Distr.
GENERAL

CES/AC.49/2003/13*
7 October 2003

ENGLISH ONLY

STATISTICAL COMMISSION and UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE (UNECE)

INTERNATIONAL LABOUR ORGANISATION (ILO)

CONFERENCE OF EUROPEAN STATISTICIANS

Joint UNECE/ILO Meeting on Consumer Price Indices
(Geneva, 4-5 December 2003)

WEIGHT UPDATES AND REVISIONS IN THE CPI

Invited paper submitted by the International Monetary Fund (IMF)**

I. INTRODUCTION

1. Periodically, statistical agencies must revise CPI weights yet, traditionally, CPIs are released as non-revised series. The often-heard argument for this is that some important users, such as government social security funds and parties to labor agreements, escalate contracts with the CPI and revisions to the index would complicate the administration of these contracts. CPIs thus tend to use fixed basket formulas with periodic item and weight revisions that are "linked into" the official series sometime after the reference period of the weights. This obviates the need to revise data between the reference period of the weights, the natural linking period, and the month when the revised series actually supplants the old series.

---

* Due to the late submission, this paper could not be translated.

** Prepared by Mr. Kim Zieschang, Statistics Department, International Monetary Fund. The views and opinions contained in this paper are the author's and do not necessarily represent the views and opinions of the IMF.

2. There always is such an interval between the reference period of the weights and the earliest feasible link period, because it takes time to process the data for revised index weights. These data typically come from household expenditure surveys. The collection and processing lags of these surveys range from about six months to two or three years or more.[1]

3. Practical CPI procedure had fit comfortably within these seemingly immutable practical constraints, until the advent of the Chained CPI for All Urban Consumers (C-CPI-U) issued by the US Bureau of Labor Statistics beginning in August 2002. The US C-CPI-U, which is published as a supplement to the official Laspeyres index, is the product of research undertaken by the BLS initially from 1989-1993, and issued in the December 1993 *Monthly Labor Review*, on so-called *superlative* price index formulas. Superlative formulas use weighting information from the current period and a recent past period in the context of periodic index chaining. Superlative indices closely approximate *economic* index numbers that reflect households' proclivity to substitute away from items becoming relatively more expensive, other things—such as disposable income, environmental amenities, and tastes and preferences—equal. The traditional, Laspeyres-type formulas of CPIs prior to the C-CPI-U are able to reflect these substitution effects only partially through periodic chaining.[2] Yet any upward errors from their incomplete reflection of substitution effects tend to be propagated through chaining, limiting the safe frequency with which they can be chained.

4. The impetus for creating the C-CPI-U grew out of public discussion of the impact on indexed items in the central government budget of the upward substitution bias in the Laspeyres formula used by the then, as well as current, official US CPI. (The discourse also included other, perhaps even more important sources of bias, such as quality change.) This discussion was largely instigated by the aforementioned late 1993 special issue of the *Monthly Labor Review* in the context of a contemporaneous debate on what to do about large prospective central government deficits. Its outcome was the now well-known Boskin Commission Report of 1996. Thus, the interests of new and influential group of users, the Treasury and the budget committees of the legislature, finally have propelled superlative index numbers from an academic notion to a published statistical series.[3] In so doing, they have revived the moribund notion that an accurate CPI must be revisable, at least for some important purposes.

II. HOW THE C-CPI-U REVISION POLICY WORKS

5. It seems reasonable in drawing out some principles for a CPI revision policy to begin with practice in the US C-CPI-U.[4] The US C-CPI-U's target concept is a monthly chained Törnqvist index. The revision policy takes account of the lag in arrival of monthly expenditure weights from the US Continuing Consumer Expenditure Survey (CCEX) and focuses on using weights that produce the best current estimates for month to month Törnqvist "links" in the index. The index level over a period of interest is the product of these monthly links from the beginning to the end of the period.

6. When first released, the *initial* C-CPI-U link for a given month uses the most recent available biannual expenditure weights available in the official Laspeyres index. At this writing in September 2003, the most recent available weights to average the short term price

relatives from August to September were for 1999-2000 (the same currently used in the official Laspeyres index).

7. When a new year of weights becomes available in January 2004, the *interim* C-CPI-U August-September index link will use the average of the share weights from 2001-2002 (the same used in the official Laspeyres index at that time).

8. The *final* August-September 2003 link is compiled when weights become available for the month of September 2003 in January 2005, in which case the weights for the August-September link are the average of the monthly weights for August and September 2003. Thus, in the end the August-September link follows the Törnqvist formula exactly.

9. The index level for September 2003 relative to, say, December 1999 is the product of the *final* monthly links up through the month for which monthly consumer expenditure data are available, multiplied by a sequence of *interim* links for which approximately two-year old biannual weights are available (relative to the month of the link), multiplied by a sequence of *initial* links for which approximately three-year old biannual weights are available (relative to the month of the link).

III. SOME PRINCIPLES IN DESIGNING A CPI REVISION POLICY FOR SUPERLATIVE INDICES

10. The following principles seem to underlie the C-CPI-U process:

- Provide relatively few revisions to any given month (BLS chooses 2 revisions from the *initial* release);
- To produce accurate preliminary estimates, consider progressively centered annual weight indices, which may approximate superlative indices. For its preliminary estimates BLS uses biannual weights centered approximately on December-January between two years;[5]
- To minimize revisions, use a formula relatively robust to moderate variations in the weights (BLS chooses the geometric family, comprising the weighted Jevons or Cobb-Douglas formula in its *initial* and *interim* estimates and the Törnqvist formula in its *final* estimates);
- To produce accurate final estimates with minimal "drift" under high frequency (monthly) chaining, use a *superlative* formula as the final result (BLS chooses Törnqvist superlative formula in its *final* estimates).

IV. REVISABLE CPIS AND "CHAIN DRIFT"

11. Despite the use of superlative chain links, high frequency chaining can still raise concerns about excessive chain "drift" in high or variable inflation situations (not currently the case in the US).[6] Making the series revisable opens possibilities to deal with this issue, as well as substitution bias. For example, to limit drift further in high inflation epochs, it may be useful test the effectiveness of the Eltetö-Köves-Schultz (EKS) "transitivizing" algorithm

applied over moving 12 month intervals, in concert with producing some direct annual change CPI indices.[7] A temporally transitive time series of indices is by definition immune to drift, because direct price indices over two or more months must equal the chain of one month links between them.[8]

12. This said, drift is not necessarily a problem. Weight changes arising from the evolution of preferences, technology, and environmental amenities over time will introduce some, conceptually defensible drift in price indices. Drift in a chained index arising from measurement errors is, however, a problem. Such errors can come from normal sampling of prices. Precision in the weights is another, possibly more important consideration in deciding how frequently to update the weights. This is largely driven by the sample sizes of the source data for the weights; for example, the sample size of the household expenditure survey. Driving the sampling errors in these sources down by blending data from the household expenditure survey and retail sales surveys could help moderate this source of error. Another approach would be to apply EKS to the most recent 12 months of price data. The latter offers a way of suppressing drift arising over a seasonal cycle, which would be influenced most by sampling error, without eliminating "normal" drift arising from underlying economic factors changing over the longer term.

V. REVISABLE CPIS AND COHERENT SYSTEMS OF ECONOMIC STATISTICS

13. As a last observation, revisability helps maintain the coherence of consumer price statistics with the overall system of macroeconomic statistics. Some countries work to keep their CPI weights reconcilable with the national accounts household expenditure aggregates.[9] This practice has merit because it provides users with transparent consistency between macroeconomic value, price, and volume statistics. A revisable CPI provides a means for sustaining this consistency more closely than generally possible for the non-revisable series.

14. National accountants usually combine retail sales survey with household expenditure survey data sources to provide a timely best estimate at annual and quarterly frequency of both the level and broad composition of household final expenditure. Retail sales survey data, while not as detailed for goods and services, provide more timely information than household expenditure surveys, usually at monthly frequency. They also are arguably more reliable in estimating the overall level of final consumption expenditure than the household expenditure surveys, whose principal strength is in product (and in most cases geographical) detail. Many countries produce these current price household expenditure aggregates quarterly. Depending on the national accounts revisions cycle, alignment of the CPI weights with quarterly national accounts might well produce revisions in a given month's data every three months until monthly information were available from the combined retail and household expenditure sources. Hence the first principle of a CPI revision policy that there should be a limited number of revisions would be difficult to support in such a highly coherent revisable CPI. The frequency of revisions might well be tempered here by how large a revision is introduced by a given quarterly weight update. If quarterly CPI weight updates produce trivial changes, annual revisions of quarterly weights might do just fine.

NOTES

[1] For at least some expenditure components of the CPI, however, such as food and household items, we are seeing the beginnings of a drastic shortening of this lag, as more and more retail establishments use electronic point of sale (POS) equipment. POS equipment scans the labels of preclassified retail inventory and enters each transaction directly into a database. Linking household data through "loyalty cards" to POS transaction data, for example, provides a complete set of prices, quantities, and household characteristics on computer media for the establishments and households included in such an arrangement. See Guðnason and Snorrason, "Use of cash register data," paper presented at the 1999 Ottawa Group meetings [http://www4.statcan.ca/secure/english/ottawagroup/ottawa/gudnason.pdf]. The chief difficulty with POS data is that the commercial product classifications such as the Uniform Product Code (UPC) have too little code space to identify all the product varieties of significance for pricing, and individual retail companies are free to set the last few digits. This makes interfirm comparability something of a nightmare at a high enough level of product description detail. A nascent international classification with potential to cope with this problem is being developed for the 2005 round of the International Comparison of Prices (ICP) Program. The ICP Structured Product Descriptions (SPDs) are an ongoing internationalization of another BLS methodology, involving the "checklist" descriptions BLS uses to probability select products within retail establishments during a so-called "disaggregation" process. The ICP SPDs focus on the descriptive rather than probability selection aspects of this methodology, and are covered in an appendix of the newly issued *Consumer Price Index Manual*.

[2] However, notwithstanding near universal application of the Laspeyres formula from the level of elementary aggregates up to all items, statistical agencies may employ sample refreshment methods on a more or less systematic basis at the level of products, as well as non-Laspeyres elementary aggregation formulas, such as the Jevons or geometric formula. These measures can go a long way toward addressing the Laspeyres substitution bias problem at the product level, where much of this substitution effectively takes place.

[3] Still another but less influential group also advocating superlative indices have been economists and business analysts interested in productivity questions, and thus concerned about the downward biases on input volume indices resulting from deflation of aggregates by upward biased price indices.

[4] The most detailed reference on this is Cage, Greenlees, and Jackman, "Introducing the Chained Consumer Price Index," presented at the 2003 meetings of the Ottawa Group [http://www.insee.fr/en/nom_def_met/colloques/ottawa/ottawa_papers.htm]. Other documents summarizing this indicator are on the BLS website [http://www.bls.gov/cpi/superlink.htm].

[5] See Masato Okamoto, "Midpoint-year basket index as a practical approximation to superlative index," paper presented at the 2001 Ottawa Group meetings in Canberra, Australia (http://www4.statcan.ca/secure/english/ottawagroup/pdf/06/Midpoint-year%20basket%20index%20-%20Okamoto.pdf). This connection is heuristic only, since BLS cannot fully implement Okamoto's method because the biannual weights are not

sufficiently timely, but rather are progressively more timely between the initial and preliminary releases. By the time weights centered in the same year as the prices are available, BLS has moved to the final monthly weights and the monthly chained Törnqvist superlative formula and has no need of the Okamoto superlative approximation.

[6] On chain index drift, see Bohdan Schultz, 1983, "Linking Price Index Numbers," in W.E. Diewert and C. Montmarquette (eds.), Price Level Measurement, Ottawa: Statistics Canada, pp. 537-566.

[7] The EKS algorithm is used in the International Comparison of Prices Program and is well known in the context of international and interarea comparisons for which it was originally designed. The analog of transitivity in cross-sectional index comparisons is Irving Fisher's so-called "circular test" in temporal indices. EKS can impose the transitivity property in any context, including time series and fits especially well with the Törnqvist index number. See Caves, Christensen, and Diewert, 1982, "Multilateral Comparisons of Output, Input, and Productivity Using Superlative Index Numbers," *Economic Journal* 92:73-86, and Kokoski, Moulton, and Zieschang, 1999, "Interarea Price comparisons for Heterogeneous Goods and Several Levels of Commodity Aggregation," in *International and Interarea Comparisons of Income, Output, and Prices*, Heston and Lipsey, eds, Cambridge, Massachusetts: NBER, 123-170. The latter paper raises the possibility of time series applications in concluding remarks, pp. 141-142.

[8] This procedure would adjust the monthly index links so their products conform to direct indices for corresponding multi-month periods. By implication, at least some direct indices over multi-month periods would have to be compiled; for example, a series of indices comparing the current month price with the same month one year ago with a Törnqvist formula whose weights are the average of the current month expenditure shares with the corresponding shares of the same month of the previous year.

[9] The national accounts aggregates relevant for the CPI are the final consumption expenditure (SNA code P.31) or the total final expenditure (sum of SNA codes P.31 for final consumption and P.51 for fixed capital formation) of the household institutional sector (SNA code S.14). Most countries use a variant of final consumption excluding consumption employers provide as compensation in kind in the form of work related benefits. However, the scope of a few CPIs produced under "inflation measurement" concepts is household total final expenditure excluding employers' compensation in kind and the imputed rent of owner occupants.
