United Nations ECE/CES/GE.22/2010/

Economic and Social Council

Distr.: General
10 April 2010
Original: English

Economic Commission for Europe

Conference of European Statisticians

Group of Experts on Consumer Price Indices

Tenth session

Geneva, 10-12 May 2010
Item x of the provisional agenda

Revising the Consumer Price Indexes for the Eastern Caribbean Currency Union

By Paul Armknecht[1]
Price Statistics Advisor
Caribbean Regional Technical Assistance Centre

Abstract

The ECCU countries are going through the process of developing new market baskets for the CPI and introducing them. At the same time they are making improvements to the classification structure, the methods, and procedures used to compile the CPI in line with those discussed in the *CPI Manual*. This process has taken several years and is nearing completion. The ECCU countries have received substantial technical assistance from CARTAC to help with the process. They also are introducing the PIPS software that will assist them in implementing many of the improved methods described in the *CPI Manual*. As the project comes to completion, these countries will be among the first wave to upgrade their CPI to the new international standards reflected in the *CPI Manual*.

[1] The author is grateful to Mick Silver for his many useful comments. The views expressed herein are those of the author and should not be attributed to CARTAC, the IMF, its Executive Board, or its management.

Revising the Consumer Price Indexes for the Eastern Caribbean Currency Union

Introduction

The Consumer Price Index (CPI) measures the average change in the prices of goods and services purchased by households for their day-to-day living needs. The CPI is intended to measure the rate of price inflation for household consumption or, alternatively, changes in their cost of living. The latter reflects the changes in the amounts that the households need to spend in order to maintain their standard of living. Most CPIs are calculated as weighted averages of the percentage price changes for a specified set, or "basket," of consumer products.

The weights reflect the spending patterns of households during some previous period. Much depends on how appropriate and timely the weights are. The more current the weights, the more reflective they are of current consumer spending patterns. It is, therefore, important to update the weights used in the CPI periodically.

Over the past several years the statistical offices in the Eastern Caribbean Currency Union (ECCU)[2] have been working to derive new market baskets for their CPIs and introduce updated methods and procedures that meet international standards. This is also important to the regional monetary authority, the Eastern Caribbean Central Bank (ECCB) for developing a regional CPI that uses the same concepts, methods and procedures across countries. St. Lucia was the first country to introduce a new basket in 2008. Anguilla, Antigua and Barbuda, St. Kitts and Nevis, and St. Vincent and the Grenadines should complete the same process by Spring 2010.

In order to update the market baskets, Household Income and Expenditure Surveys (HIESs) are required. As part of Caribbean Development Bank projects to study poverty conditions in many ECCU countries (Country Poverty Assessment reports), member countries were provided grants to conduct Living Conditions Surveys (LCS). Most countries took the opportunity to expand the LCS and make it a full-blown HIES by adding more detailed questions on expenditures on normal goods and services and including a diary survey of households' daily expenditures.

At the request of most ECCU countries, the Caribbean Regional Technical Assistance Centre (CARTAC), located in Barbados, assisted in developing consumption expenditure estimates from the HIESs and in deriving the new CPI market baskets. CARTAC also provided training

[2] The ECCU has eight members—Anguilla, Antigua and Barbuda, Dominica, Grenada, Montserrat, St. Kitts and Nevis, St. Lucia, and St. Vincent and the Grenadines. The Eastern Caribbean Central Bank is the monetary authority and central bank for the currency union.

on the new *CPI Manual* (2004)[3] for all its members and is providing technical guidance to those requesting it on the use of new CPI software developed by the IMF's Statistics and Technology and General Services Departments—the Price Index Processor System (PIPS).

With the introduction of the revised CPIs, the ECCU statistical offices will enhance their CPIs with the following:

- Improved, expanded samples of outlets reflecting those normally visited by households. This improves the coverage of the index by including newer outlets that are important. Outlets in geographic areas not previously covered have also been added. The CPI is now more representative of the entire country, not just the capitol city area.

- New item structures that include both updated spending patterns and new products (e.g., adding college tuition, cell phones, MP3 players, computers, and health and auto insurance to the CPI basket). This improves the CPI by reflecting current patterns in consumer purchases.

- The **C**lassification **o**f **I**ndividual **Co**nsumption by **P**urpose (COICOP), the international standard recommended in the *CPI Manual*. This replaces regional classification structures that were somewhat different with 10 divisions rather than the 12 provided in COICOP.

- The introduction of owner's equivalent rent to make the CPI conceptually consistent with the *System of National Accounts* and to include owner-occupied housing in the index for the first time in many ECCU countries.

- Use of the Price Index Processor System (PIPS)—discussed in the *Notes on the Calculating the CPI* section of this paper. The use of PIPS introduces major improvements to the methods and procedures currently used in the country-CPIs, namely:

  - Compilation of indexes that use geometric averaging rather than arithmetic averaging to reduce the biases documented in the *CPI Manual* that result from arithmetic averages of prices and arithmetic averages of price relatives;

  - Imputation of missing prices that uses the price change of related products vs. carrying forward the last reported price, which has an inherent bias;

[3] International Labor Organisation *et al*. (2004). CARTAC's training was conducted in Castries, St. Lucia from June 15–26, 2009 with support from the St. Lucia Statistics Department.

  - Quality adjustments of prices to reflect pure price changes using techniques documented in the *CPI Manual*; and

  - Introducing new goods into the product and item structure using the techniques documented in the *CPI Manual*.

- Greater comparability across the region in terms of classifications, structure, methods, and procedures used in the CPIs.

Purpose and Uses of the CPI

The CPI is widely used as a general measure of inflation because of three important characteristics. It is published frequently, usually every month but sometimes every quarter. It is available quickly, usually about two weeks after the end of the month or quarter. The CPI is also usually not revised.

Accordingly, the CPI is closely monitored and attracts a lot of publicity. Because the CPI provides timely information about the rate of inflation, it has also come to be used for a wide variety of purposes. For example:

- The CPI is often widely used to adjust pensions and social security benefits for inflation.

- The CPI is also used to index other payments, such as wages, rents, interest payments, and the prices of bonds.

- The CPI is commonly used as a proxy for the general rate of inflation, even though it only measures consumer inflation.

- The CPI is used by some governments or central banks to set inflation targets for purposes of monetary policy.

- The price data collected for CPI purposes can also be used to compile other indices, such as the price indices used to deflate household consumption expenditures in national accounts, or for measuring differences in the purchasing power of money in different countries.

In addition to these major uses of the CPI, it has acquired a unique status as one of the key economic statistics in most countries. There are several factors that help to explain this:

- All households have their own personal inflation experience and often compare it to the CPI. The general public are very conscious of changes in the prices of consumer goods and services, and the direct impact those changes have on their living standards. Interest in the CPI is not confined to the press and politicians.

- Changes in the CPI tend to receive a lot of publicity. Their publication can make headline news. The CPI is a high-profile statistic.

- The CPI is published frequently, usually each month, so that the rate of consumer inflation can be closely monitored. The CPI is also a timely statistic that is released very soon after the end of the period to which it refers.

- The CPI is a statistic with a long history, and people have been familiar with it over time.

- The CPI may be a relatively reliable price index compared with the price indices for some other flows. Although price changes for certain kinds of consumer goods are difficult to measure because of quality changes, price changes for other kinds of goods and services, such as capital goods and government services, especially public services, tend to be even more difficult to measure.

- Most countries have deliberately adopted a policy of not revising the CPI once it has been published. This makes it more attractive for many purposes, especially those with financial consequences such as indexation. The lack of revisions may perhaps create a somewhat spurious impression of certainty, but it also seems to enhance the credibility and acceptability of the index.

Because of the importance of the CPI to the multitude of users, the ECCU statistical offices are making every effort to improve the accuracy and relevance of their CPIs. The following sections provide an overview of the various procedures and methods that are being implemented during the CPI revision process including the use of PIPS as the compilation engine.

The CPI Processes

Developing the CPI market basket

The CPI market basket is developed from detailed expenditure information provided by families and individuals on what they actually bought. For the rebased CPIs, this information was collected from the HIES occurring during 2006 and 2008, depending on the country. During the HIES period, a sample of families from around the country provided information about their spending habits for major consumer goods and services in an interview survey. To collect information on frequently purchased items, such as food and personal care products, families also kept diaries listing everything they bought during a 2-week period.

Over the period covered by the HIES, expenditure information came from samples that covered approximately 8-10 percent of the households in the country. As noted, the sample

included weekly diaries and quarterly interviews which were combined to determine the importance, or weight, of the item categories in the CPI structure.

The CPI represents all goods and services purchased for consumption by households. The statistical offices have classified all expenditure items into more than 200 categories, according to the United Nation's Classification of Individual Consumption by Purpose (COICOP), the international standard. [4] The categories are arranged into the 12 Divisions below and examples of categories in each are also given:

COICOP Division	Category examples
01 FOOD AND NON-ALCOHOLIC BEVERAGES	breakfast cereal, milk, coffee, chicken
02 ALCOHOLIC BEVERAGES AND TOBACCO	wine, beer, cigarettes
03 CLOTHING AND FOOTWEAR	men's shirts and pants, women's dresses and blouses, shoes
04 HOUSING, WATER, ELECTRICITY, GAS, AND OTHER FUELS	rent of primary residence, owners' equivalent rent, electricity
05 FURNISHINGS AND HOUSEHOLD SUPPLIES	bedroom and living room furniture, detergents, cleaning supplies
06 HEALTH	prescription drugs and medical supplies, physicians' services, eyeglasses and eye care, hospital services
07 TRANSPORT	new vehicles, airline fares, gasoline, motor vehicle insurance
08 COMMUNICATION	postage, telephone services, cellular phones and service
09 RECREATION AND CULTURE	televisions, computers, software and accessories toys, pets and pet products, sports equipment, admissions
10 EDUCATION	primary, secondary, and college tuition and fees
11 RESTAURANTS AND HOTELS	full service meals at restaurants and cafes, snacks, lodging

[4] The actual coding manuals include several thousand codes at a 7-digit level. The item categories used are those that were the most important at the 7-digit level in food and beverages and at the 5-digit level for nonfood items.

COICOP Division	Category examples
12 MISCELLANEOUS GOODS AND SERVICES	haircuts and other personal services, cosmetics, watches and jewelry, legal and funeral expenses

Also included within these groupings are various government-charged user fees, such as water and sewerage charges, auto registration fees, and vehicle inspections. In addition, the CPI includes taxes (such as value added and excise taxes) that are directly associated with the prices of specific goods and services. However, the CPI excludes taxes (such as income and Social Security taxes) not directly associated with the purchase of consumer goods and services.

The CPI does not include investment items, such as stocks, bonds, real estate, and life insurance. (These items relate to savings and not to consumption expenses.)

The Appendix provides a table with an example from an ECCU country comparing the old market basket and the new basket on the COICOP classification structure. The introduction of imputed rents for owner-occupied housing results in a large increase in the weight for the Housing section of Division 04.

Price collection

The process by which the CPI measures price changes to consumers each month requires the efforts of many staff and the cooperation of individuals in households and retail outlets throughout the country. The prices are usually collected during the first or second week of the month when enumerators make visits to gather price information from selected supermarkets, department stores, service stations, doctors' offices, rental units, etc.[5] For the entire month, thousands of prices are recorded from areas throughout the countries. During each monthly visit, the enumerators collect price data on specific goods or services with precisely defined qualities or characteristics. If the selected item is available, the enumerator records its price. If the item is not available or if there have been any changes in the quality of the item since the last time the price was collected, the enumerator selects a new item or records the change in features that affect quality. By collecting price data on a clearly defined market basket of products that consumers purchase for their day-to-day living, the statistical office ensures that the CPI will provide an accurate measure of price changes.

Statistical offices do not have resources to price every good or service in every retail outlet in every area of the country, nor would it be an efficient use of resources. A sample of shops

[5] The one exception to monthly price collection in the ECCU is Anguilla, where prices are collected in the middle month of the quarter, i.e., February, May, August, and November.

normally visited by households is used to make the CPI representative of the prices paid for all consumer goods and services. In fact, the CPI samples are developed largely from data reported by households in the HIES on the places where they frequently make purchases. The specific products surveyed in each store represents both those identified by households as purchased there and those products identified by the outlets themselves as most popular in terms of sales.

Data Review

Once the enumerators have collected the prices in a given month, they are entered in the CPI database. One advantage of using PIPS is that it provides statistical routines that can be employed to identify data that are potentially outside the bounds of statistical expectations. (These are discussed in more detail in the next section.) The unusual prices are checked for accuracy and consistency. Respondents may be contacted/revisited to provide explanations for the unusual changes. Any necessary corrections or adjustments are then made.

Next, the PIPS calculates weighted indexes and percentage changes for further review by analysts that show price changes for each category of item, each local area, and the country as a whole. The statistical officers analyze the data and prepare the written release that is provided to other government agencies, the media, and the public. The entire process of reviewing, analyzing, and publishing the data is finished in about 4 weeks after the last data are collected.

The following section discusses the detailed methods, procedures, and formulas for calculating the CPI. The ECCU countries will be using the PIPS to compile the CPI as described here.

Notes on Calculating the CPI[6]

Many countries use the standard Laspeyres-type arithmetic mean of price relatives to compile their consumer price indices. Although the term ”Laspeyres” is often used to describe the formula, three points should be noted. First, the Laspeyres index requires that the weights reference period is the same as the price reference period, which is generally not the case. The weights are from a previous survey period since it takes some time to compile the weights for use with a price reference period of, say, January 2010. As noted below, the resulting index may more formally be a Young index or a Lowe index, as would be the case if the weights are updated for price change between the weight reference and price reference periods. The term “Laspeyres” or “Laspeyres-type” is used hereafter with this in mind. Second, the CPI is generally considered as being compiled in two stages, the

[6] This section is largely derived from the Price Index Processor System’s “CPI Compilation System User Manual” (August 2009) prepared by Paul Armknecht and Gangti Zhu. It is available on the Internet at http://www.unece.org/stats/downloads/SW_CPI_PPI/pips.html.

elementary level using an equally-weighted geometric mean (Jevons) index and the weighted higher level using a geometric Young index. Third, a modified/two-stage formulation is used, as outlined below.

Elementary index formulas

The *CPI Manual*, 2004 (Chapters 1 and 20) favors the use of the geometric mean formula (Jevons index) on axiomatic grounds. The Jevons index is defined as following:

$$(1) \quad P_J = \prod_{i=1}^{N} \left(\frac{p_i^t}{p_i^0} \right)^{1/N}$$

The arithmetic mean of price relatives (Carli index) is biased, especially in a chained form, and the ratio of arithmetic means of prices (Dutot index), which was previously used in many ECCU CPIs, is only suitable for strictly homogeneous varieties. Also see Silver and Heravi (2007) for a more detailed discussion on these differences.

The Standard Laspeyres-type formula

The standard Laspeyres-type formula is generally used for aggregating the CPI from the most detailed level of the basket at which item weights are available up to the all items level. It compared the current period cost of the market basket (the numerator) with the cost of the market basket in the base period (the denominator) and can be written as:

$$(2) \quad I_{0 \to t} = \frac{\sum_{i=1}^{N} q_i^0 p_i^t}{\sum_{i=1}^{N} q_i^0 p_i^0} = \sum_{i=1}^{N} \left[\frac{q_i^0 p_i^0}{\sum_{i=1}^{N} q_i^0 p_i^0} \times \left(\frac{p_i^t}{p_i^0} \right) \right]$$

where *i = 1, ... , N* stands for the products comprising the consumption basket and symbols 0*,t* respectively designate the price reference period (or the base price period and the current price period). The symbols *p* and *q* designate the prices and the quantities of the products in question, respectively. The ratio (p_i^t / p_i^0) is the price relative from the base period for item *i* (sometimes called the long-term price relative).

By expressing the consumption expenditure's share for the item *i* as a ratio of the total expenditure during the base period, we obtain the following weight:

$$(3) \qquad w_i^0 = \frac{q_i^0 p_i^0}{\sum_{i=1}^{N} q_i^0 p_i^0}$$

Using the preceding expression, formula (2) can be written in a slightly different form as:

$$(4) \qquad I_{t\to 0} = \sum_{i=1}^{N} w_i^0 \left(\frac{p_i^t}{p_i^0} \right)$$

However, these versions of the Laspeyres formula do not provide the flexibility required for economies that are going through significant and rapid changes with new or improved products continuously being introduced.

The modified or two-stage Laspeyres approach

There are several reasons why the Modified Laspeyres Approach is superior to the standard formula. First, in the standard formula, we are calculating price relatives for the current period from the base period. In practice, the editing of the current period's price data is done by comparing the prices for the current collection period for an item with those for the same item in the previous period. Any large variations falling outside a predetermined range check (e.g., 0.8000 to 1.1000) might indicate either the wrong item has been priced or some kind of error has been made in recording the price. With formula (4), this comparison cannot easily be made as it uses, for each item *i*, the price relatives of current period to the price reference period $\left(\frac{p_i^t}{p_i^0} \right)$.

Second, the standard formula involves a comparison of changes in prices for each item over long time periods, requiring the continuity of priced item specifications. In practice, varieties become permanently missing or unrepresentative and need to be replaced with new varieties for which there is no price in the reference period 0 to compare with. In these circumstances, it is advisable to apply a modified version of the Laspeyres formula that makes use of $\left(\frac{p_i^t}{p_i^{t-1}} \right)$, the price relative from the previous period so that a new variety can be introduced as soon as two successive price quotes are available.[7]

[7] A new replacement item may be introduced using long-run price relatives by imputing its price in the base period using the price change of the product group to which it belongs. For example, if a flat-screen television set was introduced say in January 2008 to replace the previous model in the CPI for an index whose base was 2000, then the price of the flat-screen television set in 2000 may be imputed on the basis of prices changes of all electronic goods between 2000 and January 2008. This requires a continuing heroic assumption about price changes. The two-stage approach will use actual price changes of the old set up to January 2008 and of the new set afterwards, with a one-off

(continued)

Third, when varieties are temporarily missing, imputed prices may be used based on the overall price change of the product group in question. Imputations over the short run are preferred over long-run imputations because the long-term relatives may not be reflective of the same price history as the missing variety.

The basic formula for computing the Laspeyres index can be written as:

$$(5) \quad I_{0\to t} = \frac{\sum_{i=1}^{N}\left(\frac{p_i^t}{p_i^{t-1}}\right)\times q_i^0\, p_i^{t-1}}{\sum_{i=1}^{N} q_i^0\, p_i^0} \times 100$$

where $p_{t-1,i}q_{0,i} = q_i^0\, p_i^0 \times \frac{p_i^1}{p_i^0} \times \frac{p_i^2}{p_i^1} \times \ldots \times \frac{p_i^{t-11}}{p_i^{t-2}}$

Formula (5), which is arithmetically equivalent to formulas (2) and (4), is considered more versatile than the formula using a long-term price relative from the base period, as the linking process used facilitates the introduction of new varieties and/or items or substitution when the need arises and enables more reasonable imputations.

Formula (5) can also be rewritten as:

$$(6) \quad I_{0\to t} = \sum_{i=1}^{N} w_i^0 \times \left(\frac{p_i^t}{p_i^{t-1}}\right) \times \left(\frac{p_i^{t-1}}{p_i^0}\right)$$

which can be interpreted as:

$$(7) \quad I_{0\to t} = \sum_{i=1}^{N} w_i^{*t-1} \times \frac{p_i^t}{p_i^{t-1}}$$

where $w*_i^{t-1} = w_i^0 \times \frac{p_i^{t-1}}{p_i^0}$ is an updated weight sometimes referred to as a "cost weight" of item *i*.

---

assumption only required: that the price difference between the two in January 2008 reflects their quality.

In other words, to obtain the index for the current period *t*, the Modified Laspeyres Approach involves multiplying individual price relatives of the latest price compared period ($\frac{p_i^t}{p_i^{t-1}}$) by the previous period's updated weight ($w *_i^{t-1}$), and then summing them.[8]

The Modified Laspeyres formula has obvious advantages over the standard Laspeyres formula when we consider the problems arising from permanently unobservable varieties, and the need in due course to bring in a new variety to replace the missing one. There is a need, as explained in footnote 6, to impute a base period price if the standard Laspeyres formula is used. Such imputation is unnecessary while using the Modified Laspeyres formula, in which case the current period weight for the replacement item is obtained by simply multiplying the last updated weight for the replaced item by the current period's short-term price relative of the replacement item.

The computation system uses the two-stage approach to calculate the CPI based on monthly price quotations and weights information. The price index is assigned a value of 100 in the price reference period and the value of the index for other periods, after subtracting unity, represents the average percentage change in price levels from this reference period.

Instead of holding the expenditure weight reference period at 0, the CPI can be compiled as a weighted geometric/arithmetic average of the individual price relatives holding constant the expenditure shares at period *b*. The resulting index is called a *Young* index. In this case, the formula should be revised as:

$$(8) \quad I_{0\to t} = \sum_{i=1}^{N} w_i^b \times \left(\frac{p_i^t}{p_i^{t-1}}\right) \times \left(\frac{p_i^{t-1}}{p_i^0}\right)$$

The weight reference period *b* is likely to precede the price reference period 0 because it takes time to collect and process the expenditure data. For example, the monthly CPI will run from January 2010 onwards, with January 2010=100, but the expenditure values may have been derived from the Household Income and Expenditure Survey (HIES) taken during a 2006 to 2008 period (depending on the country).

[8] Formula (6) can also be interpreted as:

$$I_{0\to t} = \sum_{i=1}^{n} w_i^0 \times STPR_{t\to t\text{-}1,i} \times LTPR_{t\text{-}1\to 0,i}$$

where $STPR_{t\to t\text{-}1,i}$ is the short-term price relative of item *i* for current period (= p_i^t / p_i^{t-1}) and $LTPR_{t\text{-}1\to 0,i}$ is the long-term price relative of item *i* for the previous period (= p_i^t / p_i^0).

In such a case, we have the choice of assuming that *either* the expenditure shares in period *b* remain constant, as in equation (8), *or* the quantities of period *b* remain constant and the weights are updated for price changes between the weight reference period and the price reference (base) period. In the latter instance, the index form is a Lowe index.

The geometric Laspeyres or Young indices

The *CPI Manual* notes that there are three target indexes—Fisher, Törnqvist, and Walsh—which are, in principle, the least biased and most appropriate indexes for compiling a CPI. It recommends that, if possible, statistical offices should strive to produce one of these measures. Each of these indexes, however, use symmetric averages of weights from the current and weight reference periods and include a geometric averaging component.[9] Most statistical offices cannot produce these types of indexes on a monthly basis.[10]

One problem with fixed based indexes like the Laspeyres and Lowe indexes is that they do not allow for any substitution in what consumers purchase through time. They assume that consumers purchase the same fixed quantities of goods and services as in the past. In fact, we know that consumers change their purchasing patterns in response to relative changes in prices. The traditional Laspeyres-type indexes tend to drift upward over time versus the target indexes because the fixed quantities tend to give more importance to those items that have the larger relative price changes and less importance to those that have lower than average price changes.

A realistic solution to this problem is to use a formula that allows for substitution by consumers similar to that of the target indexes. Such a formula is the geometric mean aggregation formula that can easily be applied to the Laspeyres and Young indexes by converting them from using arithmetic averages to using geometric averages.[11]

The use of the geometric formula is also more consistent with the Jevons index at the elementary level than a Laspeyres-type arithmetic aggregation. Geometric means are (i) not as sensitive as arithmetic means to the extreme values, (ii) are circular, i.e., fulfill a multi-period transitivity property; that is: the product of the price index change going from period 1 to period 2 *times* the price index change going from period 2 to period 3 should equal the

[9] Törnqvist uses a geometric aggregator of price relatives weighted by an arithmetic average reference and current period expenditure shares; Fisher is the geometric mean of an harmonic and arithmetic aggregator of respective reference and current period weighted averages of price relatives; and Walsh uses a geometric mean of reference and current period quantities for an arithmetic fixed basket comparison.

[10] Some countries such as the United States and Sweden produce one of these indexes regularly, but it is subject to revision when the most recent set of weights become available. See, for example, http://www.bls.gov/cpi/cpisuptn.htm for a note on the U.S. Törnqvist version and Ribe (2004) for the Swede's Walsh version.

[11] The *CPI Manual* (paragraph 1.40) suggests that statistical offices should treat the use of one of these indexes as a serious practical possibility.

price index going directly from period 1 to 3; and (iii) are more likely to lie between the Laspeyres and Paasche bounds, a desirable property discussed in the *CPI Manual* (see paragraphs 1.91–1.96).

In the geometric version of the modified Laspeyres index, a weighted geometric average is taken of the price relatives using the expenditure shares of period 0 as weights. It is defined as:

$$(11) \quad I_{0\to t} = \prod_{i=0}^{N}\left[\left(\frac{p_i^{t-1}}{p_i^0}\right)\times\left(\frac{p_i^t}{p_i^{t-1}}\right)\right]^{w_i^0}$$

Similarly, there is a geometric version of the *Young* Index if period $b \neq 0$, that is, the expenditure shares are different from the price reference period 0:

$$(12) \quad I_{0\to t} = \prod_{i=0}^{N}\left[\left(\frac{p_i^{t-1}}{p_i^0}\right)\times\left(\frac{p_i^t}{p_i^{t-1}}\right)\right]^{w_i^b}$$

Because the weights used in the CPI are taken directly from the recent HIES conducted sometime between 2006 and 2008 (depending on the specific country) and the base price reference period is for January 2010, the Geometric Young index is used. Balk (2009) discusses the relationship between this index (also referred to as the Cobb-Douglas price index) and the hypothetical Konüs cost-of-living index.

Matched Price Observations and Imputing Missing Prices

A price relative is calculated for each variety comprising the CPI basket. The calculation of price relatives would be simple if a complete set of price quotations were available for the current and previous month. In reality, this does not always happen. Quite often, some of the respondents are unable to provide a price for a particular variety because it is out of stock. Whenever a particular price observation is missing from either the previous month or the current month, the corresponding price observations are eliminated from the other period. This is equivalent to imputing the price of variety 1 in period *t* by the short-run price change of the other varieties in the product group. This ensures that the price relatives are calculated on the basis of "matched observations", i.e., a consistent sample of price quotations in each period.

In the following example we consider that item's prices are collected for four representative varieties *1*,*2*,*3*, and *4*. In the current month *variety 1*'s price cannot be collected (is missing).

Variety *v* of item *I*	Month *t-1*	Month *t*	Price relative in *t*
Variety 1	1.50	-	-
Variety 2	1.25	1.25	1.0000
Variety 3	1.25	1.50	1.2000
Variety 4	1.50	1.50	1.0000
Geometric average price of all varieties	**1.3693**	**1.4116**	
Geometric average price relative for matched observations (Variety 2, Variety 3, Variety 4)	**-**	**-**	**1.0627**
Geometric average of matched prices	**1.3283**	**1.4116**	
Short-term relative for item *i* using matched sample average prices	**1.0627 (= 1.4116/1.3283)**		

For the CPI calculation of month *t*, the geometric average price relative of 1.0627 should be calculated based on matched observations as $(1.0000*1.2000*1.0000)^{(1/3)}$ = ***1.0627***.

The month's short-term price relative for item *i* is then used to impute the missing price for Variety 1: 1.0627 x 1.5 = 1.594.

Imputing Missing Indexes

A missing price index is estimated using its parent index as the proxy, i.e., if a specific product's index in an outlet is missing due to the missing prices, the index of the product or item it belongs to will be taken to be representative. For higher levels, the next available index data in the same group or item for the missing index will be used. Holding missing prices for a variety constant by carrying the last observation forward (i.e., making the short-term price relative for that variety equal to 1.0) during a period of high inflation would cause short-term distortions in the index. This is because it would understate inflation while the variety was unavailable and then show a large increase in the index when the variety became available.[12] Further, we do not use the price for the same variety in another outlet to represent the missing price in this outlet. Thus if the Coca-Cola price is missing in one

[12] This is particularly important for seasonal items for which the spike in the price change on the return of the item when back in season might be very large and undermine the credibility of the index. The carry forward method also distorts price change measurement for items that do not return, but are replaced by comparable items (or non-comparable ones with quality adjustments). The overall effect is to introduce undue stability into the CPI.

particular outlet, we will not take another outlet's Coca-Cola price as a proxy. Instead we will take the soft drink index in the same outlet to represent that of Coca-Cola.

Since parent-group price changes are always calculated as geometric mean changes, imputations are based on geometric means.

If no price is collected for any variety covered by a product (the prices for whole product are missing), its price relatives will be imputed using average price relatives from the item group of the missing price.

Detection of Outliers

What is an outlier

An observation that is unusually large or small relative to the other values in a price relatives data set is called an outlier. Outliers are the observations that appear to be inconsistent with the remainder of the collected data.

There are several possible sources for outliers:

1. The price quotation of a transaction or variety is observed, recorded, or entered into the computer incorrectly.

2. The price quotation comes from a different population, or the quality of that transaction/variety has changed.

3. The price quotation entered is correct, but represents a rare event or novel phenomenon.

Outliers occur when the relative frequency distribution of the data set is extremely skewed. Such distributions have a tendency to include extremely large or small observations.

Two procedures statistical office can use to identify the possible errors and outliers are outlined. The first one is a non-statistical procedure, which is to find whether a specific price observation falls outside some pre-specified acceptance interval. Generally, a 20% threshold is used to identify large and unusual changes when prices are first entered. The second procedure is a statistical measure in which three methods are implemented. In both cases outlier detection will *not* result in automatic deletion. Often price changes are undertaken after some time and the "pent-up" prices changes are unusually large. To delete them would bias the index downwards. The outlier detection is to alert the compiler about a possible error that needs further investigation.

Three approaches for identifying outliers in the current period's prices are outlined below.

Method of using z-scores

In a z-score test, the mean and standard deviation of the entire data set are used to obtain a z-score for each data point, according to following formula:

$$Z_i = \frac{(x_i - \bar{x})}{s}$$

Where $\bar{x}$ is the arithmetic mean and s the standard deviation, i.e.

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

If the observations have a bell shaped distribution (normal distribution), the interval from $\bar{x} - s$ to $\bar{x} + s$ will contain approximately 68% of the measurements; the interval from $\bar{x} - 2s$ to $\bar{x} + 2s$ will contain approximately 95% of the measurements, and the interval from $\bar{x} - 3s$ to $\bar{x} + 3s$ will contain approximately all of the measurements.

In the case of price relatives, we do not know the underlying distribution of data set. Many studies of price change show that price relatives are not normally distributed. Thus we appeal to Chebyshev's theorem which applies to all possible distributions. According to Chebyshev's theorem, for any set of measurements and any number $k \geq 1$, the interval from $\bar{x} - s$ to $\bar{x} + s$ will contain at least (1 - 1/k2)*100 percent of the measurements.

Thus, at least 88.8 percent of all the observations in a data set will have z-score less than 3 in absolute value i.e. fall into the interval $(\bar{x} - 3s, \bar{x} + 3s)$, and at least 75 percent will fall within 2 standard deviations, where $\bar{x}$ is the mean and *s* is the standard deviation of the sample. Therefore, the observations with z-score greater than 3 will be potential outliers.

The Z-score method may be biased by the problem that both the mean and standard deviation are affected by the outliers.

Box Plots Method

Another procedure for detecting outliers is to construct box plots of the price relatives data. They make no distributional assumptions and, since they rely on the median and quartiles as parameters, the method of detection is not influenced by the outliers themselves. Below are the steps implemented in constructing the box plots for the software.

- The median M, lower and upper quartiles, QL and QU, and the inter-quartile range, IQR= QU - QL are calculated for the data set.

- Two sets of limits on the box plot are constructed: inner fences are located a distance below QL and above QU; outer fences are located a distance of below QL and above QU.

Observations that fall between the inner and outer fences are called "suspect" outliers.

Log-normal Method

Another method that is used to identify the possible errors and outliers is to use 2 standard deviations from the log-normal distribution, excluding price relatives of 100 (no change of prices from previous to current period).[13] It takes the natural logarithms of price relatives data, which is assumed log-normally distributed. The standard deviation and mean of the log of all price relatives in the sample are calculated. Those price relatives that fall outside of 2 standard deviations (with 95% confidence level) are considered as possible outliers.

Summary

The ECCU countries are going through the process of developing new market baskets for the CPI and introducing them. At the same time they are making improvements to the classification structure, the methods, and procedures used to compile the CPI in line with those discussed in the *CPI Manual*. This process has taken several years and is nearing completion. Conducting the HIES put some strain on the statistical office resources, but revised CPIs will make the measurement of inflation more accurate and relevant.

In completing the CPI revisions, the ECCU countries have received substantial technical assistance from CARTAC to help with the process. The assistance has included working with each country to introduce the PIPS software. PIPS will enable them in implementing many of the improved methods described in the *CPI Manual*.

The paper provides a brief overview of the new methods and procedures that are being implemented with PIPS at this time. These represent major enhancements to current practices in the region. As the project comes to completion, these countries will be among the first wave to upgrade their CPI to the new international standards reflected in the *CPI Manual*.

[13] The reason to exclude those price relatives of 100 is that there often are many "no change" price relatives. By including them we will observe a bimodal or Saddle distribution. Such a distribution is not normally distributed we cannot apply a 2-sigma limit with 95% confidence to detect outliers.

References

Armknecht, P. and G. Zhu, (August 2009) "CPI Compilation System User Manual," available on the Internet at http://www.unece.org/stats/downloads/SW_CPI_PPI/pips.html.

Balk, B. M., 2009, "Lowe and Cobb-Douglas CPIs and their Substitution Bias," Proceedings of the Eleventh Meeting of the International Working Group on Price Indices (Neuchâtel, Switzerland). Available on the Internet from www.ottawagroup.org.

Balk, B. M. and W. E. Diewert, 2003, "The Lowe Consumer Price Index and its Substitution Bias," Discussion Paper No. 04-07 (Department of Economics, University of British Columbia, Vancouver). Available on the Internet from http://econ.arts.ubc.ca/diewert/hmpgdie.htm

Caribbean Development Bank, 2008, *Country Poverty Assessment Reports*, various countries, available on the Internet at http://www.caribank.org/titanweb/cdb/webcms.nsf/75f7ba2c7557c2e50425745900719b7a!OpenView

Caribbean Regional Technical Assistance Centre, 2009, "Work Plan for November 2009–April 2010, National Accounts and Price Statistics," pp. 36–54. Available on the Internet at http://www.cartac.org/.

International Labor Organisation (ILO), International Monetary Fund, Organization for Economic Cooperation and Development, Statistical Office of the European Community, United Nations Economic Commission for Europe, and World Bank, 2004, *The Consumer Price Index Manual: Theory and Practice*, (Geneva, ILO). Available on the Internet at http://www.ilo.org/public/english/bureau/stat/guides/cpi/index.htm

Ribe, M., 2004, "Consumer Price Index (CPI) for Sweden: Improved CPI construction from January 2005, Technical description," Statistics Sweden, Pm11444.doc.

Silver, M. and S. Heravi, 2007, "Why Elementary Index Number Formula Differ: Price Dispersion and Product Heterogeneity," *Journal of Econometrics*, Vol. 140 (October), pp. 874–83.

U. S. Bureau of Labor Statistics, 2003, Note on a New Supplemental Index of Consumer Price Change, available on the internet at http://www.bls.gov/cpi/superlink.htm

Appendix

Comparison of the Current and Revised CPI Basket on the COICOP Structure (Country Example)

Item Code	Item	Current Weights	Revised Weights	Change
	All Items--COICOP	1000.00	1000.0	
01	**FOOD AND NON-ALCOHOLIC BEVERAGES**	**536.10**	**219.1**	**-317.00**
01.1.1	***Bread and Cereals***	138.70	44.1	-94.60
01.1.2	***Meat***	70.60	43.2	-27.40
01.1.3	***Fish***	46.60	15.9	-30.70
01.1.4	***Dairy Products, Cheese and Eggs***	48.50	26.0	-22.50
01.1.5	***Oils and Fats***	41.80	8.8	-33.00
01.1.6	***Fruits***	27.10	6.8	-20.30
01.1.7	***Vegetables***	96.70	30.3	-66.40
01.1.8	***Sugar and Confectioneries***	47.20	12.9	-34.30
01.1.9	***Other Foods***	4.80	10.4	5.60
01.2.2	***Coffee, Tea and Cocoa***	11.90	3.5	-8.40
01.2.1	***Soft Drinks, Fruit & vegetable juice***	2.20	17.2	15.00
02	**ALCOHOLIC BEVERAGES AND TOBACCO**	**9.50**	**38.7**	**29.20**
02.1	***Alcoholic Beverages***	7.20	36.8	29.60
02.2	***Tobacco***	2.30	1.9	-0.40
03	**CLOTHING AND FOOTWEAR**	**88.70**	**32.2**	**-56.50**
03.1.2	***Clothing***	54.40	22.2	-32.20
03.1.1	***Textile Materials***	16.60	1.8	-14.80
03.1.4	***Tailoring and Dressmaking***	3.70	-	
03.2	***Footwear***	14.50	8.3	-6.20
04	**HOUSING, WATER, ELECTRICITY, GAS**	**127.70**	**300.6**	**172.94**
	AND OTHER FUELS			
04.1.1	***House Rent****	55.70	192.6	136.90
04.9.1	***House Tax***	20.20	-	
04.9.2	***House Insurance***	7.30	-	
04.4.1	***Water Consumption***	9.80	13.1	3.30
04.3	***House Repairs and decoration***	4.90	25.4	20.50
04.5	***FUEL AND LIGHT***	29.80	69.5	39.70
05	**FURNISHINGS, HOUSEHOLD EQUIPMENT**	**61.50**	**65.9**	**4.40**
	AND ROUTINE HOUSEHOLD MAINTENANCE			
05.1.1a	***Furniture and furnishings***	27.40	18.5	-8.90
05.4	***Glassware, Tableware & Household Utensils***	7.20	2.0	-5.20

Comparison of the Current and Revised CPI Basket on the COICOP Structure (Country Example)

Item Code	Item	Current Weights	Revised Weights	Change
05.3	*Household Appliances*	13.80	11.9	-1.90
05.6.1a	*Soaps and Detergents*	9.30	17.8	8.50
05.6.1b	*Other Household Supplies*	2.50	1.2	-1.30
05.6.2	*Domestic Servants*	1.30	14.5	13.20
06	HEALTH	10.10	17.9	7.80
07	TRANSPORT	44.00	118.4	74.40
08	COMMUNICATION	24.50	94.1	69.60
09	RECREATION AND CULTURE--COICOP	42.20	38.1	-4.10
10	EDUCATION--COICOP	8.00	13.2	5.20
11	RESTAURANTS AND HOTELS	13.7	18.7	5.00
11.1	Meals at restaurants, cafes & the like	13.70	18.3	4.60
12	MISCELLANEOUS GOODS & SERVICES--COICOP	34.00	43.1	9.10

** The large difference in rent is the result of including imputed rents for owner occupied housing. Previously, owner-occupied housing was not included in the CPI.