# A Note on Long Run Analytical CPI Series that Avoid Discontinuities on Rebasing

Kim Zieschang and Mick Silver[1]
Statistics Department
International Monetary Fund

October 4, 2007

For presentation to the Ottawa Group on Price Statistics
Ottawa, Canada
October 10-12, 2007



## I. Introduction

Consumer price indexes (CPIs) need to have regularly revised weights to stay representative of household consumption patterns. While some countries update CPI weights annually to keep adjustments to the weights incremental, many countries update much less frequently and the weight adjustments often are substantial. When inflation is not highly correlated across the items comprising the index, large changes in the weights can mean significantly different estimates of aggregate inflation estimates between the old and new indexes from the beginning of the new weight reference year forward. This has implications for the methodology compilers use to link the new series onto the old.

Conventional methodology links the revised series into the published (old) series by, in effect, extrapolating the old index forward by short-term changes in the revised index in the month when the new index data are ready for dissemination. This has the advantage of producing a series with relatively smooth month to month changes that does not revise previously published data. However, as the delay in introducing the revised series gets longer (and it often is two or more years), the potential amount of uncaptured trend inflation from this practice can get large. Measuring its magnitude is straightforward. Normalize both series to the base year of the revised series. (Say, this is 2005.) Then

---

[1] Chief and Senior Economist, respectively, Real Sector Division, IMF Statistics Department. This note is part of a larger paper in preparation that considers price index linking options involving longer revision periods than are considered in this note. We sketch these options in our concluding section here. We would like to acknowledge advice on benchmarking methods from quarterly national accounts compilation from Maria Mantcheva of the IMF Statistics Department's Real Sector Division.

compare the level of the new series in the link month with the level of the old series in the link month. (Supposing revision is very timely, say, this is January 2006.) Our experience shows this difference between “new” and “old” index levels can easily be in excess of five index points by the end of the new weight reference year when linking after a long hiatus in updating the index weights (say, 10 years or more) and during a period of inflation that varies across products. Clearly, the level of the revised index relative to its weight reference year is the more accurate and thus the most credible, yet the conventional approach to linking suppresses this in the interest of smoothness.

How to resolve the apparent conflict between smoothness and trend? Price index makers have been less sympathetic to users’ dual desires on this score than national accountants. National accountants liberally use so-called benchmarking techniques to adjust high frequency series that are required to have specific linear relationships to accurate but low frequency benchmark levels of comparable information. For example, they routinely benchmark quarterly Gross Domestic Product (GDP) by industry to annual estimates of same, revising the previously published quarterly series when the less timely, but more accurate annual data arrive. They do this because the quarterly value added information generally is based on monthly and quarterly extrapolator series, and are presumptively less accurate estimates than would be obtained from a direct survey that may be conducted annual or lower frequency. This paper addresses a similar problem—providing smoothness while retaining accuracy in the long-term trend as new weighting information arrives from survey sources—in the price index context. We consider essentially the same benchmarking algorithm, the proportional Denton method, used for quarterly national accounts.

Within the benchmarking context, as noted, the smoothness/trend tradeoff cannot be solved without revising back data. The revision period would start with the beginning of the reference period of the new index weights. This is unlikely to be controversial if the idea of revising previously published data is granted in the first place (admittedly, for some price index compilers, a tall order).

If the compiling agency will not or cannot revise published information, we take the perspective of a sophisticated user faced with information suggesting standard linking is affecting the long term trend in the index, and desiring a method for adjusting the monthly series minimally so that the annual trend is preserved. We will consider two cases.

In the first case, monthly data on the rebased, new index prior to the link month are not available from the compiling agency, so the user has only the monthly values of the old index during the natural base year of the new index and up until the link month. The published dates of the new index thus will exclude its monthly values during its natural base year. However, the compiling agency does provide the level of the new index in the link month relative to its natural base. This may seem somewhat artificial to a compiler, as it should, in principle, be possible to compile the new index on a monthly basis for the months of its own base year. This situation may well be the one faced by users of the data, however, because the compiler does not release these data.

In the second case, we the compiling agency releases, for the information of users, monthly data on the new index relative to its natural base, beginning with the first month of the new base year, even though these monthly values will not be the same as the previously published information on the old index for the same period.

We show how to smooth this break while minimizing the size of revisions using a variant of the proportional Denton algorithm, the most commonly applied national accounts benchmarking technique. We consider case one only in this note, as application of the method to case 2 is identical, differing only in context and interpretation.

## II. Linking as a benchmarking problem

To fix ideas, we consider a numerical example of data from two series shown in columns 2 and 3 of Table 1:

$P_{1990,2004}^{200312},\ldots,P_{1990,2004}^{200501}$ represents the monthly values from January 2003 (200301) through January 2005 (200501) of a price series averaging together price relatives whose average level is 100 in 1990 using weights from 1990. The subscript '1990, 2004' should be read 'with weights from 1990, normalized to 2004 = 100'. This is thus the "old" index normalized to 2004 = 100.

$P_{2004,2004}^{200501},\ldots,P_{2004,2004}^{200512}$ represents the monthly values from January 2005 (200501) through December 2005 (200512) of a price series averaging together price relatives whose average level is 100 in 2004 using weights from 2004. The subscript '2004, 2004' should be read 'with weights from 2004, normalized to 2004 = 100'. This is thus the "new" index normalized to its weight and price reference (natural) base, 2004 = 100.

In this example, the data available on the new series exclude the months of its reference year, 2004, so we are considering case 1 described in the introduction. Column 4 of Table 1 shows a conventionally linked series with link month January 2005, labeled as $P_{conventional,2004}^{ym}$. As expected, it shows a an unbroken trend in short term price developments. However, the difference between the level of the old index on a 2004 = 100 basis (column 2) in January 2005 and the new index (column 3) is a sizeable 6 percent. Yet the conventionally linked index sits at the level of the old index in January

Table 1. Comparison of linking methods using a numerical example

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Year/month | Old series 2004=100 | New series 2004 = 100 | Conventionally linked | Linked and benchmarked |
| 200112 | 94.92 | | 94.92 | 94.92 |
| 200201 | 97.32 | | 97.32 | 97.32 |
| 200202 | 98.60 | | 98.60 | 98.60 |
| 200203 | 96.64 | | 96.64 | 96.64 |
| 200204 | 97.60 | | 97.60 | 97.60 |
| 200205 | 98.55 | | 98.55 | 98.55 |
| 200206 | 99.50 | | 99.50 | 99.50 |
| 200207 | 100.58 | | 100.58 | 100.58 |
| 200208 | 98.44 | | 98.44 | 98.44 |
| 200209 | 97.65 | | 97.65 | 97.65 |
| 200210 | 100.18 | | 100.18 | 100.18 |
| 200211 | 98.97 | | 98.97 | 98.97 |
| 200212 | 101.02 | | 101.02 | 101.02 |
| 200301 | 99.31 | | 99.31 | 99.31 |
| 200302 | 99.78 | | 99.78 | 99.78 |
| 200303 | 99.50 | | 99.50 | 99.50 |
| 200304 | 100.43 | | 100.43 | 100.43 |
| 200305 | 98.20 | | 98.20 | 98.20 |
| 200306 | 100.41 | | 100.41 | 100.41 |
| 200307 | 101.77 | | 101.77 | 101.77 |
| 200308 | 101.50 | | 101.50 | 101.50 |
| 200309 | 103.41 | | 103.41 | 103.41 |
| 200310 | 100.35 | | 100.35 | 100.35 |
| 200311 | 103.24 | | 103.24 | 103.24 |
| 200312 | 101.52 | | 101.52 | 101.52 |
| 200401 | 98.66 | | 98.66 | **97.96** |
| 200402 | 99.29 | | 99.29 | **98.06** |
| 200403 | 99.65 | | 99.65 | **98.09** |
| 200404 | 101.54 | | 101.54 | **99.82** |
| 200405 | 100.86 | | 100.86 | **99.22** |
| 200406 | 100.60 | | 100.60 | **99.22** |
| 200407 | 98.12 | | 98.12 | **97.23** |
| 200408 | 99.13 | | 99.13 | **98.88** |
| 200409 | 101.93 | | 101.93 | **102.53** |
| 200410 | 101.22 | | 101.22 | **102.87** |
| 200411 | 101.18 | | 101.18 | **104.09** |
| 200412 | 97.82 | | 97.82 | **102.04** |
| 200501 | 102.45 | 108.54 | 102.45 | 108.54 |
| 200502 | | 105.86 | **99.92** | 105.86 |
| 200503 | | 109.36 | **103.23** | 109.36 |
| 200504 | | 107.84 | **101.79** | 107.84 |
| 200505 | | 105.17 | **99.27** | 105.17 |
| 200506 | | 107.73 | **101.69** | 107.73 |
| 200507 | | 107.55 | **101.52** | 107.55 |
| 200508 | | 107.67 | **101.63** | 107.67 |
| 200509 | | 108.68 | **102.58** | 108.68 |
| 200510 | | 109.46 | **103.32** | 109.46 |
| 200511 | | 109.65 | **103.49** | 109.65 |
| 200512 | | 114.29 | **107.88** | 114.29 |

and effectively suppresses this difference. We have bolded the data for January 2005 forward in Column 4 of Table 1 to indicate that, while the conventionally linked series reflects the short-term changes in the new index series, the linking has permanently affected its level from January 2005 forward. Conventional linking of this series thus biases the long term trend of inflation down.

Our problem, then, is to link these two series (1) without affecting the level of the new series relative to its natural base after its natural base year (i.e., without suppressing the 6 percent higher level of inflation at the end of 2004 as measured by the new series), (2) without revising the old series prior to the natural base year of the new series (the link year), (3) while minimally adjusting the most current available monthly data during the link year, and (4) while enforcing the constraint that the average of the adjusted monthly data during the link year is equal to 100.

There are a variety of approaches to setting this problem up, depending on the objective function for minimizing the adjustments to the most current available monthly data. See, for example IMF (2001), Chapter 6, Annex 6.1. We will adopt the criterion of minimizing the squared differences between successive monthly proportional changes in the original and adjusted index during the link year, 2004.[2] National accountants know this as the proportional Denton (1971) method of benchmarking. We therefore solve the following constrained minimization to determine an adjusted monthly price index $\hat{P}^{t}_{*,2004}$, $t$ = 200401, ..., 200401:

$$\min \sum_{t=200401}^{200501} \left[ \frac{\hat{P}^{t}_{*,2004}}{P^{t}_{1990,2004}} - \frac{\hat{P}^{t-1}_{*,2004}}{P^{t-1}_{1990,2004}} \right]^2$$

Subject to

$$\hat{P}^{200312}_{*,2004} = P^{200312}_{1990,2004}$$

$$\tfrac{1}{12} \sum_{t=200401}^{200412} \hat{P}^{t}_{*,2004} = 100$$

$$\hat{P}^{200501}_{*,2004} = P^{200501}_{2004,2004}$$

In matrix form, this problem is

---

[2] As noted by Denton (1971), if the last month of 2004, just prior to the link year, 2004, is outside the range of adjustment, then

$$\min \hat{P}'\Lambda\hat{P} \text{ subject to } X'\hat{P} = K$$

where

$$\hat{P} = \begin{bmatrix} \hat{P}_{*,2004}^{200312} \\ \hat{P}_{*,2004}^{200401} \\ \hat{P}_{*,2004}^{200402} \\ \vdots \\ \hat{P}_{*,2004}^{200501} \end{bmatrix}, \ \Lambda = \Gamma'\Gamma ,$$

$$\Gamma = \begin{bmatrix} \frac{1}{P_{1990,2004}^{200312}} & 0 & 0 & \cdots & 0 & 0 \\ -\frac{1}{P_{1990,2004}^{200312}} & \frac{1}{P_{1990,2004}^{200401}} & 0 & \cdots & 0 & 0 \\ 0 & -\frac{1}{P_{1990,2004}^{200401}} & \frac{1}{P_{1990,2004}^{200402}} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{P_{1990,2004}^{200412}} & 0 \\ 0 & 0 & 0 & \cdots & -\frac{1}{P_{1990,2004}^{200412}} & \frac{1}{P_{1990,2004}^{200501}} \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{12}\iota & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$K = \begin{bmatrix} P_{1990,2004}^{200312} \\ 100 \\ P_{2004,2004}^{200501} \end{bmatrix}$$

$$\iota = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}^{(12\times 1)}$$

Following Denton (1971), the Lagrangian form of this minimization is

$$\min_{\hat{P},\lambda} \hat{P}'\Lambda\hat{P} - 2\lambda'\left(K - X'\hat{P}\right),$$

whose first order conditions are

$$\begin{bmatrix} \Lambda & X \\ X' & 0 \end{bmatrix}\begin{bmatrix} \hat{P} \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ K \end{bmatrix},$$

and thus the solution of this minimization is

$$\begin{bmatrix} \hat{P} \\ \lambda \end{bmatrix} = \begin{bmatrix} \Lambda & -X \\ X' & 0 \end{bmatrix}^{-1}\begin{bmatrix} 0 \\ K \end{bmatrix}.$$

From this, in particular,

$$\hat{P} = P_{1990,2004} + \Lambda^{-1}X\left(X'\Lambda X\right)^{-1}\left(K - X'P_{1990,2004}\right),$$

where

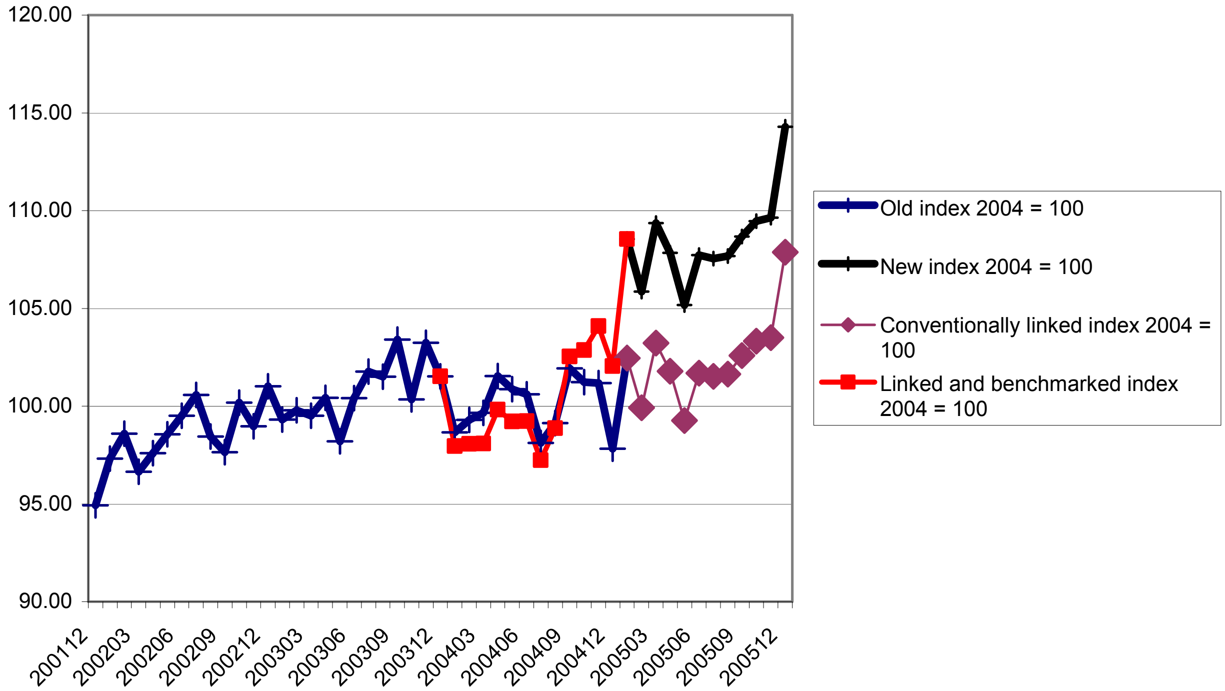
$$P_{1990,2004} = \begin{bmatrix} P_{1990,2004}^{200312} \\ \vdots \\ P_{1990,2004}^{200501} \end{bmatrix}.$$

Column 5 of Table 1, labeled $P_{*,2004}^{ym}$, displays the results of this benchmarking calculation, and Figure 1 plots the information in Table 1. The Denton linked series lies clearly above the conventionally linked series by the middle of 2004, displaying markedly higher trend inflation and a similar seasonal profile in the link year, 2004, to the new index in the year following the link year, 2004. On the other hand, the Denton procedure introduces a stronger seasonal pattern than the original series displays.

The somewhat exaggerated seasonal pattern is introduced by the constraint that the adjusted index must average to 100 in 2004. As noted in IMF(2001), there are methods for influencing the shape of seasonality induced by benchmarking procedures, and practitioners generally try to avoid applying these techniques to seasonal series.[3] One reasonable approach would be to further revise the data for 2004 as the seasonal profile for 2005 accumulates by incorporating penalties to the algorithm making the seasonal

[3] National accountants tend to seasonally adjust quarterly data before benchmarking them to annual totals.

**Figure 1. Comparison of Conventionally and Benchmark Linked Series**



profile more similar to that of 2005. Our example here happens to produce a trough centered in June 2004, similar to the June trough during 2005.

## III. Observations

Users of price indexes value both accuracy in long-term trend as well as short-term smoothness. We suggest that these twin objectives will be very difficult to achieve simultaneously without implementation of an explicit revision policy, analogous to the practice in national accounts. It is understood that in most countries there also is a vocal constituency of users intolerant of revisions to published data. Prominent in this group are users who need price indexes for contract escalation in commercial, labor, and pension agreements. We think this constituency can be served by publishing the revisable series alongside the traditional, non-revised index.[4] In the interim, compilers should initiate a dialog with users and contract designers to explore the implications of using revisable series in escalation.

Users also value explanations of the contributions index components make to aggregate inflation. Our note has not dealt with this question in the context of the specific linking algorithm we examine here. It would be a useful next step to work out these contribution to change calculations for this case. We expect to be able to exploit work that may already have been done along these lines in the national accounting benchmarking literature.

## IV. Alternative approaches with more information and longer revision periods

The benchmark-type link discussed here has been applied to keep the revision period as short as possible, and does not require as inputs any more information than the compiler has published. In particular, it does not require information on the new index during the new weight reference period. What if there is additional information?

[4] The United States Bureau of Labor Statistics, for example, publishes its revisable Chained Consumer Price Index for All Urban Consumers (C-CPI-U) alongside the non-revised indexes, the Laspeyres CPI for All Urban Consumers (CPI-U) and the CPI for Urban Wage-earners (CPI-W). See http://stats.bls.gov/news.release/pdf/cpi.pdf. We note, however, that while the C-CPI-U is revisable, as a month-weighted, monthly chained, superlative (Törnqvist) price index, it does not need the type of linking we consider in this note because its weights update at the same frequency as the index. See http://stats.bls.gov/cpi/superlink.htm. Benchmarking techniques such as those considered here could be applied to the currently non-revised CPI-U and CPI-W. However, the weights of these indexes are moving averages of multiple years and are updated frequently enough that the within-link-period trend discrepancy between index versions is likely to be small. It is this discrepancy benchmarking eliminates, preserving trend while preventing a series break.

If the basic heading price index series from which the all items index and the weights for the benchmark periods (in our earlier example, 1990 and 2004) are available, the Laspeyres index can be compiled between the weight reference years using either benchmark as the weight reference.

A straightforward approach to producing a blended estimator producing no series breaks is to take a sliding weight average of the two indexes month by month. A natural set of weights when the time in number of months from the beginning of the first weight reference period (e.g., 199001) to the beginning of the period following the second weight reference period (e.g., 200501) is $N$ is $w^t = (N - t)/N$. This 'sliding weight' estimator is thus

$$\tilde{P}^t_{*,B} = w^t P^t_{B_1,B} + (1 - w^t) P^t_{B_2,B}$$

Where $t = 0$ just prior to the first benchmark (e.g., 198912 in our example) and $t = N$ just after the second benchmark (e.g., 200501), $B$ is the normalization year in which the index is presented, $B_1$ is the first weight reference period (1990 in our example), and $B_2$ is the second weight reference period (2004 in our example). By the end of the last month of the second weight reference period (e.g., 200412), the weight of the index with the newer weights is $(N - 1)/N$. The months of the adjusted series would thereafter be the new Laspeyres index.

Of course, with this information set, we also can compare the average price levels in the two weight reference years using a superlative index formula. Interpolation between the two weight reference years could be done by Denton-type benchmarking the old monthly series between the two superlative annual levels. (In our earlier example, the level of 2004 would be 100 and the level of 1990 would be the level of the superlative in 1990 relative to 2004.) To pick up the most recent trend information, an additional benchmarking constraint also would set the adjusted series in the first month after the second weight reference year equal to the level of the newly reweighted Laspeyres series relative to the second weight reference year. (In our earlier example, this would be January 2005 of the new index with 2004 weight reference base.) The remainder of the series would then be the new Laspeyres series, until a new, say, third weight reference year is set. At that point, a third benchmark year could be added to the series benchmark constraints, with the tail of the series comprising the most current observations being the Laspeayres index relative to the new, third weight reference period. This process of extension with a Laspeyres index, followed by superlative annual benchmarking, when a new weight reference is available, could be done indefinitely. This process closely parallels the typical approach national accounting benchmarking practice takes to determining the level of the current price aggregate that implicitly underlies the index we are considering when new, 'benchmark' source information becomes available.

References

Denton, Frank T., Adjustment of Monthly or Quarterly Series to Annual Totals: An Approach Based on Quadratic Minimization, *Journal of the American Statistical Association* 66(333), March 1971.

International Monetary Fund, *Quarterly National Accounts Manual*, Washington: International Monetary Fund, 2001.