International Comparison Program

Computation of National Annual Average Prices

Draft version

Global Office

Operational Guide

Table of Contents

Computation of National Annual Average Prices[1]

1. Introduction

In the International Comparison Program (ICP), price deflators, specifically purchasing power parities (PPPs), are calculated based on the national annual average prices of a set of pre-designated products and services. In the case of the Household Final Consumption surveys where prices are collected on a quarterly basis, annual prices for the reference year are computed by taking the arithmetic means of quarterly-collected prices. Also, national prices are derived from averaging prices collected from multiple locations within a country. It is important to note that price sets are (1) first nationalized every quarter and then (2) annualized, and not the other way round. In the case where annual prices are computed first before being nationalized, it will yield different results if regional weights vary quarter by quarter. In reality, it is not always the case that all four quarter prices are available from every selected location. Sometimes data may be partially available depending on the regional and product-specific situations.

However, not all surveys follow the same frequency of data collection. For some surveys such as compensation of employees, construction or education, prices are collected once a year. Thus, they provide annual average prices. If education surveys are conducted in multiples locations of a country or data on compensation of government employees have been collected for both central and local government, it is important to annualize them first before nationalizing, which differs from the guidelines for household consumption surveys. Sometimes mid-year prices are regarded as annual prices in the case of machinery and equipment, and construction surveys.

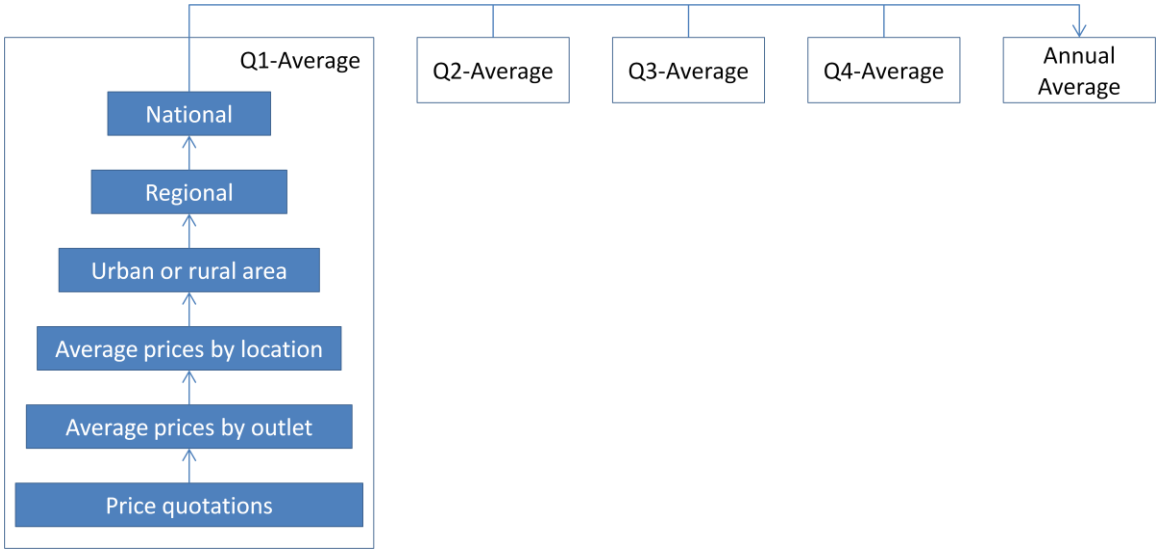
As the Household Final Consumption survey serves as the main survey of the ICP data collection, this chapter will focus more on this survey with quarterly price collection, explaining the basics of computing national annual average prices as well as special cases for dealing with missing data points.

2. Computation of National Averages

The calculation of national average prices will be carried out along two perpendicular dimensions: space and time. Spatial averaging will be aligned with the geographical stratification of price surveys, whereas temporal averaging will reflect the survey schedule. Both averaging processes will be carried out in keeping with the quarterly and annual sequencing of data validation, submission and processing.

While countries may choose or may be encouraged to compute monthly averages for validation purpose, to see how collected prices have evolved from month to month during a given quarter and whether this escalation is consistent with the monthly CPIs, in the context of ICP, monthly averaging is not considered as a mandatory step in the overall averaging process. Though it is not shown in the diagram below, it is mentioned herein for the sake of exhaustiveness only.

[1] This chapter was drafted by Nada Hamadeh and Min Ji Lee, with input from Michel Mouyelo-Katoula, based on material drafted by Yuri Dikhanov and Nada Hamadeh in the 2005 ICP.



Sequence of the averaging process

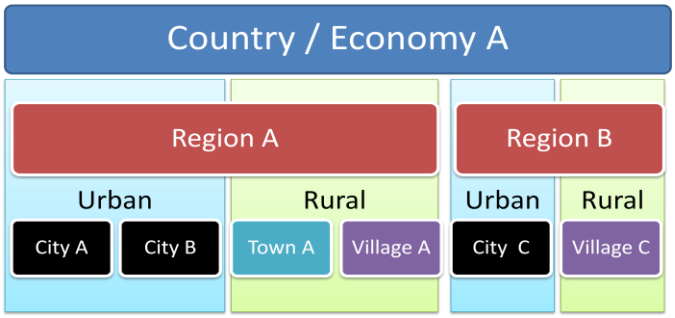
The National Coordinators are encouraged to conduct price collection in multiple locations within a country so that the outcomes will be representative national average prices. If conducted as recommended, outlets will be evenly selected from both urban and rural areas. In most countries, urban areas are generally more populous than rural areas and a good portion of population might be concentrated in few numbers of cities. In such case, where prices are collected in multiple locations with different population distributions that yield different levels of economic activity, the question arises on how to take this inequality in population density and economic activity into account. Countries should opt to use weights at various levels including national, regional (sub-national), urban/rural, location and outlets depending on their data availability at those levels. In the following section, we will discuss with examples what the options are and how weights should be treated in computation.

2.1. Without geographical weights

Difficulties arise in assigning weights to various locations as weights differ by product or a group of products, and information at the detailed level or even at the higher levels may not necessarily be available all the time. Therefore, for practical reasons, it is easier to treat every location equally, i.e. giving an even weight $1/n$ to every location, if price collection takes place in $\mathbf{n}$ locations. This is equivalent to assigning no weights as no factors other than the products' prices are taken into account during the averaging process. The method applied in the 2005 round involved taking simple arithmetic averages of prices from all locations. This may not accurately reflect reality, but we can still assume that weights are implicitly taken into consideration. Given a higher density of population it is probable that there will be more outlets to be priced in populous areas and higher level of sales come out from these areas. Countries were recommended to design their survey in a self-weighted way, ensuring that a larger number of outlets are included from populous regions, thus yielding a larger number of price quotations with a bigger impact on the overall average prices.

2.2. With geographical weights

If weights by regions within a country are readily available, it is recommended since the 2005 round that the hierarchical weighting system be used in the ICP, meaning that at any level, the total consists of its components. By that we need a breakdown of a country into different regions based on population, sales volume, expenditures, etc.



Consider a country case with a breakdown into two regions A and B, assuming the most detailed level of obtainable information is urban/rural data within a region. Then the hierarchical weighting scheme starts from the top level: for example, if Region A has a total weight of 70 (Ex: 70% of expenditures or sales), and Urban and Rural are in the proportions of 3 to 2 for that region, then Region A – Urban has a weight of 3/(3+2)*70=42, and so on. Then the rightmost column for below table, i.e. the lowest level information, will be taken into calculation as weights of each location. Note that the national coordinators should provide the weights for the lowest level where they can be reliably obtained. If weights are provided as sales volumes or other figures that do not sum up to 100, weights can be normalized accordingly.

Total 100	**Region A** 70	**Urban** 42
		Rural 28
	Region B 30	**Urban** 20
		Rural 10

Each product/item can have distinct weights associated with it. Weights, ideally, should reflect quantities of the product sold. So if, for product 1, Urban areas of Region A are responsible for 42% of total sales, this will be the weight used in determining average prices. At the same time, product 2 might have a different weight for that particular area. However, in many cases, it will be difficult to find such detailed breakdowns and some simplified procedures will need to be put in place. For example, if it is known that Urban areas of Region A are responsible for 40% of *national* food (or personal consumption expenditures), and there is no detailed data on individual products in existence, the 40% share for Urban areas of Region A can be used for *all* products comprising *national* food (or personal consumption) expenditures.

$$P_{National} = \frac{\sum_{i=1}^{n} P_{Region\ i} * W_{Region\ i}}{\sum_{i=1}^{n} W_{Region\ i}}$$

In summary, the table below captures possible options at each level, even though the list may not be completely exhaustive. In each location, multiple outlets are selected where a number of quotations are collected for a given product. These quotations are averaged using a simple unweighted arithmetic average. This average is usually self-weighted in the sense that populous locations tend to have more outlets and thus are more sampled. For the higher levels, countries are provided with options to weigh prices with sales volume, expenditure, population density or any combination of these indicators. It is worth noting that the same measure does not necessarily need be applied across levels. For instance, sales volume can be used as weights at the urban/rural level for some regions, but population can be used at the national level for simplicity. The table below is applicable, not only to products, but also to higher levels such as class, group, and category. If weights are not available at lower levels, higher level weights can be applied to all products within the level in question.

Level	Option 1	Option 2	Option 3	Option 4	Option 5
National	Unweighted	Sales Volume	Expenditure	Population	Combination of these indicators
Regional	Unweighted	Sales Volume	Expenditure	Population	Combination of these indicators
Urban/Rural	Unweighted	Sales Volume	Expenditure	Population	Combination of these indicators
Location/ Outlet type	Unweighted (self-weighted)				

Weighting options

Varying quarterly weights

As a survey for the same group products is conducted once in a quarter for four times in the reference year, seasonal items in the ICP surveys such as fruits and vegetables may have different consumption patterns by region and by quarter. In such a case, weights will be derived in a similar manner as the previous case but for every quarter. The data requirements will be more extensive in this case. All regional and urban/rural breakdowns will need to be replicated for each quarter. As the National Coordinators will be assigning different weights for each quarter, national annual average prices will be calculated based on the simple weighting method covering all quarters and all regions. See *3.6 Annual weighted average based on variable quarterly weights* for a numerical example.

$$P_{National\ Annual} = \frac{\sum_{i=1,j=1}^{i=n,j=4} P_{Region\ k,Quarter\ j} * W_{Region\ k,Quarter\ j}}{\sum_{i=1,j=1}^{i=n,j=4} W_{Region\ k,Quarter\ j}}$$

Numerical Examples

The computation of national annual averages without geographical weights can be considered as the simplified case where equal weights are assigned to each region/area. The numerical examples introduced in this section will highlight special cases and will mainly deal with the regional weighting cases *for a single quarter*.

2.2.1. All weights and all average prices provided

Items	Regional Quarter Prices Matrix Locations				Weights Matrix Locations				National Average
	A	B	C	D	A	B	C	D	
Product 1	99.01	51.30	99.00	78.42	213.00	342.00	678.00	88.00	**85.28**
Product 2	50.74	88.60	67.94	77.30	213.00	342.00	678.00	88.00	**71.14**
Product 3	8.88	33.13	79.98	57.55	213.00	342.00	678.00	88.00	**54.89**
Product 4	30.66	31.33	57.12	97.40	213.00	342.00	678.00	88.00	**48.86**
Product 5	92.96	45.84	92.04	89.98	213.00	342.00	678.00	88.00	**80.09**

In this simple case, all the weights at the regional level are available and thus average prices are computed using the Arithmetic Mean. Thus the national average price for Product 1, for instance, is:

$$\frac{99.01 * 213 + 51.3 * 342 + 99 * 678 + 78.42 * 88}{213 + 342 + 678 + 88} = 85.28$$

2.2.2. Weights completely missing for one or more products

Items	Regional Quarter Prices Matrix Locations				Weights Matrix Locations				National Average
	A	B	C	D	A	B	C	D	
Product 1	99.01	51.30	99.00	78.42	0.00	0.00	0.00	0.00	**81.93**
Product 2	50.74	88.60	67.94	77.30	213.00	342.00	678.00	88.00	**71.14**
Product 3	8.88	33.13	79.98	57.55	213.00	342.00	678.00	88.00	**54.89**
Product 4	30.66	31.33	57.12	97.40	0.00	0.00	0.00	0.00	**54.13**
Product 5	92.96	45.84	92.04	89.98	213.00	342.00	678.00	88.00	**80.09**

In the example illustrated above, weights are missing completely for Product 1 and Product 4, but are available for the remaining products. The National Coordinators may choose to use unweighted averages for those products only (Product 1 and Product 4). Alternatively, based on experts' opinions, the National Coordinators can borrow weights from a similar product or a group of similar products. Weighted averages will be calculated for the remaining products (Product 2, Product 3, and Product 5).

2.2.3. Weights partially missing for one or more products

Items	Regional Quarter Prices Matrix Locations				Weights Matrix Locations				National Average
	A	B	C	D	A	B	C	D	
Product 1	99.01	51.30	99.00	78.42	0.00	0.00	0.00	0.00	**81.93**
Product 2	50.74	88.60	67.94	77.30	213.00	342.00	0.00	88.00	**88.00**
Product 3	8.88	33.13	79.98	57.55	213.00	0.00	678.00	88.00	**62.50**
Product 4	30.66	31.33	57.12	97.40	0.00	0.00	0.00	0.00	**54.13**
Product 5	92.96	45.84	92.04	89.98	213.00	342.00	678.00	88.00	**80.09**

In the above example, weights are completely missing for Product 1 and Product 4. These two products will be treated as in Case 3.3 above. Weights for Product 2 and Product 3 are partially available. If missing weights cannot be retrieved, they will be deemed as zero. Accordingly, average prices for locations with missing weights will not be considered in the calculation of the average price for the respective products.

2.2.4. Some average prices missing

Items	Regional Quarter Prices Matrix Locations				Weights Matrix Locations				National Average
	A	B	C	D	A	B	C	D	
Product 1	99.01	51.30	99.00	78.42	0.00	0.00	0.00	0.00	**81.93**
Product 2	50.74	0.00	67.94	77.30	213.00	342.00	0.00	88.00	**58.51**
Product 3	0.00	33.13	79.98	57.55	213.00	0.00	678.00	88.00	**77.40**
Product 4	30.66	31.33	57.12	97.40	0.00	0.00	0.00	0.00	**54.13**
Product 5	92.96	45.84	92.04	89.98	213.00	342.00	678.00	88.00	**80.09**

This case is quite similar to the previous case but differs in the sense that some average prices are missing. Consider the case of Product 2. The average price for Region B is missing as well as the weight for Region C. The price from Region C for Product 2 will not be taken into consideration as explained in the previous case, and, naturally, the missing average price for Region B will not be used either.

2.3. With calibration factors

In practice, many countries collect prices only from the capital city or from a limited number of locations, and nationalize such prices when relevant calibration factors are available. This type of adjustment factors can come from CPI indices, ICP 2005 data or even from ICP 2009 exercise depending on the regional practice. When using CPI indices to make the capital city prices national, the National Coordinators should be cautious about the classification of products as ICP products might not necessarily follow their national CPI classification. Once calibration factors are verified for a given group of products, national prices can be estimated by multiplying the surveyed prices by these adjustment factors. Some countries may show a certain degree of differences between capital and national while for others this difference may be less significant.

	Country 1			Country 2			Country 3		
	Capital City	Adj. Factor	Nat'l Price	Capital City	Adj. Factor	Nat'l Price	Capital City	Adj. Factor	Nat'l Price
Product 1	199.01	0.95	189.06	78.42	1.00	78.42	8.56	0.98	8.39
Product 2	250.74	0.95	238.20	77.30	1.00	77.30	13.20	0.98	12.94
Product 3	140.30	0.95	133.29	57.55	1.00	57.55	7.40	0.98	7.25
Product 4	250.66	0.84	210.55	97.40	1.00	97.40	16.06	1.00	16.06
Product 5	192.96	0.84	162.09	89.98	1.00	89.98	15.94	1.00	15.94

3. Computation of Annual Averages

3.1. Annual weighted average based on varying quarterly weights

An ideal scenario where the National Coordinators are able to reasonably estimate varying quaterly weights, especially for seasonable products, the national annual average price will be computed based on all available weights and prices simultaneously. Treating missing weights and missing prices will not differ from cases illustrated earlier.

Product 1	Regional Quarterly Prices Matrix				Weights Matrix				Annual Av. Price
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	
Region A	41.03	20.83	48.87	95.89	213.00	250.00	225.00	200.00	46.20
Region B	31.18	42.31	10.51	41.02	342.00	300.00	380.00	300.00	
Region C	85.32	39.17	34.30	45.74	678.00	780.00	675.00	600.00	
Region D	84.32	80.80	50.32	39.10	88.00	75.00	90.00	100.00	
Quarterly national prices / weights	64.10	38.80	31.15	52.37	1321.00	1405.00	1370.00	1200.00	

If the National Coordinators price national data every quarter and then try to annualize those quarterly prices, it is crucial to apply different weights for each quarter. See the calculation below:

$$P_{National,Quarter1} = \frac{41.03 * 213 + 31.18 * 342 + 85.32 * 678 + 84.32 * 88}{213 + 342 + 678 + 88} = 64.10$$

$$P_{National,Annual} = \frac{64.10 * 1321 + 38.80 * 1405 + 31.15 * 1370 + 52.37 * 1200}{1321 + 1405 + 1370 + 1200} = 46.20$$

3.2. Annual average without varying quarterly weights

In practice, once all prices have been validated, annual average prices will be computed based on cleaned data sets of quarterly prices without varying quarterly weights. Note that this annual averaging has to correspond to the reference year. If price collection has taken place in a year other than the reference year, necessary adjustments have to be made. Similarly, if average prices for some quarters are missing, adjustments need to be made. As mentioned above, the ideal case is where we have all four quarter average prices and can compute a simple arithmetic average. Consider the example below where a country collected prices for four products (1, 2, 3 and 4) and nationalized quarterly prices:

	Quarterly National Average Prices Matrix				Annual Av. Price
	Q1	Q2	Q3	Q4	
Product 1	7.22	7.25	7.21	7.23	**7.23**
Product 2	7.22	7.2	7.23	7.2	**7.21**
Product 3	7.22	7.25	7.25	7.23	**7.24**
Product 4	7.2	7.2	7.2	7.2	**7.2**

3.3. Missing Monthly or Quarterly Prices

It is possible that countries may not have priced all four quarters, for example, due to socioeconomic reasons. If some quarter price information is missing, there is a need to use CPI indices to impute missing data points, provided that the national coordinators are able to provide sound CPI data. The example below illustrates the case of a country where quarterly average prices are partially missing.

	CPI Matrix				Quarterly National Average Prices Matrix				Annual Av. Price
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	
Product 1	102.3	103.5	103.4	103.3	7.17	7.25	7.21	7.23	**7.22**
Product 2	102.3	103.5	103.4	103.3	7.12	7.2	7.23	7.2	**7.19**
Product 3	102.3	103.5	103.4	103.3	7.12	7.2	7.25	7.23	**7.20**
Product 4	102.3	103.5	103.4	103.3	7.12	7.2	7.19	7.19	**7.17**

Note that the 1st quarter average price is missing from the original price collection. This missing average price is imputed using CPI indices from the next available quarter. If the 2nd quarter price is also missing, then we can take the 3rd quarter data for imputation and so on. In this case, 1st quarter average price can be computed as 7.17 = 7.25*(102.3/103.5). For Product 4, three quarter average prices are computed using one data point and relevant CPI indices.

A similar approach can be applied when countries conduct price collection in periods before or after the reference year. This may arise when few countries or regions lag behind in price collection due to administrative or preparatory difficulties, or when regional agencies agree with countries on survey cycles where only a portion of the surveys take place in the reference year.

The example below shows a case of a country where all price collection would have taken place in 2012 instead of 2011. Note that a quarterly national average price for Q3 2012 is 7.23, which is higher than that of Q2 2012 (7.2), while CPI indices are actually higher in Q2 2012 (103.5) than Q3 2012 (103.4). This is not unlikely as CPI indices are often available at higher levels, and not at the individual product level. When backcasting 2012 quarter prices to 2011, it is important to keep in mind that imputation should be on a quarter-on-quarter (QOQ) basis. Thus in our example, Q1 2011 would be 7.08 = 7.12*(101.7/102.3), Q2 2011 will be 7.12 = 7.2*(102.4/103.5) and so forth.

	2011				2012			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
CPI Matrix	101.7	102.4	102.0	102.1	102.3	103.5	103.4	103.3
Quarterly National Average Prices	7.08	7.12	7.13	7.12	7.12	7.2	7.23	7.2

If prices as well as CPI indices are available for every month, it is better to impute data for missing months by applying monthly CPI indices rather than using quarterly indices so that seasonality is reflected in a more accurate way. The method for imputation of missing months will be exactly the same as the case of quarterly average prices, addressed above. That is, CPI indices are to be applied at the lowest category available in which products with missing prices belong and backcasting will be based on the month-to-month basis. For instance, in the example below, January 2011 average price is imputed as 7.05 = 7.12*(101.6/102.6).

	2011				2012			
	Jan	Feb	Mar	Apr	Jan	Feb	Mar	Apr
CPI Matrix	101.6	101.2	102.3	102.5	102.6	101.7	102.6	104.3
Quarterly National Average Prices	7.05	7.08	7.11	7.12	7.12	7.11	7.13	7.24

3.4. Averaging Methods

In mathematics, various methods exist for calculations of averages; among which the most common methods are Arithmetic Mean, Geometric Mean, and Harmonic Mean. In previous ICP rounds as well as the 2011 round, average prices can be computed using any of these three methods.

To illustrate, suppose we have **n** quotations for a product. The average prices can be computed as follows:

$$Arithemetic\ Mean = \frac{1}{n}\sum_{i=1}^{n} P_i$$

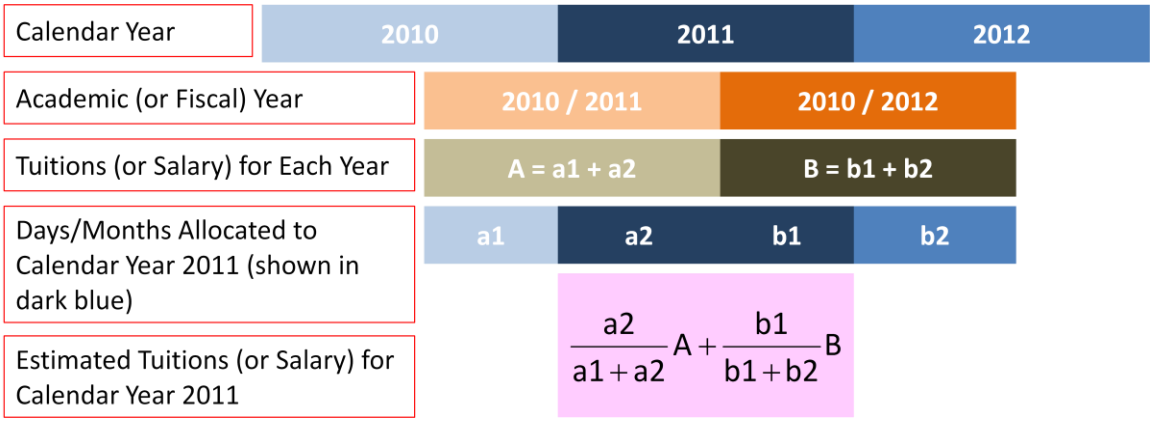
$$Geometric\ Mean = \sqrt[n]{\prod_{i=1}^{n} P_i}$$

$$Harmonic\ Mean = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} 1/P_i}$$

Although the National Coordinators have the discretion to choose which method they will be adopting, the recommended method for the ICP 2011 round is the arithmetic mean for its simplicity and to ensure consistency across countries and regions.

4. Special Surveys – Compensation of Government Employees and Private Education

For ICP Special Surveys including *Compensation of Government Employees* and *Private Education*, it is likely that the surveyed information will cover two calendar years. When the government fiscal year differs from the calendar year, *Compensation of Government Employees* will inevitably contain salary information from two fiscal years. Similar logic applies to *Private Education* as the academic year differs from the calendar year for more than half of the countries in the world. As these two special surveys pose a similar question, we provide what is required in both cases.



Assume one academic year (or fiscal year) stretches from one calendear year to the subsequent year. Then we first compute how many days or months, depending on availability, falls into the calendar year 2011 and then multiply a percentage of days/months in 2011 by the price for that academic (or fiscal) year. Once adding up the calendar year portions from two academic (or fiscal) years, we get annualized prices for the reference year of the ICP. Note that this is not a weighted method but simple summation of parts.