivier Dupriez, World Bank  
March 2007

# 1. Introduction

The 2003-2006 round of the International Comparisons Program (ICP) is a global statistical initiative established to produce internationally comparable price levels, expenditure values, and Purchasing Power Parity (PPP) estimates.

Among other applications, PPPs are used for the computation of global and regional poverty estimates in countries where poverty is prevalent. The current practice makes use of the \$1/day and \$2/day poverty lines, converted into respective national currency units by using available PPP estimates. Over the last few years, various limitations of the current approach have been identified. One of them is that the currently available PPPs do not adequately represent the consumption patterns of the poor.<sup>1</sup>

Purchasing Power Parities are averages of price ratios between countries. Their calculation requires:

- (i) *A commonly agreed list of goods and services representative of the consumption patterns of the populations of interest.* The ICP subdivided GDP expenditures into 155 standard categories called basic headings, 110 of which refer to household consumption.
- (ii) *A set of average prices per basic heading, for each country.* Price data are collected in each ICP-participating country for 1,000-plus closely specified items, which are then averaged by basic heading.
- (iii) *The weight of each basic heading in the total consumption of the population of interest.* Available PPPs were calculated using weights obtained from national accounts data which represent national consumption patterns, not the specific—and potentially significantly different—consumption patterns of the poor.

A Poverty Advisory Group established under the ICP Global Office recommended that poverty-specific PPPs be computed for countries where poverty is prevalent. To compute such poverty PPPs, a specific set of weighting coefficients representing the consumption patterns of the poor must be derived for each country. The data needed to establish such sets of weights can only be obtained from nationally representative sample household surveys.

Practically, deriving such sets of weights requires the production of a database on household consumption by basic heading, for as many ICP countries as possible. To maximize its usefulness and versatility, a set of additional variables providing information on household characteristics is included in the database. This technical note aims to describe the database and the procedures applied to create it.

The output database provides data in a common, user friendly format. It must be noted however that the use of a common data dictionary does not make data fully comparable across countries. Data were obtained from national sample surveys, characterized by major differences in their scope, coverage, timing and methods. These differences make it impossible to harmonize the data.

As of March 2007, the database covers 63 countries. For each country, the most recent dataset accessible to us was processed. In some cases, more appropriate or more recent surveys may exist and we will seek to obtain authorization to use them. To make the database more relevant for other research projects, we will also try and process data for more than one point in time per country. These data are however not yet accessible to the public. The ICP will seek approval from countries to disseminate them.

In this note, we only present data for 10 countries. The list of the ten surveys is in appendix 1. These data must be considered as provisional as adjustments and quality control are still being made.

---

<sup>1</sup> Another limitation results from the fact that the poor do not necessarily purchase goods from the same type of outlets as the non poor and that the typical size of their purchases is different as well, i.e. that the set of prices for the poor might be different from the set of prices for the non poor.

## 2. The output

The work consists of deriving a standardized set of data files from existing household survey datasets, for as many ICP countries as possible. For each country, the output of this work includes these standardized data files, the Stata programs used to generate them (and the corresponding Stata log files), and a set of Excel summary tables. More precisely, the following output is generated:

1. Three Stata data files per country:
  - a. One file named *cccyyyy\_hld.dta* (where *ccc* is the ISO 3-letter country code, and *yyyy* the reference year) contains the household characteristics variables (including the household total annual consumption). This file contains one record per item per household.
  - b. One file named *cccyyyy\_ori.dta* provides household consumption by item, based on the country-specific list of items covered by the survey. One variable links each item to the corresponding ICP basic heading. This file contains one record per item consumed per household.
  - c. One file named *cccyyyy\_ppp.dta* provides data on household consumption by ICP basic heading. This file contains one record per basic heading consumed per household.A detailed data dictionary is provided in appendix 2.

2. One Stata program per country, and the related generic Stata programs and log files:

The standardized data files listed above are generated using various Stata programs (Version 8 or 9), which constitute important metadata. One single program named *cccyyyy\_prp\_ppp.do* file is produced for each country, which contains all instructions needed to generate the standardized file from the original dataset. Running this file requires some generic sub-program files (*hh\_label\_eng.do*, *ex\_label\_eng.do*, and *icp\_check.do*), and lookup data files (*basic\_headings.dta*, and *bh\_nat\_accounts.dta*).

The program files generate various log files which are also preserved as they constitute important metadata: *cccyyyy\_prp.log*, *cccyyyy\_chk.log*, and *cccyyyy\_fix.log*. These log files are in text format, readable by any text editor or word processor. They include various quality control tables.

3. Summary country tables:

Summary tables are generated (in Excel format). These tables provide the calculated consumption shares, as well as various information on the survey questionnaire content and others. An example of one of these tables is presented in appendix 3.

We insist again on the fact that these datasets are “standardized” only to the extent that they use a common data dictionary, and that we tried to apply similar approaches to aggregate and annualize consumption data and to fix outliers. But as there are considerable differences between questionnaires and data collection methods (period of data collection, sampling, recall periods, etc.), the data cannot be made fully comparable.

### 3. The Process

For all countries, the standardization process starts with the original survey datasets provided by the national data producers. These datasets are supposed to be the “final”, i.e. fully edited. We avoid using pre-aggregated data generated by other users, unless the source of these data files is known and sufficient documentation is available.

The process consists of six main tasks:

1. Extracting household characteristics
2. Calculating annual consumption for all goods and services covered by the survey
3. Detecting and fixing outliers in consumption values
4. Mapping all goods and services to the corresponding ICP basic headings
5. Splitting the values stored in “fake” basic headings
6. Running quality control tables

These steps are described in detail below, and presented in fig.1. The process is implemented using Stata (Version 8 and 9).

Two more steps will be added at a later stage:

7. Anonymizing datasets
8. Packaging data and documentation for public dissemination (when authorized by countries)

#### 3.1. *Extracting household characteristics*

Generating the household-level file *cccyyyy\_hld.dta* consists mainly of extracting and recoding variables from the source data files. This is a straightforward process for most variables, but difficult for others (e.g., mapping the country-specific variable “level of education” to the standard one).

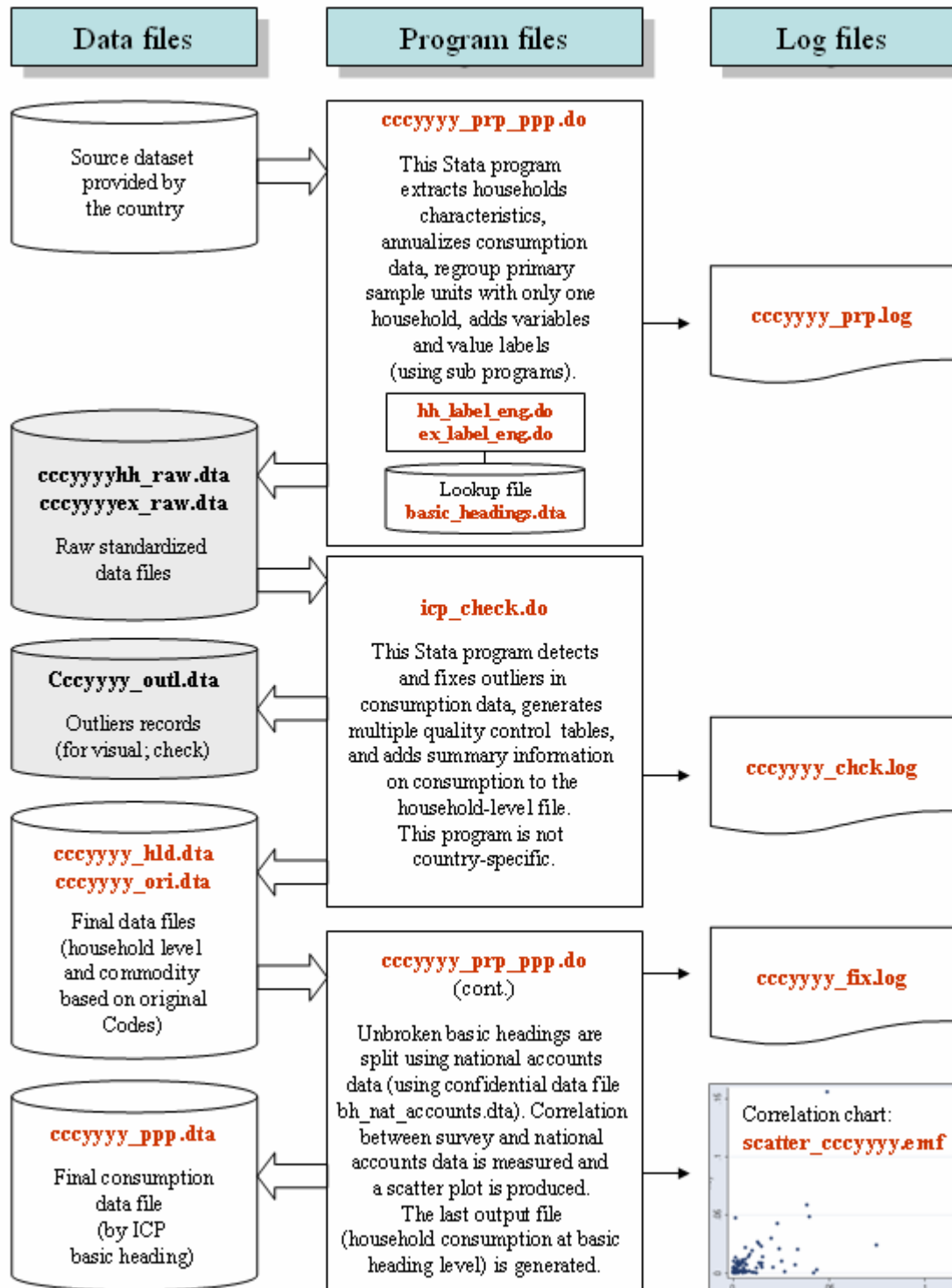
To allow calculation of sampling errors, the data files include variables identifying the stratum and primary sample unit (psu) for each record. Primary sample units containing only one household are grouped with their neighbor. In many cases however, datasets are provided with very little –if any– documentation of the sampling methodology. The creation of the variables *stratum* and *psu* is often challenging, and in some cases is the result of a best-guess.

#### 3.2. *Annualizing consumption values*

Survey questionnaires collect data using recall period that vary depending on the type of goods and services. All consumption data have to be annualized. In some cases, the process is straightforward and simply consists of applying a multiplying factor to the data. This is the case for most purchased food products and regularly purchased non-food products and services. It becomes more complex for home produced and received goods and services, for which prices have to be imputed. It is also more complex for durables goods, for which an annual use value has to be calculated when possible (not all questionnaires include the necessary information) using depreciation rates. Another major problem is the assessment of rental value (using regression models) for owner-occupied dwellings, in countries where rental market is often very limited. Annualizing health expenditure is also impossible in many cases, when a monthly recall period was used (applying a multiplying factor to such data of 12 would not make sense).

When available, we reuse (or adapt) the programs that were prepared by the data producer to produce the consumption aggregates, in order to maintain consistency with official estimates. This is however rarely possible, as these programs are almost never provided, making replication almost impossible.

**Fig. 1 – The data preparation process**



**Notes:**

In file names: ccc = ISO country code, and yyyy = survey reference year  
 [.log] files are text files readable in any text editor. [.do] files are Stata syntax programs.  
 [.dta] files are Stata data files (version 8). [.emf] is a Windows extended metafile.

### 3.3. Detecting and fixing outliers in consumption data

Datasets obtained from countries are supposed to be “clean”. Many however contain outlying values in consumption data. We apply a standard control to detect and fix them. Outliers are identified and fixed at the commodity-level. As we cannot expect all households to consume a minimum of each item, we only detect the “top” outliers, not the “bottom” ones.

For some items (e.g., food, transport services, personal effects), the outliers are detected using *per capita* values. For others (e.g., TV, car, rent), the detection is made using *per household* values. We consider as an outlier any value that exceeds the amount consumed at the 75<sup>th</sup> percentile plus 5 times the interquartile range<sup>2</sup>:

$$\text{outlier if value} > q75 + 5 * \text{IQR}$$

Outlying values are replaced by the weighted mean of the valid positive values (non outlier and > 0), calculated separately for urban and rural areas.

This detection rule is relatively conservative, and the resulting proportion of outlying records is low (see table 1). The impact of fixing these outliers on the total consumption and on its distribution is however significant. In one extreme case (Bolivia), fixing the 1.68 percent of outlying records reduces the mean per capita consumption by 44.3%. Table 1 also presents the impact on the Gini coefficient, and on the ratio of consumption of the 5<sup>th</sup> to 1<sup>st</sup> population quintiles.

**Table 1. Impact of outliers fixing on mean per capita consumption and inequality indicators**

	Country									
	Latin America						Asia			
	ARG	BOL	BRA	COL	PER	PRY	BGD	IDN	LKA	THA
% records outliers	1.21	1.68	1.62	1.98	1.00	1.32	0.86	1.07	0.90	1.14
<b>Mean per capita consumption</b> (before and after fixing outliers)										
Before fix	2,894	8,861	4,458	3,073,842	2,653	3,846,610	10,211	2,589,971	32,911	32,941
After fix	2,697	4,935	4,098	2,850,889	2,450	3,495,931	9,638	2,415,341	30,674	30,768
Difference (pc)	-6.81	-44.31	-8.08	-7.25	-7.65	-9.12	-5.61	-6.74	-6.80	-6.60
<b>Gini coefficient</b> (before and after fixing outliers)										
Before fix	0.466	0.696	0.543	0.485	0.436	0.489	0.327	0.333	0.365	0.425
After fix	0.438	0.488	0.519	0.457	0.404	0.456	0.298	0.299	0.332	0.396
<b>Ratio 5<sup>th</sup> to 1<sup>st</sup> quintile</b> (before and after fixing outliers)										
Before fix	23.64	74.64	37.29	25.99	15.6	2755	7.23	7.33	9.12	13.00
After fix	20.17	25.16	32.13	22.11	13.1	22.86	6.18	6.11	7.62	11.14

Notes:

- AGR=Argentina; BOL=Bolivia; BRA=Brazil; COL=Colombia; PER=Peru; PRY=Paraguay; BGD=Bangladesh; IDN=Indonesia; LKA=Sri Lanka; THA=Thailand.
- Data refer to survey year as indicated in Appendix 1
- The mean per capita consumption is in local currency
- The Gini coefficient is calculated on the per capita expenditure
- Quintiles are population quintiles, by per capita expenditure (the 1<sup>st</sup> quintile being the poorest)

<sup>2</sup> The interquartile range (IQR) is the range between the third and first quartiles.

The Stata program used to detect and fix outliers generates various control tables available in the log files showing the number and proportion of outliers by product (separately for purchased, home produced, and received items). Outliers in food consumption often seem to come from errors in quantity measurement units (many values are around 1,000 or 100 times the mean of valid values, indicating that grams (or ml) and kilos (or liters) may have been mixed, or that decimal points may have been missed). Interpretation of outliers in items such as education and health is more difficult.

Further work is needed to better understand the outliers and define the best method to fix them. This issue should receive greater attention from data analysts.

### **3.4. Mapping commodities to ICP basic headings**

The ICP categorizes household consumption into 110 basic headings (which correspond largely to the COICOP classification). These basic headings can be aggregated into 91 classes, 43 groups, and 13 categories.<sup>3</sup> Out of these 110 basic headings, 107 can possibly be obtained from household surveys. Three of them cannot:

- Financial intermediation indirectly measured (FISIM);
- Purchases by residential households in the rest of the world (survey questionnaires do not distinguish the place of purchase; we therefore assume that all purchased items have been acquired in the country);
- Purchases by non-residential households in the economic territory of the country (non residential household are not in the sample frames, and not covered by household budget surveys).

None of the household survey questionnaires has been specifically designed to collect data by basic heading. Some have been designed based on the COICOP classification, and provide a relatively good match with the ICP basic headings. But in many cases, goods and services are not provided with sufficient detail, and/or do not cover all basic headings.

To allow all survey records to be mapped to an ICP basic heading, we created 43 additional “fake” basic headings. They aim to provide a correspondence for items in the questionnaire that would correspond to more than one basic heading. For example, if the questionnaire collected data on “Gas and electricity”, the value could not be mapped to an ICP basic heading as the ICP has two distinct basic headings for “Electricity” and “Gas”. The value will then be mapped to our “fake” basic heading “Electricity, gas and other fuels”. All basic headings in Table 2 whose label starts with “UNBR” correspond to such “fake” (or “unbroken) basic headings, created for the sole purpose of the standardization process. Values mapped to UNBR headings will later be split into valid headings (see 3.5 below). More detailed information on the content of these “fake” basic headings is provided in Appendix 4.

---

<sup>3</sup> A detailed description of these basic headings is available in an ICP document, *Classification of Final Expenditure on GDP*, Paris, April 2003.

**Table 2. ICP basic headings for household consumption – Codes and labels**  
(ICP\_SEQ is a variable we created for convenience)

ICP_SEQ	ICP Code				ALL
	11	00	00	0	
1	11	01	01	9	<i>UNBR Food and non-alcoholic beverages</i>
2	11	01	10	9	<i>UNBR Food</i>
3	11	01	11	9	<i>UNBR Bread and cereals</i>
4	11	01	11	1	Rice
5	11	01	11	2	Other cereals, flour and other products
6	11	01	11	3	Bread
7	11	01	11	4	Other bakery products
8	11	01	11	5	Pasta products
9	11	01	12	9	<i>UNBR Meat</i>
10	11	01	12	1	Beef and veal
11	11	01	12	2	Pork
12	11	01	12	3	Lamb, mutton and goat
13	11	01	12	4	Poultry
14	11	01	12	5	Other meats and meat preparations
15	11	01	13	9	<i>UNBR Fish and seafood</i>
16	11	01	13	1	Fresh, chilled or frozen fish and seafood
17	11	01	13	2	Preserved or processed fish and seafood
18	11	01	14	9	<i>UNBR Milk, cheese and eggs</i>
19	11	01	14	1	Fresh milk
20	11	01	14	2	Preserved milk and other milk products
21	11	01	14	3	Cheese
22	11	01	14	4	Eggs and egg-based products
23	11	01	15	9	<i>UNBR Oils and fats</i>
24	11	01	15	1	Butter and margarine
25	11	01	15	3	Other edible oil and fats
26	11	01	16	9	<i>UNBR Fruits</i>
27	11	01	16	1	Fresh or chilled fruits
28	11	01	16	2	Frozen, preserved or processed fruit and fruit-based products
29	11	01	17	9	<i>UNBR Vegetables</i>
30	11	01	17	1	Fresh or chilled vegetables other than potatoes
31	11	01	17	2	Fresh or chilled potatoes
32	11	01	17	3	Frozen, preserved or processed vegetables and vegetable-based products
33	11	01	18	9	<i>UNBR Sugar, jam, honey, chocolate and confectionery</i>
34	11	01	18	1	Sugar
35	11	01	18	2	Jams, marmalades and honey
36	11	01	18	3	Confectionery, chocolate and ice cream
37	11	01	19	1	Food products n.e.c.
38	11	01	20	9	<i>UNBR Non-alcoholic beverages</i>
39	11	01	21	1	Coffee, tea and cocoa
40	11	01	22	1	Mineral waters, soft drinks, fruit and vegetable juices
41	11	02	01	9	<i>UNBR Alcoholic beverages, tobacco and narcotics</i>
42	11	02	10	9	<i>UNBR Alcoholic beverages</i>
43	11	02	11	1	Spirits
44	11	02	12	1	Wine
45	11	02	13	1	Beer
46	11	02	21	1	Tobacco
47	11	02	31	1	Narcotics
48	11	03	01	9	<i>UNBR Clothing and footwear</i>

49	11	03	10	9	<i>UNBR Clothing</i>
50	11	03	11	1	Clothing material, other articles of clothing and clothing accessories
51	11	03	12	1	Garments
52	11	03	14	1	Cleaning, repair and hire of clothing
53	11	03	20	9	<i>UNBR Footwear</i>
54	11	03	21	1	Shoes and other footwear
55	11	03	22	1	Repair and hire of footwear
56	11	04	01	9	<i>UNBR Housing, water, electricity, gas and other fuels</i>
57	11	04	11	1	Actual and imputed rentals for housing
58	11	04	31	1	Maintenance and repair of the dwelling
59	11	04	40	9	<i>UNBR Water supply and miscellaneous services relating to the dwelling</i>
60	11	04	41	1	Water supply
61	11	04	42	1	Miscellaneous services relating to the dwelling
62	11	04	50	9	<i>UNBR Electricity, gas and other fuels</i>
63	11	04	51	1	Electricity
64	11	04	52	1	Gas
65	11	04	53	1	Other fuels
66	11	05	01	9	<i>UNBR Furnishing, household equipment and routine household maintenance</i>
67	11	05	10	9	<i>UNBR Furniture and furnishings, carpets and other floor coverings</i>
68	11	05	11	1	Furniture and furnishings
69	11	05	12	1	Carpets and other floor coverings
70	11	05	13	1	Repair of furniture, furnishings and floor coverings
71	11	05	21	1	Household textiles
72	11	05	30	9	<i>UNBR Household appliances</i>
73	11	05	31	1	Major household appliances whether electric or not
74	11	05	32	1	Small electric household appliances
75	11	05	33	1	Repair of household appliances
76	11	05	41	1	Glassware, tableware and household utensils
77	11	05	50	9	<i>UNBR Tools and equipment for house and garden</i>
78	11	05	51	1	Major tools and equipment
79	11	05	52	1	Small tools and miscellaneous accessories
80	11	05	60	9	<i>UNBR Goods and services for routine household maintenance</i>
81	11	05	61	1	Non-durable household goods
82	11	05	62	1	Domestic services
83	11	05	62	2	Household services
84	11	05	62	9	<i>UNBR Domestic services and household services</i>
85	11	06	01	9	<i>UNBR Health</i>
86	11	06	10	9	<i>UNBR Medical products, appliances and equipment</i>
87	11	06	11	1	Pharmaceuticals products
88	11	06	12	1	Other medical products
89	11	06	13	1	Therapeutic appliances and equipment
90	11	06	40	9	<i>UNBR Out-patient and hospital services</i>
91	11	06	20	9	<i>UNBR Out-patient services</i>
92	11	06	21	1	Medical services
93	11	06	22	1	Dental services
94	11	06	23	1	Paramedical services
95	11	06	31	1	Hospital services
96	11	07	01	9	<i>UNBR Transport</i>
97	11	07	10	9	<i>UNBR Purchase of vehicles</i>
98	11	07	11	1	Motor cars
99	11	07	12	1	Motor cycles
100	11	07	13	1	Bicycles
101	11	07	14	1	Animal drawn vehicles

102	11	07	20	9	<i>UNBR Operation of personal transport equipment</i>
103	11	07	22	1	Fuels and lubricants for personal transport equipment
104	11	07	23	1	Maintenance and repair of personal transport equipment
105	11	07	24	1	Other services in respect of personal transport equipment
106	11	07	30	9	<i>UNBR Transport services</i>
107	11	07	31	1	Passenger transport by railway
108	11	07	32	1	Passenger transport by road
109	11	07	33	1	Passenger transport by air
110	11	07	34	1	Passenger transport by sea and inland waterway
111	11	07	35	1	Combined passenger transport
112	11	07	36	1	Other purchase transport services
113	11	08	01	9	<i>UNBR Communication</i>
114	11	08	11	1	Postal services
115	11	08	21	1	Telephone and telefax equipment
116	11	08	31	1	Telephone and telefax services
117	11	09	01	9	<i>UNBR Recreation and culture</i>
118	11	09	10	9	<i>UNBR Audio-visual, photographic and information processing equipment</i>
119	11	09	11	1	Audio-visual, photographic and information processing equipment
120	11	09	14	1	Recording media
121	11	09	15	1	Repair of audio-visual, photographic and information process. equipment
122	11	09	20	9	<i>UNBR Other major durables for recreation and culture</i>
123	11	09	21	1	Major durables for outdoor and indoor recreation
124	11	09	23	1	Maintenance and repair of other major durables for recreation and culture
125	11	09	30	9	<i>UNBR Other recreational items and equipment, garden and pets</i>
126	11	09	31	1	Other recreational items and equipment
127	11	09	33	1	Garden and pets
128	11	09	35	1	Veterinary and other services for pets
129	11	09	40	9	<i>UNBR Recreational and cultural services</i>
130	11	09	41	1	Recreational and sporting services
131	11	09	42	1	Cultural services
132	11	09	43	1	Games of chance
133	11	09	51	1	Newspapers, books and stationery
134	11	09	61	1	Package holidays
135	11	10	11	1	Education
136	11	11	11	1	Catering services
137	11	11	21	1	Accommodation services
138	11	12	01	9	<i>UNBR Miscellaneous goods and services</i>
139	11	12	10	9	<i>UNBR Personal care</i>
140	11	12	11	1	Hairdressing salons and personal grooming establishments
141	11	12	12	1	Appliances, articles and products for personal care
142	11	12	21	1	Prostitution
143	11	12	30	9	<i>UNBR Personal effects n.e.c.</i>
144	11	12	31	1	Jewellery, clocks and watches
145	11	12	32	1	Other personal effects
146	11	12	41	1	Social protection
147	11	12	51	1	Insurance
148	11	12	60	9	<i>UNBR Financial services n.e.c.</i>
149	11	12	61	1	FISIM
150	11	12	62	1	Other financial services n.e.c.
151	11	12	71	1	Other services n.e.c.
152	11	13	11	1	Purchases by residential households in the rest of the world
153	11	13	11	2	Purchases by non-residential hhlds in the economic territory of the country

Mapping a value to an ICP basic heading consists or recoding the source code (the item code in the source dataset) into the corresponding ICP basic heading. Four situations can occur:

- **One to one relationship.** In some cases, there is a perfect match between an item in the survey questionnaire and a basic heading. The recoding is straightforward.
- **Many to one relationship.** In other cases, more than one item in the survey questionnaire will be mapped to a single basic heading. In such cases, mapping is also straightforward. This will usually be the case for food items, like fruits, vegetables, etc. In the Brazil dataset for example (where the diary method was used to collect data on daily consumption), 274 different items are mapped to basic heading “Fresh or chilled vegetables other than potatoes”. In Bangladesh, only 22 items are mapped to this basic heading (see Appendix 4).
- **One to many relationship.** An item in the questionnaire may correspond to more than one basic heading. In such cases, the item is mapped to a code corresponding to one of the “fake” basic headings created for that specific purpose. We already provided above the example of a questionnaire that would collect data on “Gas and electricity”. As the ICP has two disting basic headings for “Electricity” and “Gas”, the value would be mapped to the “fake” basic heading “Electricity, gas and other fuels”. The share of total consumption in such categories varies with the level of detail available in the survey dataset, but is generally relatively low (see “Share in UNBR items” in table 3 below).
- **No data available.** None of the 60 surveys processed so far collected data on all 107 basic headings. The extent of the data gaps varies a lot. Table 3 shows the number of basic headings (out of the 107 ones for which data could be obtained from sample surveys) for which no data is available. This number is calculated after we split the UNBR categories (see 3.5 below). Table 3 also shows the share of these non-covered basic headings in the national accounts data. These shares are relatively low in most cases (except in Bolivia, where “Dental services” and “Passenger transport by air” account for a large part of the missing consumption in the survey data), although not insignificant.

**Table 3. Mapping of source items with ICP basic heading – Summary for ten countries**

	Country									
	Latin America						Asia			
	ARG	BOL	BRA	COL	PER	PRY	BGD	IDN	LKA	THA
Number of goods and services found in source dataset										
All goods and services	92	147	6927	95	506	268	328	298	426	562
<i>Food and beverages</i>	29	63	3187	25	373	186	124	211	258	141
<i>Non food</i>	63	84	3740	70	133	82	204	87	168	421
Share of consumption mapped to UNBR codes	36.1	10.6	0.3	20.3	0	6.7	0.3	3.9	4.4	3.7
Basic headings with no data (after splitting) and their share in national accounts consumption data										
Number (out of 107)	20	28	5	30	23	23	29	15	11	6
Share in cons. (%)	3.79	7.75	0.07	5.47	4.35	4.14	1.20	0.53	2.60	4.26

### 3.5. Splitting unbroken basic headings

The calculation of poverty PPPs requires expenditure shares by “real” basic headings, since average prices are made available at that level (no price data is provided for our “fake” basic headings). Therefore, consumption data mapped to the 43 “fake” basic headings have to be redistributed to ICP basic headings. To do this, we use national accounts data provided by the countries<sup>4</sup> as follows:

1. We review the survey questionnaires to identify the basic headings among which the UNBR amount will have to be distributed.
2. We extract and normalize the share (in the national accounts data) of these basic headings.
3. We apply these shares to the UNBR amount to split it into the destination basic headings.

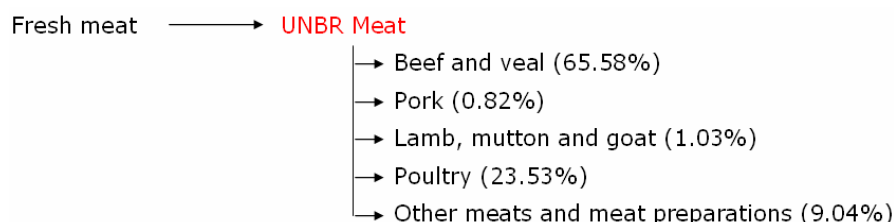
#### Examples:

Suppose that the national accounts data provide the shares for fresh meat consumption presented in column (1) below:

Basic heading	(1) % of total consumption	(2) % of consumption of fresh meat	(3) % of consumption of selected headings
Beef and veal	4.4736	65.5784	
Pork	0.0559	0.8197	
Lamb, mutton and goat	0.0703	1.0306	3.0671
Poultry	1.6051	23.5288	70.0221
Other meats and meat preparations	0.6169	9.0426	26.9109
<b>Total fresh meat</b>	<b>8.8217</b>	<b>100.0</b>	<b>100.0</b>

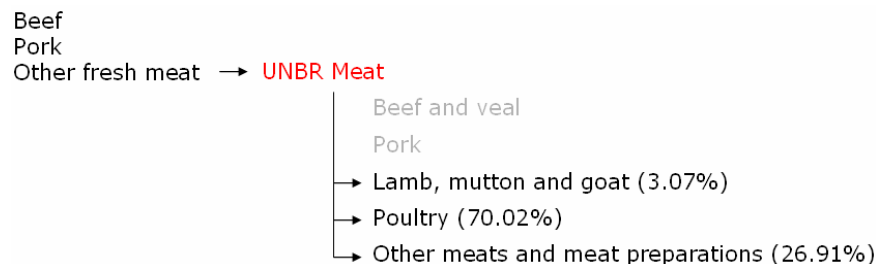
#### Case 1:

The survey collected data on consumption of “Fresh meat” without any more detail. We will then split the value into the 5 corresponding basic items. We normalize the shares so that their total is 100%, (column (2) in the table) and allocate accordingly.



#### Case 2:

The survey collected consumption data with some but not all necessary details. For example, the questionnaire includes “Beef”, “Pork”, and “Other fresh meat”. We normalize the shares of the basic headings excluding Beef and Pork (column (3) of the table) and allocate unbroken amount among the three basic items not listed in the questionnaire.



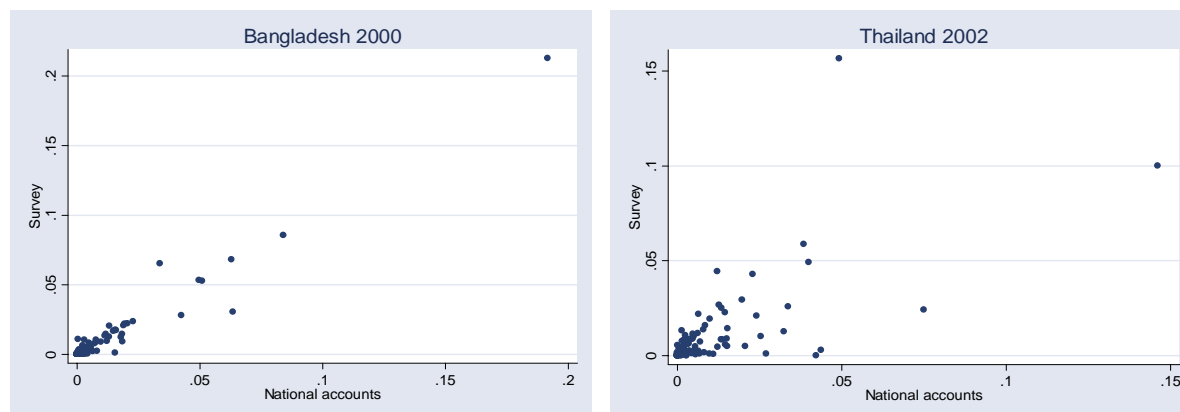
<sup>4</sup> And Stata code adapted from a program kindly provided by Angus Deaton.

This splitting method applies the same proportions to all households. It is highly probable that households at different income levels would have different shares, but such information is not available. In cases where the share of consumption in UNBR basic headings is low, the impact of this limitation is very marginal. It may be significant in the cases of Argentina and Colombia, where a large part of the total consumption (see Table 3) has to be redistributed into basic headings.

### 3.6. Running quality control tables

The Stata programs developed to generate the standardized data sets produce multiple tables for visual data quality and consistency control. These tables are available in the various corresponding log files. Among others, we produce tables showing:

- The extrapolated population (total, urban and rural) which can be compared with the expected population. This table aims to ensure that we use valid extrapolation coefficients, and that the dataset covers a sufficient proportion of the national population.
- The expenditure share by basic heading and the mean per capita consumption, by population quintile (quintiles based on per capita expenditure).
- Various household characteristics (ownership of durables, dwelling characteristics, access to water and electricity, and others), by population quintile.
- The list of basic headings for which data is available in national accounts but not in the survey. We then proceed to a visual check of the questionnaire, to ensure that this indeed results from a gap in the survey questionnaire and not from a recoding error in the mapping instructions.
- The correlation between national accounts and survey data. A correlation chart is also produced (see examples below). Table 4 presents the level of correlation obtained for 10 selected countries.



**Table 4. Correlation between national accounts and survey data on household consumption (total, food and beverages, and non food)**

	Country									
	Latin America						Asia			
	ARG	BOL	BRA	COL	PER	PRY	BGD	IDN	LKA	THA
Total	0.677	0.565	0.873	0.757	0.320	0.659	0.972	0.844	0.540	0.675
Food and beverages	0.942	0.884	0.869	0.330	0.729	0.884	0.971	0.921	0.783	0.677
Non food	0.657	0.520	0.870	0.791	0.275	0.621	0.975	0.924	0.336	0.675

## 4. Some lessons learned and recommendations

The observations below refer to the data from the 60-plus countries processed so far, not necessarily to the ten datasets used for the examples in this note.

- Ideally, we would process data from the most recent household budget survey, in order to derive consumption patterns for the same year as the year when ICP price data were collected (or a year as close as possible). But data are not easily accessible. For that reason, we did not always work on the most appropriate datasets.
- For some countries, we obtained multiple versions of a same survey dataset, sometimes significantly different from each other. Unfortunately, the source and specificities of these multiple versions is very difficult to track down. As a result, we cannot always guarantee that we were working on the most recent version of the dataset.
- The metadata (data documentation) is in most cases insufficient. In many cases accessing basic information such as the list of codes proved difficult. Sampling methods are often not documented at all, and the variables describing the various levels of stratification and primary sample units are very difficult to identify. Survey reports are sometimes difficult to locate and obtain.
- The consumption aggregation method applied by data producers (or by those who helped them process the data) is often a black box. In very few cases, this crucial phase of the data processing was fully documented and provided with the corresponding programs. Proper sharing of such information would not only reduce the cost and time of analyzing data, it would also ensure consistency of outputs generated by different analysts.
- Detecting and fixing outliers in consumption and expenditure data seems to have been made on a relatively ad-hoc basis by data producers, and is rarely documented. This is an important issue for inequality and poverty analysis. More research on this issue should be undertaken, and guidelines should be developed.
- The design of many questionnaires is unsatisfactory for our purpose (and probably for many other purposes that would require good data on household consumption patterns). Few data producers design their consumption/expenditure questionnaires using the COICOP classification. This would however be good practice, and should be promoted.
- Annualizing data on health is impossible in many cases, due to the one-month or two-weeks recall period.
- Many questionnaires include information on ownership of durables, but do not collect sufficient information (purchase/resell value; date of acquisition) to estimate an annual use value.
- The consistency between survey and national accounts consumption data varies a lot, and is very low in some cases. Some research work is needed to better understand and find solutions to this issue.
- Building this consumption database is a very time consuming exercise. We hope it can be expanded and serve many other purposes and users. This would however require that countries authorize its public dissemination. Countries should be encouraged to share their data. Technical support must be made available to those who are willing to do so but need technical assistance to formulate their dissemination policy and to properly document, and disseminate their data.

## Appendix 1 – List of surveys used in our examples

The following datasets have been used for illustrating this technical note:

Argentina (ARG)	Encuesta Nacional de Gastos de los Hogares	1996
Bolivia (BOL)	Encuesta Continua de Hogares	2002
Brazil (BRA)	Pesquisa de Orçamentos Familiares	2002
Colombia (COL)	Encuesta Nacional de Calidad de Vida	2003
Paraguay (PRY)	Encuesta Permanente de Hogares	2000
Peru (PER)	Enquesta Nacional de Hogares	2003
Bangladesh (BGD)	Household Budget Survey	2000
Indonesia (IDN)	SUSENAS	2002
Sri Lanka (LKA)	Sri Lanka Integrated Survey	2002
Thailand (THA)	Socioeconomic Survey	2002

## Appendix 2 – Data dictionary

### *Household-level file*

Files CCCYYYY\_HLD where CCC = ISO country code (string, 3 digit) and YYYY = survey year

<b>Name and type</b>	<b>Labels and codes</b>	<b>Instructions / notes</b>
<b>ihsn_no</b> <i>String</i>	<b>IHSN Survey ID</b>	<p>The International Household Survey Network (IHSN) number is the unique identification code of each survey listed in the IHSN central survey catalog. This variable is included to create a link to the IHSN database (source of metadata).</p> <p>The IHSN number is provided by the IHSN secretariat (contact: secretariat@ihsn.org)</p>
<b>country</b> <i>Numeric</i>	<b>Country code</b>	ISO 3166 3-digit numeric country code. The list of country codes is available at <a href="http://www.iso.org">www. ....</a>
<b>surveyr</b> <i>Numeric</i>	<b>Year of survey</b>	By convention, this is the year when data collection started (format YYYY).
<b>hid</b> <i>Numeric</i>	<b>Household ID</b>	Household unique identifier, computed as a sequential number with a range from 1 to N where N is the total number of households in the data file.
<b>ori_hid</b> <i>Numeric</i>	<b>Household ID in source dataset</b>	<p>Household identification number as available in the original dataset.</p> <p>In cases where more than one variable is used as an ID in the original dataset, a new variable is created by concatenating these variables. The purpose is to allow matching the new dataset with the original data files if needed.</p>
<b>stratum</b> <i>Numeric</i>	<b>Stratum</b>	<p>Code of the stratum, taken from the sample design information.</p> <p>If the sample is not stratified, the variable is created with all values as missing.</p>
<b>psu</b> <i>Numeric</i>	<b>Primary sampling unit</b>	<p>Codes for the primary sampling unit, taken from the sample design information.</p> <p>A unique code is created that identifies each PSU. In some datasets, the identification of the PSU may require more than one variable (e.g., the stratum variable + the PSU variable). In such case, a variable is created by combing these various elements.</p>
<b>geo_1</b> <i>Numeric</i>	<b>Sub-national code (level 1)</b>	Highest sub-national administrative level (country-specific) at which sample is representative (typically, this will correspond to a state or province). Country-specific labels are attached to this variable.
<b>geo_2</b> <i>Numeric</i>	<b>Sub-national code (level 2)</b>	<p>Second highest sub-national administrative level (country-specific) at which sample is representative (typically, this will correspond to a district within a state or province). Country-specific labels are attached to this variable</p> <p>For many datasets, geographic disaggregation is not possible beyond the geo_1 level. In such cases, geo_2 is created with all values as missing.</p> <p>A unique code is created that identifies each geo_2. In some datasets, the identification of this second administrative level requires more than one variable (e.g., the code of a province and district). In such case, the variable is created by combing these</p>

		various elements.																											
<b>rururb</b> <i>Numeric</i>	<b>Area of residence</b> <i>1 = Rural</i> <i>2 = Urban</i>	Urban/rural jurisdiction as defined by the country according to its own criteria. The 'semi-urban' category found in some datasets is assimilated to 'urban'.																											
<b>hhsiz</b> <i>Numeric</i>	<b>Household size</b>	Number of household members (based on country-specific definition of a household). Does not include paying boarders, domestic servants, and visitors.																											
<b>adeq_fao</b> <i>Numeric</i>	<b>Adults equivalent (FAO scale)</b>	<p>Number of adult equivalent in the household, computed based on the standard FAO scale. The variable is calculated for each household by summing up the following adult equivalent factor given to each member according to his/her age and sex:</p> <table border="1"> <thead> <tr> <th></th> <th><i>Male</i></th> <th><i>Female</i></th> </tr> </thead> <tbody> <tr> <td><i>&lt;1 yr</i></td> <td><i>0.27</i></td> <td><i>0.27</i></td> </tr> <tr> <td><i>1-3 yrs</i></td> <td><i>0.45</i></td> <td><i>0.45</i></td> </tr> <tr> <td><i>4-6 yrs</i></td> <td><i>0.61</i></td> <td><i>0.61</i></td> </tr> <tr> <td><i>7-9 yrs</i></td> <td><i>0.73</i></td> <td><i>0.73</i></td> </tr> <tr> <td><i>10-12 yrs</i></td> <td><i>0.86</i></td> <td><i>0.78</i></td> </tr> <tr> <td><i>13-15 yrs</i></td> <td><i>0.96</i></td> <td><i>0.83</i></td> </tr> <tr> <td><i>16-19 yrs</i></td> <td><i>1.02</i></td> <td><i>0.77</i></td> </tr> <tr> <td><i>20 and above</i></td> <td><i>1.00</i></td> <td><i>0.73</i></td> </tr> </tbody> </table>		<i>Male</i>	<i>Female</i>	<i>&lt;1 yr</i>	<i>0.27</i>	<i>0.27</i>	<i>1-3 yrs</i>	<i>0.45</i>	<i>0.45</i>	<i>4-6 yrs</i>	<i>0.61</i>	<i>0.61</i>	<i>7-9 yrs</i>	<i>0.73</i>	<i>0.73</i>	<i>10-12 yrs</i>	<i>0.86</i>	<i>0.78</i>	<i>13-15 yrs</i>	<i>0.96</i>	<i>0.83</i>	<i>16-19 yrs</i>	<i>1.02</i>	<i>0.77</i>	<i>20 and above</i>	<i>1.00</i>	<i>0.73</i>
	<i>Male</i>	<i>Female</i>																											
<i>&lt;1 yr</i>	<i>0.27</i>	<i>0.27</i>																											
<i>1-3 yrs</i>	<i>0.45</i>	<i>0.45</i>																											
<i>4-6 yrs</i>	<i>0.61</i>	<i>0.61</i>																											
<i>7-9 yrs</i>	<i>0.73</i>	<i>0.73</i>																											
<i>10-12 yrs</i>	<i>0.86</i>	<i>0.78</i>																											
<i>13-15 yrs</i>	<i>0.96</i>	<i>0.83</i>																											
<i>16-19 yrs</i>	<i>1.02</i>	<i>0.77</i>																											
<i>20 and above</i>	<i>1.00</i>	<i>0.73</i>																											
<b>m_00_15</b> <i>Numeric</i>	<b>Nb of males, 0 to 15 years</b>	Number of male household members aged 0 to 15 years. Undefined age are counted in "16 to 59"																											
<b>m_16_59</b> <i>Numeric</i>	<b>Nb of males, 16 to 59 years</b>	Number of male household members aged 16 to 59 years. Undefined age are counted in "16 to 59"																											
<b>m_60p</b> <i>Numeric</i>	<b>Nb of males, 60 years and over</b>	Number of male household members aged 60 years and over. Undefined age are counted in "16 to 59"																											
<b>f_00_15</b> <i>Numeric</i>	<b>Nb of females, 0 to 15 years</b>	Number of female household members aged 0 to 15 years. Undefined age are counted in "16 to 59"																											
<b>f_16_59</b> <i>Numeric</i>	<b>Nb of females, 16 to 59 years</b>	Number of female household members aged 16 to 59 years. Undefined age are counted in "16 to 59"																											
<b>f_60p</b> <i>Numeric</i>	<b>Nb of females, 60 years and over</b>	Number of female household members aged 60 years and over. Undefined age are counted in "16 to 59"																											
<b>hhcomp</b> <i>Numeric</i>	<b>Extended family</b> <i>0 = No, nuclear or unrelated</i> <i>1 = Yes</i>	An extended family is one with household members in addition to the head, his or her spouse and their children. Domestic servants and paying boarders are not considered part of the household so their presence in the dwelling has no impact on whether the household is extended or not.																											
<b>hhsex</b> <i>Numeric</i>	<b>Sex of household head</b> <i>1 = Male</i> <i>2 = Female</i>	Sex of the head of household. Each household, for the purposes of this data set, has one and only one head. The head of the household is the member declared as such by the respondent(s). In cases where more than one head is identified, the older one is considered as head.																											
<b>hhagey</b> <i>Numeric</i>	<b>Age of household head</b>	Age (in years) of the head of household. Each household, for the purposes of this data set, has one and only one head. The head of the household is the member declared as such by the respondent(s). In cases where more than one head is identified, the older one is considered as head.																											
<b>hhcivil</b>	<b>Civil status of head of</b>	Marital status of the head of household. Each household, for the																											

<i>Numeric</i>	<b>household</b> 1 = <i>Single</i> 2 = <i>Married, monogamous</i> 3 = <i>Married, polygamous</i> 4 = <i>Civil union</i> 5 = <i>Divorced/Separated</i> 6 = <i>Widowed</i>	purposes of this data set, has one and only one head. The head of the household is the member declared as such by the respondent(s). In cases where more than one head is identified, the older one is considered as head.
<b>hheduc</b> <i>Numeric</i>	<b>Education level of household head</b> 1 = <i>None</i> 2 = <i>Preschool</i> 3 = <i>Primary incomplete</i> 4 = <i>Primary completed</i> 5 = <i>Secondary incomplete</i> 6 = <i>Secondary complete</i> 7 = <i>Post secondary technical</i> 8 = <i>University or graduate studies</i> 9 = <i>Adult education or literacy program</i> 10 = <i>Not stated</i> 99 = <i>Undefined</i>	<p>Education level of the head of household. Each household, for the purposes of this data set, has one and only one head. The head of the household is the member declared as such by the respondent(s). In cases where more than one head is identified, the older one is considered as head.</p> <p>This variable is obtained by recoding country-specific information. The best possible match is sought, but in many cases the correspondence between country-specific values and the standard codes is imperfect. To assess the reliability of this variable, users are invited to consult the survey questionnaire and the standardization program.</p>
<b>ownhouse</b> <i>Numeric</i>	<b>Ownership of dwelling unit</b> 0 = <i>No</i> 1 = <i>Yes</i>	<p>Ownership status of the dwelling unit by the household residing in it.</p> <ul style="list-style-type: none"> <li>▪ <i>Yes</i> includes ownership whether or not full-payment has yet been made.</li> <li>▪ <i>No</i> includes renters, squatters, housing provided for free.</li> </ul>
<b>roof</b> <i>Numeric</i>	<b>Main material used for roof</b> 1 = <i>Adobe, wattle, mud,</i> 2 = <i>Thatch</i> 3 = <i>Wood</i> 4 = <i>Iron / metal sheets</i> 5 = <i>Cement</i> 6 = <i>Tiles/bricks</i> 9 = <i>Other</i>	<ul style="list-style-type: none"> <li>▪ <i>Adobe, wattle, mud</i> includes all building techniques that rely on earth or mud put over a frame or mixed with other materials for strength.</li> <li>▪ <i>Thatch</i> includes grass or any form of natural vegetation for roofing.</li> <li>▪ <i>Iron or metal sheets</i> are processed tin, zinc, and the like</li> <li>▪ <i>Cement</i> includes concrete and stone and cement blocks.</li> <li>▪ <i>Tiles/bricks</i> include baked bricks.</li> <li>▪ <i>Other</i> includes tin, cardboard among others.</li> </ul> <p>The best possible match is sought, but in many cases the correspondence between country-specific values and these standardized codes is imperfect. To assess the reliability of this variable, users are invited to consult the survey questionnaire and the standardization program.</p>
<b>walls</b> <i>Numeric</i>	<b>Main material used for external walls</b> 1 = <i>Adobe, wattle, mud</i> 2 = <i>Bricks</i> 3 = <i>Wood</i> 4 = <i>Iron sheets</i> 5 = <i>Cement</i> 9 = <i>Other</i>	<ul style="list-style-type: none"> <li>▪ <i>Adobe, wattle, mud</i> includes all building techniques that rely on earth or mud put over a frame or mixed with other materials for strength.</li> <li>▪ <i>Bricks</i> include baked bricks.</li> <li>▪ <i>Wood</i> includes timber and wood planks, unfinished.</li> <li>▪ <i>Iron / metal sheets</i> are processed tin, zinc and the like.</li> <li>▪ <i>Cement</i> includes concrete and stone and cement block.</li> <li>▪ <i>Other</i> includes tin, cardboard among others.</li> </ul> <p>The best possible match is sought, but in many cases the correspondence between country-specific values and these standardized codes is imperfect. To assess the reliability of this variable, users are invited to consult the survey questionnaire and the standardization program.</p>
<b>floor</b> <i>Numeric</i>	<b>Main material used for floor</b> 1 = <i>Earth</i> 2 = <i>Bricks</i> 3 = <i>Wood planks</i> 4 = <i>Polished wood/tiles</i> 5 = <i>Cement</i>	<ul style="list-style-type: none"> <li>▪ <i>Earth</i> implies dirt or mud floors.</li> <li>▪ <i>Bricks</i> include baked bricks.</li> <li>▪ <i>Polished wood/tiles</i> include finished wood floors, parquet floors, as well as ceramic tiles.</li> <li>▪ <i>Other</i> includes linoleum or vinyl flooring</li> </ul>

	9 = <i>Other</i>	The best possible match is sought, but in many cases the correspondence between country-specific values and these standardized codes is imperfect. To assess the reliability of this variable, users are invited to consult the survey questionnaire and the standardization program.
<b>rooms</b> <i>Numeric</i>	<b>Number of habitable rooms</b>	Number of rooms in the whole household dwelling unit, which may consist of one or more structure(s). It includes all rooms used for living, sleeping and eating. It excludes storerooms, bathrooms and kitchens. In the case of a one-room dwelling this variable will have the value of 1.
<b>water</b> <i>Numeric</i>	<b>Main source of drinking water</b> 1 = <i>Piped (own tap)</i> 2 = <i>Public standpipe</i> 3 = <i>Protected well</i> 4 = <i>Unprotected well</i> 5 = <i>Surface water</i> 6 = <i>Rain water</i> 7 = <i>Truck, vendor</i> 9 = <i>Other</i>	<ul style="list-style-type: none"> <li>▪ <i>Piped</i> refers to water delivered via a pipe within the house or compound, that is, own tap. It includes both interior pipe and exterior one.</li> <li>▪ <i>Public standpipe</i> refers to water delivered via pipe but not within compound (water point shared among houses). This refers to public tap and community water points.</li> <li>▪ <i>Wells</i> include springs, boreholes.</li> <li>▪ <i>Surface water</i> includes lakes, rivers and ponds.</li> <li>▪ <i>Other</i> includes water sources not mentioned elsewhere.</li> </ul> <p>If the main source of water differs between the wet and dry season, refers to the water source during dry season.</p> <p>The best possible match is sought, but in many cases the correspondence between country-specific values and these standardized codes is imperfect. To assess the reliability of this variable, users are invited to consult the survey questionnaire and the standardization program.</p>
<b>electcon</b> <i>Numeric</i>	<b>Connection to electricity in dwelling</b> 1 = <i>Yes, public/quasi public</i> 2 = <i>Yes, private</i> 3 = <i>Yes, source unstated</i> 4 = <i>No</i>	<ul style="list-style-type: none"> <li>▪ <i>Public or quasi public</i> refers to electricity from mains.</li> <li>▪ <i>Private</i> refers to electricity from generator or solar or private company.</li> </ul> <p>Note that having an electrical connection says nothing about the actual electrical service received by the household in a given country or area.</p>
<b>fuelcook</b> <i>Numeric</i>	<b>Main cooking fuel</b> 1 = <i>Firewood</i> 2 = <i>Kerosene</i> 3 = <i>Charcoal</i> 4 = <i>Electricity</i> 5 = <i>Gas</i> 9 = <i>Other</i>	<ul style="list-style-type: none"> <li>▪ <i>Electricity</i> refers to mains and electricity from generator or solar.</li> <li>▪ <i>Other</i> includes fuel derived from coffee waste, saw dust, crop residue, cow dung among others.</li> </ul>
<b>fuelligh</b> <i>Numeric</i>	<b>Main source of lighting</b> 1 = <i>Electricity</i> 2 = <i>Kerosene</i> 3 = <i>Candles</i> 9 = <i>Other</i>	<ul style="list-style-type: none"> <li>▪ <i>Electricity</i> refers to any source of electricity, mains, generator, solar, etc.</li> <li>▪ <i>Other</i> includes fuel derived from coffee waste, saw dust, crop residue, cow dung among others.</li> </ul>
<b>toilet</b> <i>Numeric</i>	<b>Main toilet facility</b> 1 = <i>Flush toilet</i> 2 = <i>Latrine</i> 3 = <i>No facility</i> 9 = <i>Other</i>	<ul style="list-style-type: none"> <li>▪ <i>Flush toilet</i> refers to flush to main sewer or septic tank.</li> <li>▪ <i>Latrine</i> is a simple but protected pit latrine. It can be covered or ventilated. It excludes open pit or uncovered latrines.</li> <li>▪ <i>No facility</i> includes, open fields, bush.</li> <li>▪ <i>Other</i> includes bucket, pan, and open pit latrines among others.</li> </ul>
<b>ownland</b> <i>Numeric</i>	<b>Ownership of land</b> 0 = <i>No</i> 1 = <i>Yes</i>	Land covers all land owned by the household, be it residential, agricultural, rented out, fallow or in use. Some countries ask about land ownership irrespective whether it is agricultural or non-agricultural. Refers to the actual property rights of the land: <ol style="list-style-type: none"> <li>1. Can have a legal document such as title deed showing proof of ownership.</li> </ol>

		<ol style="list-style-type: none"> <li>2. Does not have a legal document but have land ownership rights as per the definition of traditional land ownership system.</li> <li>3. Has some other document showing ownership (bill of sale, receipt) although this is not formally a title.</li> </ol>
<b>landsize</b> <i>Numeric</i>	<b>Land size owned (ha)</b>	Includes both residential and agricultural land. Land size should be in hectares. By convention 1 ha = 2.471 acres.
<b>llivesk</b> <i>Numeric</i>	<b>Nb of large-sized livestock owned</b>	Number of large-size livestock heads owned by the household (includes cattle, camels, donkeys and horses).
<b>mlivesk</b> <i>Numeric</i>	<b>Nb of medium-sized livestock owned</b>	Number of medium-size livestock heads owned by the household (includes sheeps, goats and pigs).
<b>poultry</b> <i>Numeric</i>	<b>Nb of poultry owned</b>	Number of birds owned by the household (includes all forms of birds such as chicken, geese, and doves).
<b>radio</b> <i>Numeric</i>	<b>Ownership of a radio</b> <i>0 = No</i> <i>1 = Yes</i>	Ownership of a radio, irrespective of who owns it within the household and regardless of what condition the asset is in. It includes radio, radio cassette, and 3-in-1 radio cassette.
<b>tv</b> <i>Numeric</i>	<b>Ownership of a television</b> <i>0 = No</i> <i>1 = Yes</i>	Ownership of a television, irrespective of who owns it within the household and regardless of what condition the asset is in.
<b>phone</b> <i>Numeric</i>	<b>Ownership of a telephone</b> <i>0 = No</i> <i>1 = Yes</i>	Ownership of a phone and/or a cell phone, irrespective of who owns it within the household and regardless of what condition the asset is in.
<b>cphone</b> <i>Numeric</i>	<b>Ownership of a cell phone</b> <i>0 = No</i> <i>1 = Yes</i>	Ownership of a cell phone, irrespective of who owns it within the household and regardless of what condition the asset is in.
<b>rfridge</b> <i>Numeric</i>	<b>Ownership of a refrigerator</b> <i>0 = No</i> <i>1 = Yes</i>	Ownership of a refrigerator in house. Refers to actual ownership of the asset irrespective of who owns it within the household and regardless of what condition the asset is in.
<b>sewmach</b> <i>Numeric</i>	<b>Ownership of a sewing machine</b> <i>0 = No</i> <i>1 = Yes</i>	Ownership of a sewing machine, irrespective of who owns it within the household and regardless of what condition the asset is in.
<b>computer</b> <i>Numeric</i>	<b>Ownership of computer</b> <i>0 = No</i> <i>1 = Yes</i>	Ownership of a computer, irrespective of who owns it within the household and regardless of what condition the asset is in. This refers to computer for household use (not for commercial use).
<b>stove</b> <i>Numeric</i>	<b>Ownership of a stove</b> <i>0 = No</i> <i>1 = Yes</i>	Ownership of a stove or cooker in house, irrespective of who owns it within the household and regardless of what condition the asset is in.
<b>bcycle</b> <i>Numeric</i>	<b>Ownership of a bicycle</b> <i>0 = No</i> <i>1 = Yes</i>	Ownership of a bicycle, irrespective of who owns it within the household and regardless of what condition the asset is in.
<b>mcycle</b> <i>Numeric</i>	<b>Ownership of a motorcycle</b> <i>0 = No</i> <i>1 = Yes</i>	Ownership of a motorcycle, irrespective of who owns it within the household and regardless of what condition the asset is in.
<b>car</b> <i>Numeric</i>	<b>Ownership of a private car</b> <i>0 = No</i> <i>1 = Yes</i>	Ownership of a car or truck. This refers to car for household use. Not a commercial vehicle. Refers to actual ownership of the asset irrespective of who owns it within the household and regardless of what condition the asset is in.
<b>oxcart</b>	<b>Ownership of an animal cart</b> <i>0 = No</i>	Ownership of an animal cart, which is used as a means of transport or a farm tool. Refers to actual ownership of the asset

<i>Numeric</i>	<i>1 = Yes</i>	irrespective of who owns it within the household and regardless of what condition the asset is in.
<b>boat</b> <i>Numeric</i>	<b>Ownership of a boat</b> <i>0 = No</i> <i>1 = Yes</i>	Ownership of a boat or canoe. Refers to actual ownership of the asset irrespective of who owns it within the household and regardless of what condition the asset is in.
<b>focon_nd</b> <i>Numeric</i>	<b>Food consumption (non deflated)</b>	Total annual household food consumption (including non-alcoholic beverages) in local currency, not deflated by regional price deflator. Includes purchased, home produced and received items.  This variable is obtained by aggregating data from the commodity-level standardized data file (cccyyyy_ppp.dta). Its exact content varies from country to country, due to differences in the questionnaire design and survey methodology. Users of the data are invited to consult the survey questionnaire and the “standardization” programs to obtain details on the content of the variable.
<b>focon_de</b> <i>Numeric</i>	<b>Food consumption (deflated)</b>	Total annual household food consumption (including non-alcoholic beverages) in local currency, deflated by regional price deflator. Includes purchased, home produced and received items.  The regional price deflators (Laspeyres or Paachse) used in/by the national data producer is used when available.  This variable is obtained by aggregating data from the commodity-level standardized data file (cccyyyy_ppp.dta). Its exact content varies from country to country, due to differences in the questionnaire design and survey methodology. Users of the data are invited to consult the survey questionnaire and the “standardization” programs to obtain details on the content of the variable.
<b>nfcon_nd</b> <i>Numeric</i>	<b>Non-food consumption (non deflated)</b>	Total annual household food consumption in local currency, not deflated by regional price deflator. Includes purchased, home produced and received items.  This variable is obtained by aggregating data from the commodity-level standardized data file (cccyyyy_ppp.dta). Its exact content varies from country to country, due to differences in the questionnaire design and survey methodology. Users of the data are invited to consult the survey questionnaire and the “standardization” programs to obtain details on the content of the variable.
<b>nfcon_de</b> <i>Numeric</i>	<b>Non-food consumption (deflated)</b>	Total annual household food consumption in local currency, deflated by regional price deflator. Includes purchased, home produced and received items.  The regional price deflators (Laspeyres or Paachse) used in/by the national data producer is used when available.  This variable is obtained by aggregating data from the commodity-level standardized data file (cccyyyy_ppp.dta). Its exact content varies from country to country, due to differences in the questionnaire design and survey methodology. Users of the data are invited to consult the survey questionnaire and the “standardization” programs to obtain details on the content of the variable.
<b>tocon_nd</b> <i>Numeric</i>	<b>Total household consumption (non deflated)</b>	Total annual household consumption in local currency, not deflated by regional price deflator. Includes purchased, home produced and received items.  This variable is obtained by aggregating data from the commodity-level standardized data file (cccyyyy_ppp.dta).Its

		exact content varies from country to country, due to differences in the questionnaire design and survey methodology (e.g., it may include or not annualized consumption of durables, etc.) Users of the data are invited to consult the survey questionnaire and the “standardization” programs to obtain details on the content of the variable.
<b>tocon_de</b> <i>Numeric</i>	<b>Total consumption (deflated)</b>	<p>Total annual household consumption in local currency, deflated by regional price deflator. Includes purchased, home produced and received items.</p> <p>The regional price deflators (Laspeyres or Paachse) used in/by the national data producer is used when available.</p> <p>This variable is obtained by aggregating data from the commodity-level standardized data file (cccyyyy_ppp.dta). Its exact content varies from country to country, due to differences in the questionnaire design and survey methodology (e.g., it may include or not annualized consumption of durables, etc.) Users of the data are invited to consult the survey questionnaire and the standardization programs to obtain details on the content of the variable.</p>
<b>wta_hh</b> <i>Numeric</i>	<b>Household weighting coefficient</b>	Weighting coefficient to be used in all calculations referring to household level data.
<b>wta_pop</b> <i>Numeric</i>	<b>Population weighting coefficient</b>	<p>Weighting coefficient to be used to obtain estimates referring to the population, in all calculations based on household level data.</p> <p>Calculated as <math>wta\_pop = wta\_hh * hhsiz</math></p> <p>The sum of wta_pop across all households must provide a realistic estimate of the population of the country.</p>

## Commodity-level file – by original commodity code

Files CCCYYYY\_ORI where CCC = ISO country code (string, 3 digit) and YYYY = survey year

Name	Labels and codes	Instructions / notes
<b>ihsn_no</b>	<b>IHSN Survey ID</b>	<i>See household-level file</i>
<b>country</b>	<b>Country code</b>	<i>See household-level file</i>
<b>surveyr</b>	<b>Year of survey</b>	<i>See household-level file</i>
<b>hid</b>	<b>Household ID</b>	<i>See household-level file</i>
<b>stratum</b>	<b>Stratum</b>	<i>See household-level file</i>
<b>psu</b>	<b>Primary sampling unit</b>	<i>See household-level file</i>
<b>geo_1</b>	<b>Sub-national code (level 1)</b>	<i>See household-level file</i>
<b>geo_2</b>	<b>Sub-national code (level 2)</b>	<i>See household-level file</i>
<b>rururb</b>	<b>Area of residence</b>	<i>See household-level file</i>
<b>hhsiz</b>	<b>Household size</b>	<i>See household-level file</i>
<b>adeq_fao</b>	<b>Adults equivalent (FAO scale)</b>	<i>See household-level file</i>
<b>src_var</b> <i>String</i>	<b>Source variable</b>	String variable providing the name of the source variable(s) in the original file.
<b>src_cod</b> <i>Numeric</i>	<b>Source code</b>	This variable provides a unique code for each commodity for which data are available in the source dataset. In many cases, the original item code is used. In some cases, when the original codes do not uniquely identify a good or service, a new value is created. All values are labeled. This variable allows the production of a table showing all consumption items available in the source dataset, and their mapping to ICP basic headings.
<b>icp_seq</b> <i>Numeric</i>	<b>ICP code sequential number</b>	This variable is used to facilitate the correspondence (mapping) between the original codes and the ICP codes. It is used for convenience of programmers and analysts. The list of ICP_SEQ codes and the corresponding ICP codes are available in table 1 above. Although the ICP lists 110 basic headings for household consumption, ICP_SEQ ranges from 1 to 153. The difference comes from the 43 “Unbroken” basic headings created to allow the mapping of items listed in the questionnaires that correspond to more than one ICP basic heading.
<b>category</b> <i>Numeric</i>	<b>ICP category</b>	ICP category code.
<b>group</b> <i>Numeric</i>	<b>ICP group</b>	ICP group code. The variable includes some non-original ICP codes, corresponding to “unbroken” groups of ICP groups (label starting with “_UNBR”).
<b>class</b> <i>Numeric</i>	<b>ICP class</b>	ICP class code. The variable includes some non-original ICP codes, corresponding to “unbroken” groups of ICP classes (label starting with “_UNBR”).
<b>basic_hd</b> <i>Numeric</i>	<b>ICP basic heading</b>	ICP basic heading code. The variable includes some non-original ICP codes, corresponding to “unbroken” groups of ICP basic headings (label starting with “_UNBR”).

<b>cons_csh</b> <i>Numeric</i>	<b>Annual consumption (purchased)</b>	<p>Annual consumption of purchased item, in local currency (not deflated by regional price deflators).</p> <p>The exact content of this variable varies from country to country, due to differences in the questionnaire design and survey methodology. Users of the data are invited to consult the survey questionnaire and the “standardization” programs to obtain details on the content of the variable.</p> <p>If the source dataset does not distinguish purchased, home produced and received items, all amounts are stored as “purchased”.</p>
<b>cons_hmp</b> <i>Numeric</i>	<b>Annual consumption (home produced)</b>	<p>Estimated value of annual consumption of home-produced item in local currency (not deflated by regional price deflators).</p> <p>The exact content of this variable varies from country to country, due to differences in the questionnaire design and survey methodology. Users of the data are invited to consult the survey questionnaire and the “standardization” programs to obtain details on the content of the variable.</p> <p>If the source dataset does not distinguish purchased, home produced and received items, all amounts are stored as “purchased”.</p>
<b>cons_gft</b> <i>Numeric</i>	<b>Annual consumption (received)</b>	<p>Estimated value of annual consumption of received item in local currency (not deflated by regional price deflators).</p> <p>The exact content of this variable varies from country to country, due to differences in the questionnaire design and survey methodology. Users of the data are invited to consult the survey questionnaire and the “standardization” programs to obtain details on the content of the variable.</p> <p>If the source dataset does not distinguish purchased, home produced and received items, all amounts are stored as “purchased”.</p>
<b>cons_tot</b> <i>Numeric</i>	<b>Annual consumption (all sources)</b>	<p>Estimated value of annual consumption (all sources) in local currency (not deflated by regional price deflators).</p> <p>The exact content of this variable varies from country to country, due to differences in the questionnaire design and survey methodology. Users of the data are invited to consult the survey questionnaire and the “standardization” programs to obtain details on the content of the variable.</p> <p>The variable must be equal to <math>\text{cons\_csh} + \text{cons\_hmp} + \text{cons\_gft}</math>.</p>
<b>reg_defl</b> <i>Numeric</i>	<b>Regional price deflator</b>	<p>When available, this variable is a Laspeyres or a Paasche regional price deflator (depending on what price deflator was readily available in the source dataset). If no price deflator is readily available, a Paasche deflator is calculated when possible. If no regional price deflator is available and if it cannot be calculated, the value “1” is imputed.</p>
<b>wta_hh</b>	<b>Household weighting coefficient</b>	<i>See household-level file</i>
<b>wta_pop</b>	<b>Population weighting coefficient</b>	<i>See household-level file</i>

### ***Commodity-level file – aggregated by BHD***

Files CCCYYYY\_PPP where CCC = ISO country code (string, 3 digit) and YYYY = survey year

Obtained by collapsing at HH/BH level and subset of variables.

<b>Name</b>	<b>Labels</b>	<b>Instructions / notes</b>
<b>hid</b>	<b>Household ID</b>	<i>See household-level file</i>
<b>stratum</b>	<b>Stratum</b>	<i>See household-level file</i>
<b>psu</b>	<b>Primary sampling unit</b>	<i>See household-level file</i>
<b>rururb</b>	<b>Area of residence</b>	<i>See household-level file</i>
<b>hhsiz</b>	<b>Household size</b>	<i>See household-level file</i>
<b>icp_seq</b>	<b>ICP code sequential number</b>	<i>See household-level file</i>
<b>basic_hd</b>	<b>ICP basic heading</b>	<i>See household-level file</i>
<b>cons_tot</b>	<b>Annual consumption (all sources)</b>	<i>See household-level file</i>
<b>wta_hh</b>	<b>Household weighting coefficient</b>	<i>See household-level file</i>

### Appendix 3 - Summary table for BANGLADESH 2000

Sample size: 7448

No	BH code	Basic heading label	% households		Source items	Share in tot. cons. (%)	
			original	mapped		original	mapped
1	1101019	UNBR Food and non-alcoholic beverages	0.00	0.00			
2	1101109	UNBR Food	0.00	0.00			
3	1101119	UNBR Bread and cereals	0.00	0.00			
4	1101111	Rice	✓ 99.85	99.85	6	21.28	21.28
5	1101112	Other cereals, flour and other products	✓ 18.37	18.37	2	0.21	0.21
6	1101113	Bread	✓ 34.10	34.10	1	0.32	0.32
7	1101114	Other bakery products	✓ 17.75	17.75	2	0.16	0.16
8	1101115	Pasta products	✓ 16.18	16.18	1	0.10	0.10
9	1101129	UNBR Meat	⊕ 4.65	0.00	1	0.09	
10	1101121	Beef and veal	✓ 40.56	43.03	2	2.15	2.20
11	1101122	Pork	✖ 0.00	4.65			0.00
12	1101123	Lamb, mutton and goat	✓ 3.57	7.89	2	0.20	0.20
13	1101124	Poultry	✓ 26.28	29.28	2	1.30	1.33
14	1101125	Other meats and meat preparations	✓ 21.40	24.53	4	0.32	0.33
15	1101139	UNBR Fish and seafood	0.00	0.00			
16	1101131	Fresh, chilled or frozen fish and seafood	✓ 97.38	97.38	14	6.80	6.80
17	1101132	Preserved or processed fish and seafood	✓ 56.07	56.07	2	1.03	1.03
18	1101149	UNBR Milk, cheese and eggs	0.30	0.00	1	0.01	
19	1101141	Fresh milk	✓ 47.38	47.38	2	1.68	1.68
20	1101142	Preserved milk and other milk products	✓ 10.50	10.67	2	0.30	0.30
21	1101143	Cheese	✖ 0.00	0.30			0.00
22	1101144	Eggs and egg-based products	✓ 66.94	66.94	3	0.85	0.85
23	1101159	UNBR Oils and fats	0.00	0.00			
24	1101151	Butter and margarine	✓ 0.12	0.12	1	0.00	0.00
25	1101153	Other edible oil and fats	✓ 99.80	99.80	5	2.17	2.17
26	1101169	UNBR Fruits	0.00	0.00			
27	1101161	Fresh or chilled fruits	✓ 65.00	65.00	17	1.74	1.74
28	1101162	Frozen, preserved or processed fruit and fruit-based products	✖ 0.00	0.00			
29	1101179	UNBR Vegetables	0.00	0.00			
30	1101171	Fresh or chilled vegetables other than potatoes	✓ 99.84	99.84	22	6.52	6.52
31	1101172	Fresh or chilled potatoes	✓ 97.64	97.64	1	1.67	1.67
32	1101173	Frozen, preserved or processed vegetables and vegetable-based products	✓ 9.37	9.37	1	0.03	0.03
33	1101189	UNBR Sugar, jam, honey, chocolate and confectionery	⊕ 1.92	0.00	1	0.01	
34	1101181	Sugar	✓ 59.04	59.60	2	0.79	0.80
35	1101182	Jams, marmalades and honey	✓ 0.35	2.24	1	0.00	0.00
36	1101183	Confectionery, chocolate and ice cream	✓ 22.15	23.12	3	0.10	0.10
37	1101191	Food products n.e.c.	✓ 99.81	99.81	13	3.03	3.03
38	1101209	UNBR Non-alcoholic beverages	0.00	0.00			

No	BH code	Basic heading label	% households		Source items	Share in tot. cons. (%)		
			original	mapped		original	mapped	
39	1101211	Coffee, tea and cocoa	✓	20.01	20.01	1	0.18	0.18
40	1101221	Mineral waters, soft drinks, fruit and vegetable juices	✓	9.98	9.98	5	0.11	0.11
41	1102019	UNBR Alcoholic beverages, tobacco and narcotics		0.00	0.00			
42	1102109	UNBR Alcoholic beverages		0.00	0.00			
43	1102111	Spirits	✘	0.00	0.00			
44	1102121	Wine	✘	0.00	0.00			
45	1102131	Beer	✘	0.00	0.00			
46	1102211	Tobacco	✓	58.97	58.97	4	1.71	1.71
47	1102311	Narcotics	✓	54.36	54.36	6	1.25	1.25
48	1103019	UNBR Clothing and footwear		0.00	0.00			
49	1103109	UNBR Clothing		0.00	0.00			
50	1103111	Clothing material, other articles of clothing and clothing accessories	✓	28.44	28.44	8	0.28	0.28
51	1103121	Garments	✓	99.92	99.92	24	5.31	5.31
52	1103141	Cleaning, repair and hire of clothing	✓	57.52	57.52	2	0.23	0.23
53	1103209	UNBR Footwear	⊙	2.83	0.00	1	0.00	
54	1103211	Shoes and other footwear	✓	97.52	97.52	8	0.75	0.76
55	1103221	Repair and hire of footwear	✓	9.88	11.59	1	0.01	0.01
56	1104019	UNBR Housing, water, electricity, gas and other fuels		0.00	0.00			
57	1104111	Actual and imputed rentals for housing	✓	98.54	98.54	2	8.55	8.55
58	1104311	Maintenance and repair of the dwelling	✓	21.97	21.97	5	1.22	1.22
59	1104409	UNBR Water supply and miscellaneous services relating to the dwelling		7.88	0.00	1	0.15	
60	1104411	Water supply	✓	0.00	7.88			0.11
61	1104421	Miscellaneous services relating to the dwelling	✘	0.00	7.88			0.03
62	1104509	UNBR Electricity, gas and other fuels		0.00	0.00			
63	1104511	Electricity	✓	36.72	36.72	1	1.44	1.44
64	1104521	Gas	✓	9.99	9.99	1	0.47	0.47
65	1104531	Other fuels	✓	92.02	92.02	6	5.26	5.26
66	1105019	UNBR Furnishing, household equipment and routine household maintenance		0.00	0.00			
67	1105109	UNBR Furniture and furnishings, carpets and other floor coverings		0.00	0.00			
68	1105111	Furniture and furnishings	✓	39.19	39.19	6	1.01	1.01
69	1105121	Carpets and other floor coverings	✘	0.00	0.00			
70	1105131	Repair of furniture, furnishings and floor coverings	✘	1.56	1.56	1	0.03	0.03
71	1105211	Household textiles	✓	48.67	48.67	11	0.62	0.62
72	1105309	UNBR Household appliances		0.00	0.00			
73	1105311	Major household appliances whether electric or not	✓	1.49	1.49	3	0.09	0.09
74	1105321	Small electric household appliances	✓	2.98	2.98	1	0.08	0.08
75	1105331	Repair of household appliances	✘	0.00	0.00			
76	1105411	Glassware, tableware and household utensils	✓	58.90	58.90	5	0.30	0.30
77	1105509	UNBR Tools and equipment for house and garden		0.00	0.00			
78	1105511	Major tools and equipment	✘	0.00	0.00			
79	1105521	Small tools and miscellaneous accessories	✘	0.32	0.32	1	0.01	0.01

No	BH code	Basic heading label	% households		Source items	Share in tot. cons. (%)		
			original	mapped		original	mapped	
80	1105609	UNBR Goods and services for routine household maintenance		0.00	0.00			
81	1105611	Non-durable household goods	✓	99.33	99.33	9	0.92	0.92
82	1105621	Domestic services	✓	7.17	7.17	3	0.87	0.87
83	1105622	Household services	✘	0.00	0.00			
84	1105629	UNBR Domestic services and household services		0.00	0.00			
85	1106019	UNBR Health		0.00	0.00			
86	1106109	UNBR Medical products, appliances and equipment		0.00	0.00			
87	1106111	Pharmaceuticals products	✓	88.35	88.35	2	1.45	1.45
88	1106121	Other medical products	✘	0.00	0.00			
89	1106131	Therapeutic appliances and equipment	✓	1.22	1.22	6	0.01	0.01
90	1106409	UNBR Out-patient and hospital services		0.00	0.00			
91	1106209	UNBR Out-patient services		0.00	0.00			
92	1106211	Medical services	✓	48.31	48.31	7	0.46	0.46
93	1106221	Dental services	✓	1.58	1.58	2	0.01	0.01
94	1106231	Paramedical services	✘	5.24	5.24	2	0.02	0.02
95	1106311	Hospital services	✓	2.15	2.15	2	0.11	0.11
96	1107019	UNBR Transport		0.00	0.00			
97	1107109	UNBR Purchase of vehicles		0.00	0.00			
98	1107111	Motor cars	✘	0.00	0.00			
99	1107121	Motor cycles	✘	0.00	0.00			
100	1107131	Bicycles	✘	0.00	0.00			
101	1107141	Animal drawn vehicles	✘	0.00	0.00			
102	1107209	UNBR Operation of personal transport equipment		0.00	0.00			
103	1107221	Fuels and lubricants for personal transport equipment	✓	0.97	0.97	2	0.28	0.28
104	1107231	Maintenance and repair of personal transport equipment	✓	9.48	9.48	5	0.60	0.60
105	1107241	Other services in respect of personal transport equipment	✘	0.00	0.00			
106	1107309	UNBR Transport services	⊕	2.15	0.00	1	0.03	
107	1107311	Passenger transport by railway	✓	1.60	1.60	1	0.03	0.03
108	1107321	Passenger transport by road	✓	71.04	71.04	4	2.35	2.35
109	1107331	Passenger transport by air	✘	0.00	2.15			0.00
110	1107341	Passenger transport by sea and inland waterway	✓	11.71	11.71	1	0.19	0.19
111	1107351	Combined passenger transport	✘	0.00	2.15			0.00
112	1107361	Other purchase transport services	✘	0.00	2.15			0.03
113	1108019	UNBR Communication		0.00	0.00			
114	1108111	Postal services	✓	0.34	0.34	1	0.20	0.20
115	1108211	Telephone and telefax equipment	✘	0.00	0.00			
116	1108311	Telephone and telefax services	✓	0.50	0.50	1	0.06	0.06
117	1109019	UNBR Recreation and culture		0.00	0.00			
118	1109109	UNBR Audio-visual, photographic and information processing equipment		0.00	0.00			
119	1109111	Audio-visual, photographic and information processing equipment	✓	19.03	19.03	8	0.41	0.41
120	1109141	Recording media	✓	2.16	2.16	2	0.10	0.10

No	BH code	Basic heading label	% households		Source items	Share in tot. cons. (%)	
			original	mapped		original	mapped
121	1109151	Repair of audio-visual, photographic and information process. equipment	✘	0.00	0.00		
122	1109209	UNBR Other major durables for recreation and culture		0.00	0.00		
123	1109211	Major durables for outdoor and indoor recreation	✘	0.00	0.00		
124	1109231	Maintenance and repair of other major durables for recreation and culture	✘	0.00	0.00		
125	1109309	UNBR Other recreational items and equipment, garden and pets		0.00	0.00		
126	1109311	Other recreational items and equipment	✘	0.00	0.00		
127	1109331	Garden and pets	✘	0.00	0.00		
128	1109351	Veterinary and other services for pets	✘	0.00	0.00		
129	1109409	UNBR Recreational and cultural services	⊙	4.05	0.00	1	0.03
130	1109411	Recreational and sporting services	✓	1.89	5.69	1	0.01
131	1109421	Cultural services	✓	16.77	19.32	5	0.09
132	1109431	Games of chance	✘	0.00	4.05		0.00
133	1109511	Newspapers, books and stationery	✓	61.67	61.67	3	1.07
134	1109611	Package holidays	✘	0.00	0.00		
135	1110111	Education	✓	60.89	60.89	12	2.79
136	1111111	Catering services	✓	61.80	61.80	6	2.07
137	1111211	Accommodation services	✓	1.33	1.33	2	0.23
138	1112019	UNBR Miscellaneous goods and services		0.00	0.00		
139	1112109	UNBR Personal care		0.00	0.00		
140	1112111	Hairdressing salons and personal grooming establishments	✓	69.47	69.47	1	0.32
141	1112121	Appliances, articles and products for personal care	✓	98.76	98.76	7	2.04
142	1112211	Prostitution	✘	0.00	0.00		
143	1112309	UNBR Personal effects n.e.c.		0.00	0.00		
144	1112311	Jewellery, clocks and watches	✓	23.52	23.52	4	0.91
145	1112321	Other personal effects	✓	18.35	18.35	6	0.08
146	1112411	Social protection	✘	0.00	0.00		
147	1112511	Insurance	✘	0.00	0.00		
148	1112609	UNBR Financial services n.e.c.		0.00	0.00		
149	1112611	FISIM					
150	1112621	Other financial services n.e.c.	✘	0.00	0.00		
151	1112711	Other services n.e.c.	✓	3.63	3.63	3	0.36
152	1113111	Purchases by residential households in the rest of the world					
153	1113112	Purchases by non-residential households in the economic territory of the country					
✓	Data collected with sufficient level of detail (for _UNBR, expected nb of reporting hholds is 0)		>	71	(ideal: 107)	328	100.00
✘	No data (not in questionnaire / cannot be annualized / no household reported it)		>	36	(ideal: 0)		
⊙	Data collected with insufficient level of detail (aggregated for multiple basic headings)		>	5	(ideal: 0)		

## Appendix 4 – List of ICP basic headings contained in the “fake” (unbroken) basic headings

ALL	Items contained in UNBR groups (see ICP_SEQ column in Table 2)																																				
UNBR Food and non-alcoholic beverages	4	5	6	7	8	10	11	12	13	14	16	17	19	20	21	22	24	25	27	28	30	31	32	34	35	36	37	39	40								
UNBR Food	4	5	6	7	8	10	11	12	13	14	16	17	19	20	21	22	24	25	27	28	30	31	32	34	35	36	37										
UNBR Bread and cereals	4	5	6	7	8																																
UNBR Meat	10	11	12	13	14																																
UNBR Fish and seafood	16	17																																			
UNBR Milk, cheese and eggs	19	20	21	22																																	
UNBR Oils and fats	24	25																																			
UNBR Fruits	27	28																																			
UNBR Vegetables	30	31	32																																		
UNBR Sugar, jam, honey, chocolate and confectionery	34	35	36																																		
UNBR Non-alcoholic beverages	39	40																																			
UNBR Alcoholic beverages, tobacco and narcotics	43	44	45	46	47																																
UNBR Alcoholic beverages	43	44	45																																		
UNBR Clothing and footwear	50	51	52	54	55																																
UNBR Clothing	50	51	52																																		
UNBR Footwear	54	55																																			
UNBR Housing, water, electricity, gas and other fuels	57	58	60	61	63	64	65																														
UNBR Water supply and miscellaneous services relating to the dwelling	60	61																																			
UNBR Electricity, gas and other fuels	63	64	65																																		
UNBR Furnishing, household equipment and routine household maintenance	68	69	70	71	73	74	75	76	78	79	81	82	83																								
UNBR Furniture and furnishings, carpets and other floor coverings	68	69	70																																		
UNBR Household appliances	73	74	75																																		
UNBR Tools and equipment for house and garden	78	79																																			
UNBR Goods and services for routine household maintenance	81	82	83																																		
UNBR Domestic services and household services	82	83																																			
UNBR Health	87	88	89	92	93	94	95																														
UNBR Medical products, appliances and equipment	87	88	89																																		
UNBR Out-patient and hospital services	92	93	94	95																																	
UNBR Out-patient services	92	93	94																																		
UNBR Transport	98	99	100	101	103	104	105	107	108	109	110	111	112																								
UNBR Operation of personal transport equipment	103	104	105																																		
UNBR Transport services	107	108	109	110	111	112																															
UNBR Communication	114	115	116																																		
UNBR Recreation and culture	119	120	121	123	124	126	127	128	130	131	132	133	134																								
UNBR Audio-visual, photographic and information processing equipment	119	120	121																																		