International Comparison Program

[04.01]

ICP Survey Framework

ICP Book

Fred Vogel

3rd Technical Advisory Group Meeting
June 10-11, 2010

Paris, France

Table of Contents

_Toc261478771

The ICP Survey Framework
ICP Book - Chapter 12

Section 1. Overview

One of the primary purposes of the International Comparison Program is to provide the means to convert national estimates of respective Gross Domestic Products to a common currency. The foundation of the ICP is that the GDP is measured strictly according to the 1993 System of National Accounts and that Purchasing Power Parities based on the comparison of national prices for a selected basket of goods and services be used as the currency converters.

Chapter 3 provides the scope of the GDP expenditures and their breakdown as required by the ICP, and Chapter 4 explains the concepts that define the prices to be collected.

The purpose of this chapter is to describe the overall survey framework used for the collection of prices for the 2005 ICP, provide a review of the effectiveness of the 2005 survey process, and conclude with a summary of steps to be taken to improve the data collection for the ICP 2011. The survey framework to be implemented in 2011 will build off requests from data users for more information about the national scope of the data collection; especially the price levels between urban and rural areas. The survey framework will also guide the regions in determining the product definitions to be priced, the number of products to be priced, and the selection of outlets for the price collection.

The next section will introduce the conceptual framework that first defines the difference between the target and ICP populations of prices and products. This concept guides the selection of products to be priced which will be explained in Section 3. Each country will want to include products widely consumed and considered representative of its price and consumption patterns. However, what is important in one country may not be so in another. Therefore, each country will need to agree to also price products that may be comparable with those in other countries even though they may not be important to their economy. The comparison of "like with like" or comparability is the heart of the ICP.

Section 4 will review the steps taken to determine the number of products to be priced within each basic heading. This is a crucial step because the PPPs for individual products vary considerably even within the same basic heading. When there is a large variation in within basic heading product PPPs by country, then more products need to be selected for the price collection. Smaller numbers of products need to be selected for pricing for those basic headings where there are only small differences in the product PPPs.

Section 5 presents the determination of the sampling frame and the selection of outlets for price collection. The number of outlets, the type of outlet, and their location and distribution across the country all need to be considered. Guidelines are provided for each of these steps.

The chapter concludes (chapter 6) with a review of steps to be taken for the 2005 ICP to improve the survey framework.

Section 2. Conceptual Framework for Price Collection

The discussion begins with an elementary review of the target population of the ICP and how that relates to the estimation of PPPs. The **target population** is the national GDP of each country and its component aggregates. The goal is to compare the national GDP's and components by converting them to a common currency. The process starts by building up the GDP expenditures in each country from the basic heading level as defined in Chapter 3. The total expenditures for a given basic heading are essentially the summation of all of the transactions (price times quantity) that take place during the year for every product included in the basic heading added across products, or $\sum\sum PQ$.

This can include hundreds to thousands of different products with the number increasing as characteristics such as type of packaging, quantity and quality variables such as brands are used to more precisely define each product. In addition, there may be thousands or millions of individual transactions taking place each year.

The **target price** for each product in a basic heading is the national annual average which is a weighted average of the price of each transaction and the quantity purchased at that time.

$$\overline{p_i} = \frac{\sum_j \left(p_{ij} q_{ij}\right)}{\sum_j q_{ij}} = \sum_j w_{ij} p_{ij} \qquad \text{where} \qquad w_{ij} = \frac{q_{ij}}{\sum_j q_{ij}}$$ The first step in the estimation of PPPs is to define the characteristics of the products so they can be compared by matching "like with like" to compute the PPP for each product which is (P_{ai}/P_{bi}) for countries (a) and (b).

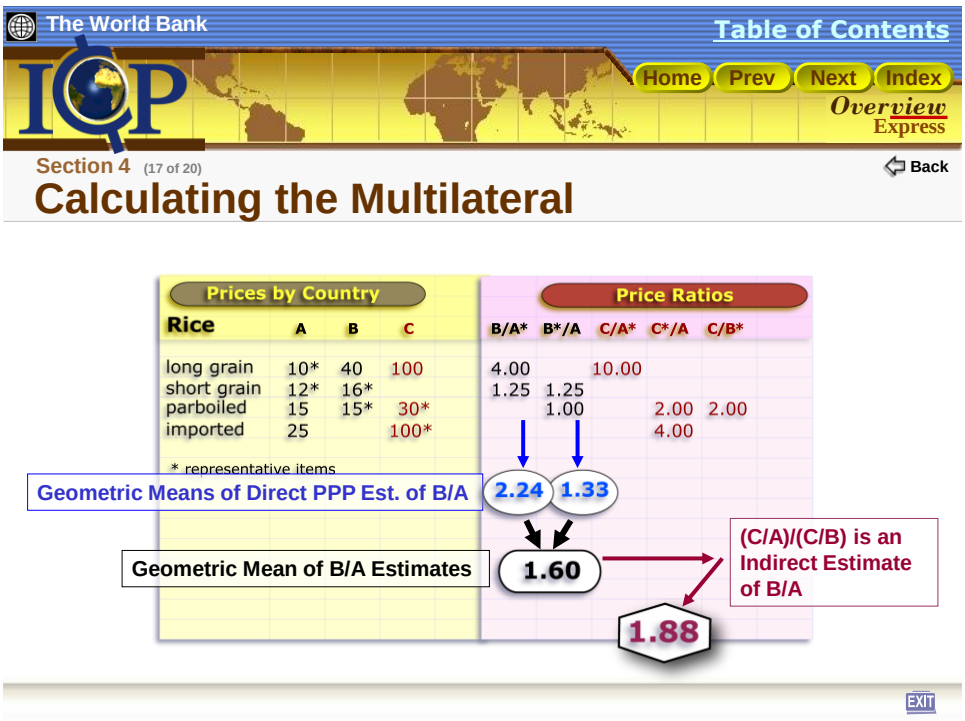
The **target PPP at the GDP** level between any two countries is the weighted average of the product PPPs across all products. Chapters 6 and 7, respectively, provide the methodology used to average product PPPs to the basic heading level and from there to the GDP.

The ICP Population is the subset of products that are available in at least two countries within a region. However, inferences based on this subset of products are made to the target population. This means each country should ensure the regional product list contains those important to its economy. **ICP prices** are not transaction prices, but prices observed in a sample of outlets which are used to estimate the national average price which by inference is equal to the **target price.** The following example illustrates these concepts.

Suppose there were only three countries in the world, and the rice basic heading contained a total of 4 products. Table 1 provides an example showing the national annual average prices for each country in its national currency. Methodology described in previous chapters describes how these product PPPs are combined to form the basic heading PPPs. The asterisks are used to identify

those products that are an important part of the country's consumption for the rice basic heading. While long grain rice is important for country A, it is less important than other rice products in countries B and C. However, countries B and C submitted prices for long grain because it was still available.

Table 1. Product PPPs within a basic heading



The target population, prices, and PPPs set the stage for the survey framework. The reality for the ICP is that:

- There can be considerable variability between the product level PPPs within a country. Note that the product PPPs for p_i^B/p_i^A range from 1.00 to 4.0. Price levels for different products vary depending on consumer preferences, volume of sales, income level of households, and location in the rural/urban domains. This points out that the subset of products to be priced needs to represent the overall price levels and variation between countries . The variation between the product PPPs will vary by basic heading. For example, the egg basic heading contains very homogeneous products while others such as garments and motor cars are very heterogeneous. This means different numbers of products need to be defined for each basic heading depending on the variability of the product PPPs.
- Not every product in say country A can be found in countries B or C, etc. In other words, there is only a subset of products within each country that can also be found in other countries. The size of this subset is influenced by the level of detail used to describe the products. For example, rice is consumed in every country; however, long grain rice, prepackaged with no broken kernels may not be available in every country. The goal is to select products from the overlapping subset that have relative price levels that represent the entire basic heading whether or not they are eligible for pricing for PPPs.
- It is not possible, even within the domain of overlapping products, to obtain a price for each one in every country. Furthermore, products are not of equal importance to every country. Within each basic heading there will be products widely consumed in Country A with price levels that are representative of the general expected level —especially for

products not in the domain that overlaps with products in other countries. In addition, there will be products that may be available in Country A, but not consumed widely and perhaps also not found in all parts of the country. However, these may be important in other countries. For these reasons, the process to define the product characteristics and to choose the products to be included in the comparison is critical.

- Countries are treated equally regardless of the relative sizes of their GDP's. This means the selection of products needs to be balanced across countries regardless of their size.
- With thousands of transactions taking place, it is not practical to collect individual transaction prices. This has two consequences. First, the quantities of each product purchased are not available which means weights are not available for averaging product PPPs to the basic heading level. The second consequence is that prices are obtained from outlets where purchases are made. The total number of outlets exceeds what can be included in the price collection. The requirement for national annual average prices means that decisions will have to be made about locations within the country where price collection will take place, the choice of outlets within those locations, and the frequency of data collection.

Section 3 which follows describes the process to describe the characteristics of the products to be priced. This process takes into account the fact that basic heading PPPs are essentially the average of the PPPs based on individual products. The goal when defining the products is to select those that collectively will provide the basic heading PPP that would be an unbiased estimate of the target population PPP. The overall reliability of the PPPs at the level of the basic heading is dependent on the product specifications, the number of products priced, where they are priced, and the number of countries that provide the prices.

Section 4 provides a process to determine the number of products to be priced followed by section 5 that provides the requirements for the sample frame for the price collection.

Section 3. Determining Product Specifications

Structured Product Definitions

A new approach for product classification and identification was developed for the 2005 ICP. This approach is based on a new international product coding system and a process to describe price determining characteristics using Structured Product Descriptions. These Structured Product Descriptions (SPDs) provided a standardized process to create the detailed specifications for products to be priced.

The first step was to harmonize existing classifications for household consumption items. This included the OECD-Eurostat Classification of Expenditures on the GDP as described in Chapter 3, and the Classification of Individual Consumption for Purpose established for Household Budget Surveys (COICOP). The OECD-Eurostat classification served as the base classification structure for the 2005 ICP to assign products to basic headings. The OECD classification included 222 basic headings. For ICP purposes, this was considered to be too detailed, especially for developing countries. As a result, the 222 Basic Headings were combined into 155. The

COICOP classification structure was mapped to the OECD structure of 155 basic headings so that countries using the COICOP classification structure could be integrated into the ICP process.

The next step after the OECD and COICOP classifications were harmonized was to add detail about price determining characteristics for products within each basic heading. The starting point was the coding system established by the US Bureau of Labor Statistics (BLS) for the Consumer Price Surveys for the Consumer Price Index. This is commonly referred to as the BLS checklist which is used during the price collection stage to identify the specifications of the products being priced. The BLS checklist is designed in such a way that it contains lists of characteristics that describe products in a consistent way. Each BLS checklist contains a list of characteristics describing a cluster of similar products within a basic heading. In some cases the product cluster was the same as the basic heading. For example, there was only one BLS product cluster each for the rice and fresh milk basic headings. However, there were several product clusters for the garment basic heading where there were separate product clusters for men, women's, and children's clothing.

The characteristics related to each BLS product cluster were used to form SPDs. There was one SPD for each product cluster within a basic heading. Annex A shows two examples of SPDs. The first is for the fresh milk basic heading which is represented by one SPD. The second SPD is for women's clothing which is one of several product clusters in the garment basic heading, each represented by an SPD. In other words, there are separate SPDs for men's and children's clothing. A review of these two examples illustrates the different number of combinations of price determining characteristics that can be used to define individual products.

The initial SPDs for each BLS product cluster were prepared by the Global Office. The next step was a review by the regional coordinators to ensure the product characteristics reflected the realities of the countries in their regions. The SPDs were updated to reflect the input from the regions. For example, type of milk—cow or buffalo—was added to the milk SPD.

The SPDs can be used to define a large number of different products, even for a basic heading as homogeneous as rice. Rice includes white and brown rice, long, medium and short grain, and varieties such as basmati sometimes sold under a brand name in a variety of package types and sizes. Quality can enter into the definition, for example, varying percentages of broken rice.

The SPDs contain the following classification variables:

- Quantity and Packaging: Simply the units in which the product is sold. The specification should provide a range of number of units, or their size or weight, that determines the product to be priced. The price per unit may be different if a liter of milk is priced vs a .25 liter container.
- Source: Usually identifies whether the product is produced domestically or imported.
- Seasonal availability: Important for fruits and vegetables and indicates whether the product is available around the year or only seasonally.
- Product characteristics: The contents of this section vary depending on the product. The number of characteristics depends on the heterogeneity of the products being specified. Note the example of women's clothing as an example.

- Brand/label name: Brands play an important role in the specification of products. International brands and model numbers may by themselves completely define a product. However, the characteristics of even branded products should be defined because they can be sold in different sizes or models. The Brand identifier should be viewed as an additional characteristic that is superimposed on an otherwise complete product specification. The following table from the ICP Handbook defines the role of brands in product specifications.

Table 2. The Role of Brands in Product Specifications.

<table>
<tr><th></th><th colspan="2">Specification names a single international brand or a cluster of international brands</th><th>Branded product, but brand not named in specification</th><th>Product without any brand</th></tr>
<tr><td>Brand value</td><td colspan="3">Some brand value exists</td><td>No brand value</td></tr>
<tr><td rowspan="2">Product searched for by price collectors</td><td colspan="2">Actual brand(s) and model(s) as specified; they should be found in most or all countries in the region</td><td>National or local brands which have a reputation only within a country or locality</td><td>Products without a brand name</td></tr>
<tr><td>One single brand</td><td>One out of a cluster of named brands</td><td>One out of a set of unnamed brands widely known within the country or locality</td><td>An unbranded product whose name or label, if any, has no significance to the buyer</td></tr>
<tr><td>Typical selling point</td><td colspan="2">The reputation of the producer and assumed quality of the product</td><td>The reputations of the producer, shop or other outlet and assumed quality of the product</td><td>Low price</td></tr>
</table>

Brands can have a significant price affect because of perceived or real quality differences. A general guideline is that price comparisons should only be made between products within the brand stratum. In other words, if a product with the same specifications has an international brand name in one country and is brandless in another country, they should be considered to be different products and not directly compared. Another guideline, however, is to exercise caution when including international brands is to ensure they are consumed widely by the consumers. In some cases, branded products may be comparable between countries, but be consumed by only a tiny number of consumers because it is a luxury item.

Using SPDs to define product specifications (PSs)

The recommended process for each region to determine the product specifications using the SPDs was for each country to first map their CPI products to the SPDs. Each mapping determined a product specification. Each country submitted this initial set of Product Specifications (PS) to

their regional coordinator. The regional coordinator reviewed these PSs to determine where there was overlap, or where a change or additional price determining characteristics would result in a product described in such a way that several countries could provide prices. This was an iterative process repeated several times and culminated in a meeting of the national coordinators where they agreed upon the final specifications of the products to be priced.

This iterative process is based on some complex concepts underlying the preparation of the product specifications. Products can be very tightly specified with absolute characteristics to be met for matching. An alternative is to tightly specify some characteristics, but leave some latitude to the price collectors for others. For example, the rice specification may call for long grain rice, but leave it to the price collector to determine the type and size of container and record those values along with the price. This provides the opportunity for the country to provide more prices, however, it can produce more variability into the matching exercise unless prices are adjusted, for example, to standard quantities or package weights.

A pre-survey is an important part of the process to define product specifications. The final test of a product specification is to determine if price collectors in each country can actually find and price the same product. A lesson learned from the ICP 2005 is that many products had to be re-defined after the first data collection because the review of the prices revealed that products were not tightly specified leading to different products being priced. In some cases, problems occurred when translating the product specifications into the local language.

Another guideline is that a picture is worth a thousand words. Where possible, photographs of the products should be part of the data collection package

Important and less Important Products (Representative and non representative products—per Eurostat-OECD)

It should be clear by now that comparability of the products being priced is an essential principle underlying the estimation of PPPs. A dilemma facing the ICP since its beginning is that even though a product may be available in several countries, it may only be important or a significant part of the consumption in a few countries. While rice is a staple in some countries, bread, cereals and meat may be more important in others. Should the product PPP for a type of rice consumed widely in country A but not in country B receive equal weight with the type more important in country B? To overcome this dilemma, the OECD-Eurostat practice is for each country to classify every product being considered for inclusion in the price collection into one of two categories—representative, or not representative but still comparable. A representative or important product is one that accounts for a significant share of a country's expenditures within a basic heading. The representative or non representative classification is determined for products within the basic heading and is country specific.

Each country will want to price products that are purchased by a large proportion of its population and account for a significant part of the total expenditures of the basic heading. While some of these products may be available in other countries, those countries may have other products more important to their economy. For this reason, the EKS* method was developed whereby countries

classify each product as representative or non representative. This provides a form of weighting as shown in table 3 below.

This example illustrates a binary comparison between two countries. There are 14 products,1-10 are Important in country J, products 8-12 are Importantin country K, and products 8-12 are Important in both countries. Prices were provided for products 13 and 14, but they were less Important in both countries. Note that the relative weight each product PPP receives is dependent on the classification of importance and the number of products priced. Product PPPs are computed by first using product prices that were representative for country J, and again for products important for country K. The geo mean of each column provides an estimate of the PPP of county K to J. The geo mean of these two estimates becomes the binary PPP between the two countries. If there are more countries, the EKS procedure is employed and brings in the less important products if they were representative in other countries.

Note that **Product PPPs carry more weight as a country provides fewer prices.** Also note that the PPPs for products 8, 9, and 10 carry more weight than the other products.

The underlying principle guiding the use of the important/less important coding is that it also provides a form of weighting for the more important products.

Table 3. Illustration of the use of representativity and the number of products priced.

Product	**Country J**	**Country K**	**J^*/K**	**J/K^***	
1	Rep	Non rep	ppp		
2	R	N	ppp		A. Each product PPP for country K carries twice weight of each product PPP for country J B. PPPs for products 8,9, and 10 carry more weight that the other product PPPs. C. Products 13 and 14 not used in binary comparison— will enter as indirect comparisons in EKS*
3	R	N	ppp		
4	R	N	ppp		
5	R	N	ppp		
6	R	N	ppp		
7	R	N	ppp		
8	R	R	ppp	ppp	
9	R	R	ppp	ppp	
10	R	R	ppp	ppp	
11	N	R		ppp	
12	N	R		ppp	
13	N	N			
14	N	N			CPRD will include PPPs for products 13 and 14.
			PPP J^*/K	PPP J/K^*	
			Geo Mean J^*/K and J/K^*		

As noted in table 1, the measurement of basic heading PPPs becomes increasingly sensitive to the choice of products the more the patterns of relative prices diverge from country to country. Table 4 shows how relative prices may be derived from the national prices using basic heading PPPs . The prices in columns 2 to 4 of table 4 are the national prices converted into the currency of Country A using the basic heading PPPs for countries B and C, or 2.356 and 88.42 respectively. Country A is the base, therefore, its PPP is 1.0. The prices in column 5 are the average international prices or the geometric averages of the converted national prices. The prices in columns 6 to 8 are the relative prices where the relative price is defined as the national price divided by the average international price for that product.

Table 4. Variability in relative prices

Product	National prices converted into the currency of A using the basic heading PPPs			International price (geometric average)	Relative prices (National price divided by the international average price)				
	A	B	C		A	B	C		Rel Std
1	10	17.82	18.10	14.78	0.67	1.21	1.22	1.00	.31
2	18	18.68	8.48	14.18	1.27	1.32	0.60	1.00	.40
3	15	16.98	14.48	15.45	0.97	1.10	0.94	1.00	.08
4	12	8.49	10.86	10.34	1.16	0.82	1.05	1.00	.17
5	12	10.19	20.36	13.55	0.89	0.75	1.50	1.00	.40
6	8	6.37	6.33	6.86	1.16	0.93	0.92	1.00	.14
PPP A = 1	1.00	2.36	88.42	Geometric average	1.00	1.00	1.00	1.00	
				Relative STD	.22	.25	.30		

Two main points need to be made from this example. The list of products should be balanced meaning that each country should have products whose relative prices are both above and below average. The other point is economic theory shows that relative prices and quantities are generally negatively correlated. In other words, products with large expenditures will have lower prices than specialty items. This implies that importantproducts should have prices below average.

In ICP 2005 countries were asked to classify all goods and services priced for Household Final Consumption Expenditure (HFCE) into two groups – "representative" and "non-representative". This was done in order to improve the accuracy of the PPPs. Prices of products that were identified as representative by a country were to be given more weight in calculating that country's PPP.

A "representative" product was defined as one whose price was deemed to be representative of the price level of the products in a Basic Heading. The countries participating in the OECD-

Eurostat and CIS groups have been identifying representative products for the last several years, but countries in other regions had difficulties in making the distinction. All countries tried to do so but it was apparent that they were using different criteria. As a result "representativity" was not taken into account in calculating PPPs in other regions for ICP 2005.

For ICP 2011 countries in regions other than OECD-Eurostat and CIS regions will be asked to classify all goods and services in Household Final Cconsumption Expenditures that are AVAILABLE as either IMPORTANT or LESS IMPORTANT. If a good/service is NOT AVAILABLE in the country, the notion of IMPORTANT / LESS IMPORTANT is not applicable to that good/service. Importance is defined by reference to the expenditure share of the item within a Basic Heading. Products that are identified as important by a country will then be given more weight in calculating its PPPs.

How to determine if a product is IMPORTANT or LESS IMPORTANT

Defining importance by reference to expenditure shares raises an obvious problem in that countries are never asked to provide expenditure weights below the Basic Heading (BH) level. The BHs are in fact defined as the most detailed level of expenditures for which countries can reasonably be asked to supply expenditure shares. Countries cannot therefore be expected to classify goods and services according to their known expenditure shares. What they are asked to do instead is to say whether, if expenditure shares were available at the product level, the expenditure shares for each product would be large or small within the BH. If it is thought that the expenditure share, if known, would probably be large, the item is classified as IMPORTANT; if small it is classified as LESS IMPORTANT.

There are three basic rules for deciding if a product is IMPORTANT or LESS IMPORTANT:

Rule 1. Is it in the CPI?
If an item is the same as, or very similar to, one that is included in a country's consumer or retail price index, the country should always classify it as IMPORTANT. (Note however, that products that are in the ICP lists but are not in the CPI may still be IMPORTANT.)

Rule 2. Use expert judgement/common knowledge.
The statistician can call upon his or her own knowledge of what are widely available and commonly bought brands of cigarettes, soap powder, biscuits, toothpaste etc.

Examples: Cheddar cheese is sold in almost all food shops but Brie is only available from speciality shops. Cheddar is IMPORTANT; Brie is LESS IMPORTANT. Kleenex face tissues are sold in every supermarket and chemist shop. A "100 piece box of Kleenex facial tissues" is an IMPORTANT product and other types are LESS IMPORTANT.

Rule 3. Ask experts
Most often the experts will be shop-keepers. The success of their business depends on knowing which products are best sellers and which are bought less often.

Example: Two kinds of breakfast cereal are specified in the product list – "Kellogs Cornflakes, family size" and "Country Store Muesli, 500 grams packet". The shopkeeper may say that both

are best sellers so both are IMPORTANT, neither is sold in large quantities so both are LESS IMPORTANT, or one or other is a best-seller so one is IMPORTANT and the other is LESS IMPORTANT.

Importance is determined within a Basic Heading

An important product is one that has a large expenditure share within the Basic Heading to which it belongs. An important product may have a very small expenditure share within HFCE as a whole but still be important within its Basic Heading.

For example, in many countries few people buyk wine but this does not mean that all the products specified under the Wine BH are LESS IMPORTANT. Within the Wine BH there may be one or two types of wine that are best-sellers and the wine merchant can almost certainly identify them. These particular wines are IMPORTANT within the BH even though their expenditure share may be very small in total HFCE.

Some Basic Headings are rather heterogeneous

Several BHs are rather heterogeneous; they contain a range of products that serve different purposes. Heterogeneous BHs should be split into homogeneous sub-groups before deciding on importance.

For example, the BH Newspapers, books and stationery is heterogeneous and should be split into newspapers, books and stationery before assigning importance to particular products. Garments is another heterogeneous BH as it includes clothing for men, women and children. It should be split into these three components before assigning importance.

Many of the heterogeneous BHs are combinations of the more homogeneous BHs in the OECD-Eurostat classification on which the ICP Expenditure Classification is based. The ICP Expenditure Classification shows which of these more homogeneous BHs have been combined. This is a useful guide in splitting up heterogeneous BHs before allocating products to the IMPORTANT and LESS IMPORTANT categories.

How countries should proceed

The importance of products needs to be taken into account both while the product lists are being drawn up and when the prices are being validated.

When the core list and the regional product lists are being drawn up, the statisticians involved must determine, by means of a pre-survey, what are important products for each BH. They must ensure that these products are included in the core list and in the regional list.
When the core list and the regional list have been finalised, the country statisticians should then consider all the products under each BH. The lists will already include products that were identified as IMPORTANT in their countries but now they will also include products identified as important by other countries in the region. Each of these should be evaluated using the three basic rules above and be classified as either IMPORTANT or LESS IMPORTANT.

LESS IMPORTANT does not mean that the product can be ignored.

Countries will provide prices for all products that they have identified as IMPORTANT. But they are also required to price products that they have classified as LESS IMPORTANT. This is essential in order to provide links with other countries.

Review of product specifications during data collection

It is essential that the product specifications be reviewed after the first data collection using the diagnostics from the Quaranta and Dikhanov tables described in Chapter 16. However, table 4 above can be used to show how to evaluate each product. The variability of the relative prices for product 2 ranges from .6 to 1.32 with a relative standard deviation of .40. This is a signal that either the product specifications are too "loose" or that one or more countries have miss-interpreted some of the specifications. The specifications for any product resulting in relative prices with relative standard deviations above .30 should be thoroughly examined. In some cases, it may turn out that the product is not comparable and should be removed from the list or redefined for the next round of data collections.

Section 4. Determining the number of products to price.

The overall reliability of the PPPs at the level of the basic heading and higher levels aggregation depends on the interaction of three factors:

- The number of products to be priced within each basic heading which will depend upon the heterogeneity of the basic heading;
- The number of products for which countries actually collect prices which is a reflection of the amount of overlap of available products across countries:
- The numbers of outlet prices collected for each product for which price surveys are actually undertaken.

At the same time that the product specifications are being developed, attention also needs to be given to the number of products to be priced in each basic heading, and also the number of countries that will be able to price them.

The purpose of this section is to illustrate the sources of variability inherent in the estimation of PPPs and how to use the knowledge about the variability to set targets for the number of products each country should try to price within a basic heading.

Table 4 above illustrates the variability of the relative product prices within the basic heading by country. This variability is measured by the relative standard deviation of the relative prices within each country which is .22 for country A and .25 and .3, respectively for countries B and C. Even though the sample of products is not from a random selection, the principles of sampling theory can be used to determine the number of products to price—annex A provides a useful

overview and a more detailed explanation is provided in Chapter 12 A. The goal is to price enough products that the sampling error of the mean of the product PPPs is within a target level of precision.

Table 5 provides some examples of the relative standard deviations from ICP 2005 and a target number of products to price. Table 2 in Chapter 12 A and examples provided there can be used as a guide to determine the number of products to price within each basic heading.

Table 5. Example of target number of products to price

Product	Relative STD of relative price ranges over countries	Target number of products	Number in 2005 Ring
Electricity	.03-.05	3	5
Rice	.10-.20	10 +	6
motorcars	.15-.25	10-45	30
Fresh/frozen seafood	.20-.25	10-45 +	10
Garments	3.0-3.5	25-100	68
Pharmaceuticals	2.5-3.5	25-100	43

Electricity is usually furnished by a very small number of providers; therefore, there is very little variability in the rates as evidenced by the relative standard deviations ranging from .03 to .05 across countries. With these small deviations, a country may only need to price 3 items to be 90 percent confident the resulting PPP is within 5 percent of the target. Products such as milk and eggs also exhibit very little price variability.

The relative price levels for rice are more variable across countries ranging from .10-.20. This suggests that more than 10 products be priced unless the country or region is willing to accept a level of precision at the 10 percent rather than the 5 percent level.

Another point to be considered at this point is the relative importance of the basic heading. If the basic heading is an important part of the consumption basket, then a country will want to have a precise measure of the relative prices. Therefore, the country or region would want to target a number of products to achieve a 5 percent level of precision. However, if the basic heading has a very small weight, then the target level of precision can be increased to 10 percent so that the resources can go towards the more important basic headings.

A final point to consider is that not every country will be able to price every product. For that reason, the target number of products will need to be increased so that each country can price the minimum number. In other words, as the size of the overlap domain decreases, it will be necessary to define more products so that each country can submit prices for a minimum number. A relevant point is that the number of products prices should be similar across the countries being prepared. Because expenditure weights are not available for individual products, the only weighting comes from the Importance classification and the number of products each country prices.

Section 5. Sample design and determining the number of price observations.

The purpose of this section is to provide the guidelines to determine the number of price observations to estimate the national average prices as well as the type of outlets and their locations where the prices will be collected. As stated in the introduction, the target price is the weighted average of the prices at which the product is sold during different times of the year and over the country using the quantities purchased by month and location. In other words, this implies the sample must relate to the entire country and to the entire year. In some countries, there are auxiliary data that can be used to calibrate capital city data to the country and a point in time price to the entire year. Chapter 13_will provide the methods to compute the national average prices for these cases.

It should be noted that the required sample size is not dependent on the size of the country, but on the heterogeneity of prices across the retail markets. The greater the price variability across the markets, the larger the sample is required for the same level of precision.

While the data collection must provide a national annual average price, each country must work within the framework of information available to make up a sampling frame and select the desired outlets. Ideally, each country should have a frame or register of all outlets frequented by consumers; each outlet should have measures of size reflecting its volume of sales. This register could be either stratified by size or samples of outlets selected using probabilities proportionate to size. The problem is that even if the measures of size were available, they may not always be reflective of the sales of individual products. For example, a meat market may also sell fruits and vegetables; however, its volume of sales is more reflective of the meat sales. Because of cost considerations, it is good practice to price what is available in an outlet once the price collector is there.

For these reasons, a purposive sample of outlets is used for the ICP price collection, as is usually the practice for their price surveys for the CPI. A starting point is the frame established for the CPI. (See the Practical Guide to Producing Consumer Price Indices—Sampling Procedures). One problem is that the CPI price surveys in many countries are urban based. This is appropriate when the relative price changes over time are the same across the urban and rural sectors.

Many countries have distinctive rural and urban sectors that exhibit very different pricing patterns and levels. Even though they reflect the same changes in prices over time, they may have different price levels. In these cases, the sample size for both the rural and urban price collections should be large enough to provide reliable estimates for each sector. Countries should use information from the most recent household expenditure survey to determine the relative coverage of the urban and rural sectors. Annex B provides an overview of the rural expenditures as a percent of the total for selected countries and product categories. Note that rural expenditures make up a large portion of the total for food items. It is also quite likely that the products consumed by the rural areas may not be the same as those consumed in urban areas. On the other hand, it is also likely that some of the products to be surveyed such as motor cars are only available in urban areas.

Probably the single largest source of questions about the 2005 ICP was about the degree of the urban/rural price coverage. The table in Annex C was provided in response to those questions and

shows the urban/rural coverage by country. This shows that there was a lack of consistency in the coverage of the rural areas which led to questions about the reliability of the data for some purposes. Therefore, a goal for ICP 2011 is to improve the coverage of the rural areas.

The selection of the outlets is especially important because different products have different distribution profiles. Some products are sold mostly in supermarkets, but may also be available in a range of other outlets from specialty shops to local traditional markets. Prices for the same product can vary from outlet type to outlet type because of varying circumstances such as services provided. For these reasons, the selection of outlets should take into account the different types of outlets and their relative share of the overall expenditures. This will usually require expert judgment because of the lack of a sample frame with expenditures by outlet or outlet type. Some guidelines or considerations follow for the selection of outlets:

- The first consideration is that the selection of outlets by type should be in proportion to the volume of their sales of the products to be surveyed.
- The next consideration is the variability of the prices within and between outlet types. The guidelines provided in Annex A regarding the sample size apply here as well and should be used to determine the number of outlets to be included in the price collection.
- The third consideration is the location of the outlets, especially the urban/rural domains. Again, the number of outlets by location should be in proportion to each area's share of the volume of sales.
- A final consideration is to determine the number of outlets or price observations that should be collected. Information about the variability of the prices is needed and decisions need to be taken about the desired level of precision following the guidelines in Annex A. and the more detailed explanation in Chapter 12A.

The advantage of selecting outlets by type and location by volume of sales is that it provides a self weighting sample, therefore, simplifying the estimation of the national average prices.

In response to the questions coming from data users, countries will be asked to provide the following indicators for each observation of product prices. This will allow the country to break down the national average prices into the urban and rural components to better understand the price distributions and to respond to questions about the national coverage of the product. This information can also provide each country a better understanding of how the underlying rural and urban price levels affected their PPPs.

Table 6. Outlet type and location indicator

Outlet Types			Location
	Types	**Examples**	**Capital city, other urban, or rural**
1	Large shops	Supermarkets, hypermarkets, department stores, etc.	
2	Medium & small shops	Minimarkets, kiosks, neighborhood shops, grocery stores, convenience stores. etc.	
3	Markets	Open markets, covered markets, wet markets, etc.	
4	Street outlets	Mobile shops, street vendors, etc.	
5	Bulk and discount shops	Wholesale stores, discount shops, etc.	
6	Specialized shops	Supply shops, hardware shops, furniture shops, etc.	
7	Private service providers	Taxi cabs, hotels, restaurants, private schools, private hospitals, etc.	
8	Public or semi-public service providers	Water suppliers, electric power companies, public schools, public hospitals, etc.	
9	Other kinds of trade	Online (Internet) shopping sites, catalogue orders, etc.	

In an ICP context, the number of outlets or price observations to take depends upon the required precision and the relative importance of the product in the basic heading. This is liable to vary from country to country so that different sizes of sample may be appropriate in different countries. The variability of the prices between outlets will also itself tend to vary between countries. The appropriate size of sample depends on the net result of a set of interacting factors. It is a matter on which national coordinators may wish to consult with regional coordinators. It must also be remembered that a product PPP is the ratio of the estimated average prices in two different countries. It might not be optimal for one country to spend a lot of resources achieving a high degree of precision in its estimated price for some particular product if other countries do not, or cannot. This is a matter which may call for some collective discussion and some general guidelines at a regional level. Such guidelines would have to be specific to a particular set of countries and particular set of products.

The difficulty and costs of collecting outlet prices could vary significantly between different types of products. When it is difficult to collect prices for a particular type of product, say because the product is not very common and found only in a very few widely dispersed outlets, it may be judged to be cost effective not to try to collect any prices for that product and concentrate on collecting prices for products that are more important and readily available. Such a strategy may increase the total number of price observations for important products but risks introducing bias by reducing the number of products priced.

Section 6. Summary

The concept of Importance will be used in the 2011 ICP which means this be determined by every country for every product priced. The Important/less Important coding should also be a part of the data validation exercise to ensure it is applied consistently across countries.

Since the importance coding will be finalized during data validation, countries will also be asked to price items they consider to be less important. Some simple guidelines should be that they price less important items available in the outlets being surveyed for the important products; they should not go to great expense to search out each one.

Guidelines based on the statistical variability of relative prices of products within basic heading were provided to determine the number of products to be priced. These along with the relative share of the basic heading to the total contribute to the decision about the number of products. Similar principles also apply to the number of individual price observations to obtain.

The final significant requirement is that each price observation be coded to indicate the outlet type and the urban/rural dimension.

Annex A. Determining the Number of products to price and the number of price observations (Yuri Dikhanov)

This annex provides the framework to determine the number of product specifications to be prepared by basic heading, the target number countries should price, and the number of price observations to be made for each product. The number of price observations will translate into the number of outlets to be selected for the price survey.

The main purpose of this annex is to examine the relationship between the size of the sample, whether it be the number of products to be priced or the number of price observations and the probable margin of error, or precision, attached to the national annual average price or the basic heading PPP. The same points about margin of error also apply to the desired level of decision for the estimated basic heading PPPs. This analysis draws on classic sampling theory. The central limit theorem states that if a population has a finite variance σ^2 and an arithmetic mean μ then the distribution of the sample mean in repeated random samples drawn from that population approaches the normal distribution with a variance of σ^2 / n and a mean μ as the sample size n increases. The sample mean provides an unbiased estimate of the population mean. The probability of the sample mean not deviating from the population mean by more than a certain amount can then be derived from the area under the normal curve. In this way, probable margins of error can be attached to sample means. An explanation of sampling errors and confidence intervals can be found in any textbook on probability and statistical theory.

In practice, the population standard deviation, σ, will not be known but may be estimated from the sample itself, from other samples drawn from the same population, or in other ways, as explained below. It is convenient to replace the estimated value of the standard deviation, s, by its value relative to the estimated mean, m: i.e., s / m. This is the relative standard error as measured during price collections to determine the national average price. It also applies to the relative standard deviations of the relative prices as evaluated in the Quaranta and Dikhanov tables.

It is then possible to construct tables showing, for example, the minimum size of sample needed to ensure that the probability of the sample mean deviating from the population mean will not exceed some specified amount. Such a table is shown below.

The table is constructed on the assumption that a 10 % level of significance is required. Its use may be illustrated by the following example. Suppose that the estimated relative standard deviation, s /m ,is 0.2 or 20% (second column) and also that the required precision level is 5% (first row). The entry in the first row and second column is 45. This means that a sample of 45 is needed to ensure that there is 90 % probability that that the sample mean does not deviate from the population mean by more than 5%.

The greater the variance in the population, the lower the level of precision in the estimated mean for any given size of sample. Conversely, the larger the size of the sample, the higher the level of precision in the estimated mean achieved for any given variance in the population. The size of sample needed to achieve a given level of precision, say 5% , may increase sharply with the relative standard deviation: for example, when s / m increases for 0.2 to 0.3 the minimum sample size needed more than doubles from 45 to 100.

Annex B. Rural expenditures as a percent of total, selected countries and products

BASIC HEADING ITEMS	Senegal	S. Africa	India	Indonesia	Brazil	Kazakhstan	Yemen
Rice	51.28	43.71	67.31	57.24	24.76	45.88	36.30
Other cereals, flour and other prod	63.44	54.06	63.18	69.92	34.06	66.33	60.84
Bread	25.54	32.37	36.36	30.88	9.61	33.44	3.53
Other bakery products	68.66	15.58	51.80	40.30	16.61	37.75	34.16
Pasta products	23.30	18.61	34.65	37.91	20.29	51.18	23.97
Beef and veal	13.25	29.50	60.05	24.00	19.94	33.90	23.29
Pork	8.48	18.81	65.26	55.76	32.08	35.12	0.00
Lamb, mutton and goat	31.39	17.01	53.20	46.05	29.41	79.06	33.72
Poultry	14.40	34.59	57.49	32.04	21.36	31.09	31.61
Other meats and meat preparation	46.37	19.34	68.68	16.35	20.99	24.11	20.61
Fresh, chilled or frozen fish and sea	27.42	16.31	68.68	44.34	30.50	34.62	22.85
Preserved or processed fish and se	46.07	41.22	59.20	50.66	15.83	20.67	31.03
Fresh milk	50.35	22.82	55.59	24.68	22.91	47.41	90.13
Preserved milk and other milk prod	17.38	30.88	49.99	22.85	13.13	40.41	50.10
Cheese	1.23	13.41	0.00	3.79	13.08	19.77	17.69
Eggs and egg-based products	5.34	33.41	56.37	41.30	23.90	34.67	29.43
Butter and margarine	7.05	24.41	46.47	11.04	13.62	40.19	51.42
Other edible oil and fats	41.80	37.28	57.76	51.48	24.20	43.78	34.01
Fresh or chilled fruits	42.81	26.59	48.22	39.20	15.10	27.22	28.27
Frozen, preserved or processed frui	71.07	18.19	45.98	44.43	28.67	27.89	35.03
Fresh or chilled vegetables other th	31.33	29.96	57.24	48.99	26.41	34.49	30.46
Fresh or chilled potatoes	26.79	38.46	66.31	50.71	23.77	42.08	30.91
Frozen, preserved or processed veg	19.85	31.57	59.37	45.20	18.16	42.07	31.41
Sugar	48.32	47.72	61.93	53.15	29.83	44.48	44.66
Jams, marmalades and honey	32.00	21.23	27.76	20.95	19.40	45.15	36.16
Confectionery, chocolate and ice cr	20.27	14.41	48.26	19.30	11.76	38.17	45.14
Food products n.e.c.	41.94	27.50	56.42	35.98	18.72	37.95	39.13
Coffee, tea and cocoa	50.33	31.18	48.59	45.55	22.88	44.24	45.78
Mineral waters, soft drinks, fruit an	16.93	22.58	29.32	13.76	11.01	29.76	20.22
Spirits	27.10	20.02	60.25	80.09	13.07	45.77	0.00
Wine	23.04	26.13	79.00	47.24	16.60	34.52	0.00
Beer	22.35	34.65	38.20	28.95	13.37	25.03	0.00
Tobacco	48.75	23.65	62.12	48.87	13.68	34.53	27.31
Clothing material, other articles of	31.57	27.61	63.07	43.78	12.92	28.91	33.56
Garments	40.00	26.02	55.92	42.74	12.75	34.11	29.89
Cleaning, repair and hire of clothing	28.58	10.77	58.69	36.92	9.16	13.55	33.84
Shoes and other footwear	29.79	27.41	50.99	40.29	12.69	36.71	31.85
Repair and hire of footwear	31.25	0.00	0.00	0.00	0.00	0.00	30.28
Actual and imputed rentals for hous	25.95	18.57	1.99	25.29	8.52	0.00	9.41
Maintenance and repair of the dwel	59.64	14.70	63.09	37.73	13.69	0.00	21.49
Water supply	15.80	8.71	18.71	8.66	3.50	18.45	18.26
Miscellaneous services relating to th	0.00	7.09	0.00	0.00	12.01	2.23	61.10
Electricity	6.30	16.99	36.61	29.57	8.55	32.82	10.88
Gas	6.08	54.03	35.12	25.25	17.19	40.62	39.78

Annex C. Outlet Information by Location

Country	Outlet information by location	
	Region or State/ Province	Urban/ Rural
Bangladesh	23 districts (out of 64)	urban (37 markets) and rural (20), mostly included in CPI
Bhutan	20 districts + 2 big cities	two cities, 21 urban towns, all major rural areas
Brunei Darussalam	3 districts	small country
Cambodia	capital + 5 province cities	urban
China	11 large cities	primarily urban
Hong Kong, China	throughout economy	urban and rural
Macao, China	entire area	
Taiwan, China	16 survey areas	urban and rural (8 cities and 34 towns/townships)
Fiji	two cities + 4 towns	small country
India	22 states	31 urban centers collect on everthing 201 rural villages collect only on food, clothing and footwear, education
Indonesia	28 provinces to represent urban-rural; West-East Indonesia; Java-outer islands; and large-medium cities.	
Iran, Islamic Rep.	30 provinces	urban in 30 provinces (30 capitals + 50 other cities) and rural in 28 provinces (62 villages)
Lao PDR	capital + 7 provinces	urban in capital and 4 provinces and rural in 3 provinces
Malaysia	14 states	urban (36 capita and urban centers) and rural (15 rural centers)
Maldives	capital + 4 other islands	small country
Mongolia	capital + 21 provinces	urban (capital + two cities) and rural (19 provinces)
Nepal	four domains (mountain, hill, terai, Kathmandu)	urban (14 centers) and rural (17)
Pakistan	4 provinces	urban (35 cities and 71 markets)
Philippines	17 regions	urban for capital and urban/rural for other regions
Singapore	throughout economy	small country
Sri Lanka	24 districts	urban and rural (17 districts have both, 3 have only urban and 4 have only rural)
Thailand	16 provines and capita	urban
Vietnam	20 provinces	urban and rural
Argentina	Grand Buenos Aires	urban
Bolivia	Capital + 3 cities	urban
Brazil	6 major cities	urban
Chile	capital and 12 cities	urban
Colombia	Capital + 3 major cities	urban
Ecuador	2 major cities	urban
Paraguay	Gran Asuncion	urban

Peru	Capital + 4 cities	urban
Uruguay	Capital + 4 cities	urban
Venezuela, RB	Grand Caracas	urban
Austria	capital city	urban
Belgium	capital city	urban
Germany	capital city	urban
Luxembourg	capital city	urban
Netherlands	capital city with main urban areas	urban
Czech Republic	capital city	urban
Hungary	capital city	urban
Poland	capital city	urban
Slovak Republic	capital city	urban
Slovenia	capital city	urban
Switzerland	capital city	urban
Denmark	capital city	urban
Finland	capital city	urban
Ireland	capital city	urban
Sweden	capital city	urban
United Kingdom	capital city	urban
Estonia	capital city	urban
Latvia	capital city	urban
Lithuania	capital city	urban
Iceland	capital city	urban
Norway	capital city	urban
France	capital city	urban
Greece	capital city	urban
Italy	capital city	urban
Portugal	capital city	urban
Spain	capital city	urban
Cyprus	capital city	urban
Malta	capital city	urban
Bulgaria	capital city	urban
Romania	capital city with main urban areas	urban
Turkey	capital city	urban
Australia	capital city with main urban areas	urban
Canada	capital city with main urban areas	urban
Japan	capital city with main urban areas	urban
Korea, Rep.	capital city with main urban areas	urban
Mexico	capital city with main urban areas	urban
New Zealand	capital city with main urban areas	urban
United States	capital city with main urban areas	urban
Israel	capital city with main urban areas	urban
Albania	capital city	urban
Bosnia and Herzegovina	capital city	urban
Croatia	capital city	urban
Macedonia, FYR	capital city	urban
Montenegro	capital city	urban
Serbia	capital city	urban

Armenia	throughout country	
Belarus		
Kazakhstan		
Kyrgyz Republic		
Moldova		
Azerbaijan	capital city	urban
Ukraine		
Georgia		
Russian Federation		
Tajikistan	food throughout country, but others in Dushanbe	
Angola	9 provinces	province capital plus two to three rural areas accessible from the provincial capital
Benin	all 12 departments	urban (urban centers) and rural (village closest to urban centers)
Botswana	every Census district has at least one collection center (32, 52% population & 69% consumption)	all town/city (100%), some urban villages (63%) and rural villages (4%)
Burkina Faso	10 regions	region center and adjacent rural area with largest population within a radius of about 20 km
Burundi	7 zones	urban (urban centers)
Cameroon	all 10 regions	urban (10 urban centers) and rural (10 rural areas)
Cape Verde	3 islands	urban and rural in all three islands
Central African Republic	7 administrative regions/10 prefectures	urban (urban centers) and rural (rural locality closest to urban centers)
Chad	8 regions	urban and additional rural markets
Comoros	all 3 islands	urban and rural (331 towns/villages)
Congo, Dem. Rep.	11 provinces	urban (10 centers) and rural (10 centers)
Congo, Rep.	11 departments	urban (6 centers) and rural (20 locations)
Cote d'Ivoire	capital + all 9 other regions (by rural/urban)	urban (10 region centers) and rural (9 largest prefectures near region centers)
Equatorial Guinea	6 provinces/8 municipality	urban/rural
Egypt, Arab Rep.	11 gvernorates (66 collection centers)	2 governorates have urban (6 each) and 9 have both (3 urban &3 rural each)
Ethiopia	13 regions	Addis + 12 urban areas + new rural areas
Gabon	five zones plus two largest cities	urban (123 outlets) and rural (only weekly markets 7)
Gambia, The	all 8 regions/8 local government area	
Ghana	10 regions	41 urban + 19 rural markets
Guinea	capital and 4 zones	urban and rural (one region capital and one rural weekly market nearby for each zone)
Guinea-Bissau	capital and 7 regions	urban and rural (except for capital all regions have both)
Kenya	all regions	Nairobi (10 areas) + 15 urban centers + 10 new rural centers
Lesotho	all 10 regions	urban + additional 1 or 2 villages for each region

Liberia	all 5 regions	urban (five region largest towns) and rural (rural area surrounding largest town)
Madagascar	6 provinces	urban (7 large urban centers, 8 other urban centers) and rural (25 rural locations)
Malawi	all 27 districts (except Island of Likoma)	4 cities and 29 rural centers (6 total for districts with the 4 cities and one each for other districts)
Mali	capital and 8 regions	urban and rural
Mauritania	13 regions	urban (13 regional centers) + rural (5 rural centers in 3 regions)
Mauritius	all 10 districts	urban and rural
Morocco	8 regions	urban (8 regional centers) and rural (14 rural markets)
Mozambique	4 provinces	urban (4 cities) and rural (2 villages in each province)
Namibia	3 zones (9 out of 13 regions)	capital + surrounding rural areas
Niger	capital + 7 administrative regions	urban (capita1 + 7 region centers) and rural (7 largest rural weekly market nearby)
Nigeria	6 zones	46/23 rural/urban centers sampled within 6 zones & by urban/rural (to have five price obs per center per item)
Rwanda	capital and all 12 provinces	urban (capital + 12 province centers + 3 other cities) and rural (one location in each province)
Sao Tome and Principe	2 islands/8 districts	urban (29 centers) and rural (14 centers)
Senegal	5 regions	urban (8 centers) and rural (5 centers)
Sierra Leone	4 provinces/regions	urban (5) and rural (3) collectio n centers
South Africa	9 provinces	urban (50 collection centers)
Sudan	16 states (northen Sudan)	urban (28 largest markets in state capital cities+one additional market in other town for each state) and rural (additional two rural village markets for each state)
Swaziland	10 towns	urban (10 centers) and rural (9 centers)
Tanzania	7 zones (11 regions)	urban (11 out of 20 CPI centers) and rural (one center each for the same 10 regions, each with 4 villages having weekly markets)
Togo	capital + 5 regions	urban and rural
Tunisia	7 regions/24 governorates	urban and rural
Uganda	capita and 4 zones	urban (six urban centers + one more in Northern zone) and rural (eight centers)
Zambia	all 9 provinces(41 districts)	urban (10 centers) and rural (38 centers)
Zimbabwe	all 10 provinces	urban (88 centers) and rural (32 centers)
Bahrain		urban
Egypt, Arab Rep.	3 regions	urban/rural
Iraq	capital and several large cities	urban
Jordan	3 regions/all kingdom governors (12)	urban/rural
Kuwait		urban
Lebanon	5 regions	Mostly urban with some coverage of rural towns
Oman	6 regions	urban

Qatar	3 regions	urban
Saudi Arabia	most regions	mostly urban with some rural coverage (Bedouin villages) for some groups
Syrian Arab Republic	Damascus metropolitan area	mostly urban with relatively low coverage of rural areas
Yemen, Rep.	most regions	mostly urban with some rural coverage for some groups

Prices for consumption only. Housing, health, education, government and gross fixed capital formation are not included.

*** Prices used in PPP:**

N national representative

U primarily urban

C only capital city

Annex D. Determining Sample sizes

The problem of determining sampling size for the ICP is far from trivial. In short, only general guidance can be given, as a more precise estimate requires prior knowledge of the population characteristics. In the ICP framework, there are two critical elements. One is determining the number of products that should be priced, the other is to establish the survey frame and select outlets for price observations. The number of price observations needed to estimate the national average prices will in most cases determine the number of outlets to be in the sample.

According to the Central limit theorem of statistics, there is a 95% chance that the [$\boldsymbol{y \pm 2SE}$] interval will contain the population mean, and a 68% chance for the [$\boldsymbol{y \pm SE}$] interval to contain the mean, where $\boldsymbol{y}$ is the sample mean, and $\boldsymbol{SE}$ is sampling error**.** This is, obviously, under the condition that the sample size is large or prices are distributed normally. The problem, of course, is that the characteristics of the distribution are not known prior to the actual price collection, and the form of the distribution is not known either. Normal distribution would be a very rough approximation in our case, as real prices, probably, conform more to a truncated log-normal distribution. One can argue that there should exist a physical [or rather economical] limit below which prices do not normally fall, determined by costs of production and distribution. The price may fall below that limit only in the case of the product going out of circulation, close-out sale or the product being a loss-leader to entice customers to come to the store.

All these cases are supposed to be filtered out at the stages of product identification and actual price collection. At the time of selecting products to be collected in a country, that country has to make sure that the products are representative and can be freely purchased. During actual price collection, the price collectors are discouraged from recording any special discounts or otherwise not normal prices [see section … for more detail].

Normal practice in sampling is to collect enough data to generate an estimate within a given confidence interval (say, the 95% or 68% intervals). In the CPI it is possible to use knowledge from previous surveys to infer the shape of the distribution and its characteristics. It is very difficult to do the same in the ICP. At the very least, it requires pre-surveys to sample at least some of the products to study their behavior. After which the results would be extended to other products from the same categories.

Unfortunately, life gets even more complicated as other errors and biases (non-stochastic ones) interplay with purely statistical sampling errors as our sample gets exposed to seasonality movements, selection bias etc. This section provides boundaries for the sample size estimate given different assumptions about distribution.

The Weak Law of Large Numbers (Chebyshev's Theorem)

The Weak Law of Large Numbers (Chebyshev's Theorem) states, with some modification relevant for our sampling problem, that:

$$P(|\bar{x}-\mu| \geq k\frac{\sigma}{\sqrt{n}}) \leq \frac{1}{k^2}$$

and written differently (with $\varepsilon = k\frac{\sigma}{\sqrt{n}}$),

$P(|\bar{x}-\mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}$, or $\delta \leq \frac{\sigma^2}{n\varepsilon^2}$, and $n \geq \frac{\sigma^2}{\delta\varepsilon^2}$.

Or, that the sampling average $\bar{x}$ converges in probability to the population mean $\boldsymbol{\mu}$ and the probability of the estimate being further away from the mean than ***k*** sampling errors ($\frac{\sigma}{\sqrt{n}}$) is less than ($\boldsymbol{1/k^2}$). The sample size is ***n***. The Chebyshev's Theorem does not say anything about the form of the distribution; it is valid for any distribution. Thus, for ***k=4.5*** the confidence interval will be approximately 95%. In the case of the normal distribution for the same confidence interval it is enough to have ***k=2***. In price measurement we can assume that ***2<k<4.5*** because we do not know exact form of the distribution. For the 68% confidence interval the result will be ***1<k<1.7***.

The Central Limit Theorem

Now, in order to determine the sample size, we will proceed as following. As we know from the Chebyshev's Teorem: $n \geq \frac{\sigma^2}{\delta\varepsilon^2}$ (for any distribution). On the other hand, the sample size can be obtained from the Central Limit Theorem. Roughly, the central limit theorem says that the sum of a number of independent samples taken from any distribution is approximately normally distributed (converges in probability to the normal distribution). The key word here is "converging", meaning that the distribution can be considered normal if ***n*** is large enough.

How large is large enough depends on the particular distribution. In the case of the normal distribution, any sample size will obviously generate normal sums. In the case of a bi-modal distribution, or an unusually skewed one the sample size should be at least 30-50. This essentially tells us that we have to try making sampling cells as homogeneous as possible. The cell in this case will be "outlet type – region" or similar breakdown. Breaking down price data by outlet type, in particular, will eliminate market segmentation that, if exists, may generate a bi-modal distribution, which is particularly difficult to evaluate properly.

It is important to make cells to correspond to smallest analytical cut of the price data. In the case of ICP, it is essential to have some regional breakdown (at least capital city-other urban areas-rural areas), as well as the outlet type breakdown. This information will make national price data

more precise, and as a by product, it can also be utilized in studies of regional and vertical poverty, in determining poverty lines and unemployment payments among other uses.

More formally, the Central Limit Theorem can be written as follows:

$$\lim_{n\to\infty} P\left(\left|\frac{x-\mu}{\sigma/\sqrt{n}}\right| \le k\right) = \Phi(k) - \Phi(-k),$$

where $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{1}{2}x^2}\, dx$.

Or, we can rewrite it as $\lim_{n\to\infty} P\left(|x-\mu| \ge k\frac{\sigma}{\sqrt{n}}\right) = 2 - 2\Phi(k)$, given that $\Phi(a) + \Phi(-a) = 1$.

An important (for us) result from the Theorem is:

$$\Phi\left(\frac{\varepsilon}{\sigma}\sqrt{n}\right) \ge 1 - \frac{\delta}{2},$$

or $\frac{\varepsilon}{\sigma}\sqrt{n} \ge \Phi^{-1}\left(1-\frac{\delta}{2}\right)$

and $n \ge \left(\frac{\sigma}{\varepsilon}\Phi^{-1}\left(1-\frac{\delta}{2}\right)\right)^2$.

The last formula is valid if σ is known. If it is not known (as it happens most of the time in price statistics), Student's ***t***-distribution is to be used ($t = \frac{x-\mu}{s/\sqrt{n}}$), where ***s*** is sample standard deviation (STD) (see Table 1). This distribution is similar to normal but with bigger tails and a lower hump in the middle.

Unfortunately, ***t***-distribution has ***n-1*** (degrees of freedom) as its parameter, which makes it somewhat more difficult to use in estimating ***n*** (a special table was prepared for this purpose, see Table 2[1]).

Thus, we can sum up that the required minimum sample size should be within the following bounds:

$$\left(\frac{\sigma}{\varepsilon}\right)^2 \left(\Phi^{-1}\left(1-\frac{\delta}{2}\right)\right)^2 < n < \left(\frac{\sigma}{\varepsilon}\right)^2 \frac{1}{\delta},$$

the lower bound being determined by the approximately normal distribution of prices and known σ and the upper bound being valid for any distribution.

And if σ is not known, the lower bound is found from:

$$\left(\frac{s}{\varepsilon}\right)^2 \left(t_{n-1}^{-1}\left(1-\frac{\delta}{2}\right)\right)^2 < n,$$

where t^{-1} is inverse t-statistics.

[Some general rules about using ***t***-statistics:
Although the *t* procedure is fairly robust, i.e. it does not change very much when the assumptions of the procedure are violated, one should always study the price data to check for skewness and outliers before using it on small samples. Here small can be interpreted as n < 10-15. If our sample is small and the data is clearly nonnormal or outliers are present, ***t*** should not be used. On the other hand, *t*-statistic can be safely used even when the sample indicates the population is clearly skewed, if n > 30. In Table 2 cases where n>30 are grayed out].

Example

Let's consider the following example:

$s = 20\%$

$\varepsilon = 10\%$

$\delta = 32\%$

It translates into the following: given sample standard deviation ***s*** (20%) [here we are using relative values, i.e. STD divided by mean], we want to determine the number of necessary measurements ***n*** needed to be 68% certain (i.e., 100% - 32%) that the sample mean falls within ***s/2*** (10%) of μ.

Chebyshev's theorem yields n=12 in this case and the Central Limit Theorem yields n=6 (see Table 2, find significance level 0.32, sample standard deviation 20% and target precision 10%). It's clear that the number of necessary measurements has to be closer to 12 if we do not know the exact distribution, but we may relax it to be 6 if the distribution is known to be close to normal.

Table 1. Inverse t-Table

df	**0.317**	0.100	**0.050**	0.001
1	1.837	6.314	12.706	636.578
2	1.321	2.920	4.303	31.600
3	1.197	2.353	3.182	12.924
4	1.142	2.132	2.776	8.610
5	1.110	2.015	2.571	6.869
6	1.091	1.943	2.447	5.959
7	1.077	1.895	2.365	5.408
8	1.067	1.860	2.306	5.041
9	1.059	1.833	2.262	4.781
10	1.053	1.812	2.228	4.587
11	1.048	1.796	2.201	4.437
12	1.043	1.782	2.179	4.318
13	1.040	1.771	2.160	4.221
14	1.037	1.761	2.145	4.140
15	1.034	1.753	2.131	4.073
16	1.032	1.746	2.120	4.015
17	1.030	1.740	2.110	3.965
18	1.029	1.734	2.101	3.922
19	1.027	1.729	2.093	3.883
20	1.026	1.725	2.086	3.850
30	1.017	1.697	2.042	3.646
40	1.013	1.684	2.021	3.551
50	1.010	1.676	2.009	3.496
60	1.008	1.671	2.000	3.460
70	1.007	1.667	1.994	3.435
80	1.006	1.664	1.990	3.416
90	1.006	1.662	1.987	3.402
100	1.005	1.660	1.984	3.390
∞	1.000	1.645	1.960	3.290

Table 2. Relationship between observed sample standard deviation and sample size, based on t-statistics

Significance level **0.317**

target precision (%)	*sample relative STD (%)*				
	10	20	30	40	50
5	6	18	38	66	102
10	3	6	11	18	27
15		4	6	9	13
20		3	4	6	8
25			3	4	6

Significance level **0.100**

target precision (%)	*sample relative STD (%)*				
	10	20	30	40	50
5	14	46	100	176	273
10	5	14	27	46	70
15		9	14	22	33
20		5	9	14	20
25			6	9	14

Significance level **0.050**

target precision (%)	*sample relative STD (%)*				
	10	20	30	40	50
5	19	65	141	249	387
10	7	19	38	65	99
15		11	19	31	46
20		7	12	19	27
25			10	14	19

Significance level **0.010**

target precision (%)	*sample relative STD (%)*				
	10	20	30	40	50
5	31	111	243	429	668
10	11	31	64	111	170
15		18	31	52	78
20		11	20	31	46
25			16	22	31

Recommended procedure to determine sample size

The following process is recommended:

(1) during the pre-survey, several price observation for at least one item per SPD should be recorded, and sample standard deviation is estimated; this standard deviation will be used to determine an approximate sample size for all items in this SPD; if there is a possibility to carry out this estimation for all items, the country should do it. This pre-survey estimation can be carried out in the capital city;

(2) based on Table 2, the sample size is determined; it is recommended that for more important items [in terms of expenditure weight] better precision is selected (say, 10% at the 95% confidence interval), and for lesser items lower precision is deemed adequate (say, 20-25%). An attempt to compute (or, at least plot) skewness should be made; however, it is unlikely that we can determine density curve with any precision, given very small sample sizes and possible selection biases at the pre-survey stage;

(3) If computed sample size drops below 10, we should use 10 as minimum, because below that sample size the assumptions about normality of the result do not hold in general; if pre-survey results in a weird-looking distribution, Chebyshev's formula needs to be used instead;

(4) If computed sample goes above 30, the normality considerations become irrelevant, and Table 2 should be used exclusively to determine the sample size;

(5) During actual price collection, the country administrator should monitor STD for each item to be collected, and, if necessary, increase sample size to achieve precision targets using Table 2. This can be done through distributing additional collection forms to price collectors for the next stage of price collection (quarter or month, depending on the frequency of collection).

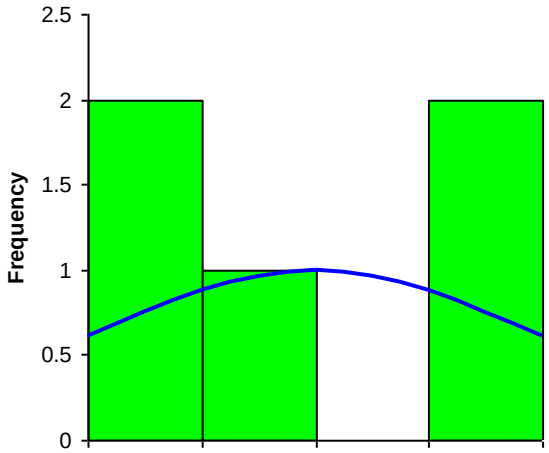
Workflow example 1

Let's consider the following example:

During the pre-survey five price observations for item ***i*** of the rice SPD were made:

p_1=1.8, p_2= 2.0, p_3=2.2, p_4=2.2, p_5=1.8

We can plot a histogram of these prices as it is shown below:



A non-normal distribution clearly can be suspected, possibly bi-modal. However, the sample size is too small to make a definite conclusion. A simple histogram is often enough to make a judgment. `Excel` can be used to plot histograms efficiently[1]. However, it is possible to use statistical packages and other tools to analyze distribution in more detail.

In this example `Analyse-It` add-on for `Excel` was used (a 30-day trial version can be downloaded from www.analyse-it.com). Any other statistical software can be used instead. The histogram above is plotted against the normal distribution line, and then two normality tests were carried out – *Kolmogorov-Smirnov* and *Shapiro-Wilk*; both tests rejected normality, but not strongly, - at the 0.15 and 0.12 significance levels, respectively, essentially supporting our observation made on the basis of the histogram[1].

Such a pattern of price observations could arise in the case when two types of outlets are mixed together, say, supermarket and village market. So it is necessary to check if it is the case here. If indeed it turns out to be the case, the observations from the outlet with smaller number of price quotes should be discarded, and analyses should be rerun.

The sample mean is 2.0 and STD is 0.20 (here `Excel` formulae **AVERAGE()** and **STDEV()** can be used), which yields $\boldsymbol{s}$ = 10%. Now, because rice is an important item the target precision ($\boldsymbol{\varepsilon}$) is selected to be 10 percent as well. Table 2 gives us sample size of seven for this case. However, because our pre-survey produced a likely non-normal distribution, t-statistics cannot be used.

We will use *Chebyshev's* formula instead ($n \geq \frac{\sigma^2}{\delta \varepsilon^2}$) with $\boldsymbol{\delta}$ = 0.05, which results in sample size of 20. This number will be used as sample size for rice. After the first batch of collected prices (for the first quarter or the first month, depending on the frequency of price collection), the computation needs to be rerun at least for some cells (*item - outlet type – region* combinations), in order to confirm the pre-survey findings. If our estimate of sample size turns out to be too optimistic, additional collection forms will be sent to price collectors starting next round of price collection.

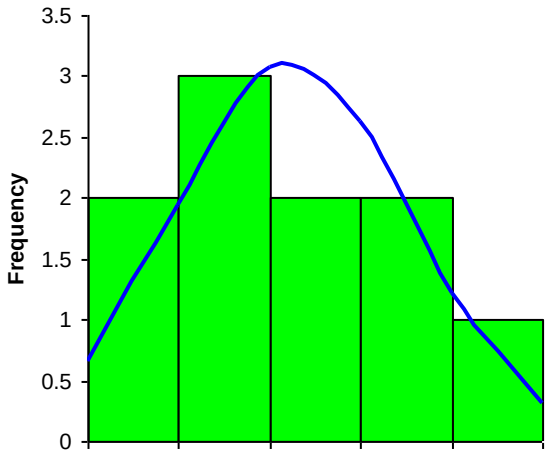
Workflow example 2

Now let's consider the following:

During the pre-survey ten price observations for item ***i*** of the breakfast cereal SPD were made:

p_1=0.8, p_2= 1.2, p_3=1.0, p_4=1.1, p_5=1.2, p_6=1.3, p_7= 1.4, p_8=1.5, p_9=1.3, p_{10}=1.7

We can plot a histogram of these prices as it is shown below:



The histogram is quite close to the normal distribution line. *Shapiro-Wilk* test produces p=0.998, strongly supporting normality and *Kolmogorov-Smirnov* test does not rejected normality [p>0.15, normally it is not shown beyond 0.15].

The sample mean is 1.25 and STD is 0.25, which yields ***s*** = 20%. Now, because breakfast cereal also happens to be an important item the target precision (ε) is selected to be 10 percent. Table 2 gives us sample size of nineteen in this case. This number will be used as sample size for the item. As in the previous example, after the first batch of collected prices, the computation needs to be rerun at least for some cells, in order to adjust sample size if necessary.