Months versus years

Ralph Turvey

Problems

A CPI, in contrast with textbook Laspeyres indexes, where P's and Q's relate to the same periods, uses a year's expenditures for estimating the weights but a month's prices for computing the index, thus estimating successive *months*' costs of an *annual* basket. This combination of data for a year and data for a month within one computation creates various problems, some of which are often ignored.

One of these problems relate to the computation of annual averages. These are used for:

1. Price updating weights from the weight reference-year to the price reference-month — which requires dividing each weight by an annual index and multiplying it by a monthly index.
2. Making whole-year to whole-year index comparisons.
3. Deflating annual series.
4. The Rothwell treatment of seasonal items, which requires weighted average prices for a whole year as reference prices.

A second problem relates to the computation of month-to-month index changes. Because sets of *monthly* weighting data are never available, only a single set of *annual* weights, it is impossible, even retrospectively, to compute true Laspeyres-, Paasche- or Fisher-type index comparisons between any pair of months. Annual weights *have* to be used.

1. Computing annual averages

Unfortunately, annual averages of prices are difficult or even impossible to calculate for most elementary aggregates on account of item replacements and missing observations.

All that can be done in such cases is to use transitive-type monthly elementary aggregate indexes. By this I mean indexes which would be transitive if there were no item replacements or missing observation. These allow calculation of ratios, denoted $\overset{year}{\underset{d}{R}}$, of the annual average of twelve monthly indexes for each year to the December index. (assuming annual chaining in December, denoted *d*).

These ratios unnecessary for 12-month comparisons

I_y^x signifies a micro index for x with y as price reference period. The 12-month change from September 1993 to September 1994 when the weights used in each year were derived from the last year but one can be expressed as

$$1-\frac{\sum W_{92} I_{year93}^{s94} \cdot \sum W_{91} I_{year92}^{year93}}{\sum W_{91} I_{year92}^{s93}}$$

when the I_{year}^{x} micro indexes relate each month's prices to an annual average of prices or one annual average to another.

Substituting $\overset{year}{\underset{d}{R}}$ gives the above twelve month change as:

$$1-\frac{\sum W_{92} \dfrac{I_{d93}^{s94}}{\overset{year93}{\underset{d93}{R}}} \cdot \sum W_{91} I_{d92}^{d93} \dfrac{\overset{year93}{\underset{d93}{R}}}{\overset{year92}{\underset{d92}{R}}}}{\sum W_{91} \dfrac{I_{d92}^{s93}}{\overset{year92}{\underset{d92}{R}}}}$$

which simplifies to

$$=1-\frac{\sum W_{92} I_{d93}^{a94} \cdot \sum W_{91} I_{d92}^{d93}}{\sum W_{91} I_{d92}^{a93}}$$

Thus the $\overset{year}{\underset{d}{R}}$ cancel out and it does not matter whether they are computed using simple averages or quantity-weighted averages of the twelve indexes for each year. Month to month comparisons can be made without using any annual averages.

Simple averages or quantity-weighted averages

Annual averages, however, *are* necessarily required for whole-year comparisons and for price-updating weights from the weight reference-year to a later year or month. Here it *does* matter whether simple or quantity-weighted averages of the twelve indexes for each year are used.

If purchase quantities and prices are positively or negatively correlated within a year, a weighted average of twelve monthly indexes will exceed or fall short of their simple average. Such correlations certainly exist for certain index components:

- The availability of fresh products varies seasonally, their prices moving inversely.
- Clothing purchases are particularly large in months when there are Sales.
- December purchases exceed their monthly average for many items, so that in a year with marked inflation there is a positive correlation.

Price-quantity correlations are likely to be small for highly aggregate subindexes, however, except for years with rapid inflation. Examples for such subindexes are provided by some UK data. These relate to seasonally unadjusted monthly retail sales volumes by three groups of outlets: Predominantly food stores, Household goods stores and Textile and clothing stores. Matching them with three approximately corresponding monthly price sub-indexes, allows calculation of the differences between weighted and simple average annual indexes shown in the table below. They are very small, but they do exist.

	Food	Household Goods	Clothing & Footwear
	Weighted minus simple average % points		
1992	0.0	0.1	0.2
1993	0.0	0.1	0.3
1994	0.0	0.1	0.3
1995	0.1	0.3	0.4
1996	0.0	0.1	0.5

As for monthly fresh product indexes, it is clear that simple averages will exceed their weighted averages because of the marked negative price-quantity correlations for such items. (I have found no data to illustrate this, though users of the Rothwell method obviously have such data.) Weight updating from the weight reference-year to the price reference-month using the simple average will yield lower price-updated fresh product weights than if the weighted averages are used.

Since monthly quantity data are unavailable for the majority of elementary aggregate indexes and of subindexes, monthly quantity weights have to be monthly deflated expenditures. Denoting the monthly indexes as I_m and the monthly expenditures as E_m, the weighted averages will thus have to be computed as:

$$\frac{\sum_{m=1}^{m=12} I_m \frac{E_m}{I_m}}{\sum_{m=1}^{m=12} \frac{E_m}{I_m}} = \frac{\sum_{m=1}^{m=12} E_m}{\sum_{m=1}^{m=12} \frac{E_m}{I_m}}$$

2. Which annual weights?

Consider a CPI which is reweighted annually, say in December. The most recent weights available in December 1998 or January 1999 will be those for 1997 unless a mixture of provisional estimates and projections is used for 1998. In either case they relate to a whole year.

Weights for current monthly indexes

For the indexes for January 1999 and succeeding months, there is no practical possibility of more up to date weights. So the twelve indexes for 1999 can only be fixed weight indexes. For these weights there is a choice between using:

(1) Rough provisional estimates and projections for 1998
(2) Estimates for 1997
(3) Estimates for 1996 (if processing 1997 data takes more than twelve months).

Weights for month to month comparisons

Index theory fails to provide any ideal for such comparisons. **The economic theory of ideal indexes and superlativeness assumes that the weight and price reference-periods coincide and are of the same length.** Since in practice weight reference-periods are whole years, the theory can only guide the construction of indexes for whole years. It cannot guide the construction of price indexes used to compare one month with another.

Shapiro, M. and Wilcox, D.,"Alternative strategies for aggregating prices in the CPI" (NBER Working Paper 5980, March 1997) provide formulae for chained Laspeyres-type, Paasche-type and Superlative–type monthly indexes using annual weights and calculate them retrospectively for December to December. Their Laspeyres-type index weights elementary aggregate indexes based on December in the preceding year by expenditures for the whole of that preceding year; their Paasche-type weights them by expenditures for the whole of the current year. Their Törnqvist-type index uses the simple mean of the weights for the two years. They note that this method could be refined in two ways. One would centre the price reference-period within the period over which the expenditures are calculated, e.g. basing the elementary aggregate indexes on mid-year rather than December prices. The other would price-update or downdate expenditures to the price reference-month (and perhaps the current month).

This all seems sensible, though the resulting Fisher-type and Törnqvist-type indexes cannot be called "superlative".

However, there is another possibility which seems better than either of them. Why not use a weights estimate for year *t* to compute the twelve-month overall index change from December *t-1* to December *t* or from January *t* to January *t+1*?

- If the index is computed as an arithmetic weighted average of all the elementary aggregate indexes it comes to the same as a Divisia index when quantities are constant throughout the year.

- If computed as a geometric weighted average of all the elementary aggregate indexes it comes to the same as a Divisia index when expenditure proportions are constant throughout the year.

However, the constant quantity assumption has been rejected earlier in this note.

A better justification is the simple statement that **weight estimates for the whole of 1998 are relevant for measuring price change over the whole year of, say, 1998 and weight estimates for 1997 are not**.

So unless experience shows that weights estimated in January 1999 for 1998 are likely to be substantially revised, they should be used in January 1999 to calculate a December 1997 to December 1998 twelve month change — and to calculate the indexes for January and the subsequent months of 1999.

This means that there would be two indexes every December, a long-term index used for annually chaining over the year just elapsed and a short-term index used during the current year. Why not? The Swedes do it, following the recommendations of a 1943 official investigation of index problems[1].

[1] *Betänkande angånde Levnadskostnadsindex,* Statens Offentliga Utredningar 1943:8. There is an annual long-term linking from December *t-1* (denoted *0*) to December *t* (denoted *1*) using weights from the year *t*. Starting with a Divisia index $\log P_{01} = \sum \int_0^1 \frac{p_t q_t}{\sum p_t q_t} \cdot d \log p_t$, the report derives the following approximation on the assumption that the quantities *q* are constant throughout the integration period of year *t* i.e. from *0* to *1*. $P_{01} = \frac{\sum p_1 q}{\sum p_0 q}$

In practice, elementary aggregate indexes *I* and expenditure weights *w* have to be used, so this becomes:

$$P_{01} = \frac{\sum \frac{p_1}{p} pq}{\sum \frac{p_0}{p} pq} = \frac{\sum I_{1p} w}{\sum \frac{w}{I_{p0}}} = \sum \frac{\frac{w}{I_{p0}}}{\sum \frac{w}{I_{p0}}} I_{10}$$

where *p* is the average price throughout the integration period of year *t* and since $I_{1p} = I_{10}$ divided by I_{p0} Thus the weights for year *t* have to be revalued to December *t-1* prices to be consistent with the Divisia approximation. Is this done in the Long term index and if so, how is it done? The revaluation is, in practice done, at a lower aggregation level than that for which actual year t weights are obtained, and this has been a source of concern. Year t weights are obtained for some 90 National Accounts groups, whereas the revaluation is done at a level of about 300 item groups.

Matters are simpler if constant expenditure shares, *a*, rather than constant quantities, *q*, are assumed, for then the approximation comes out as: $P_{01} = \prod (I_{10})^a$. The report suggested that this formula might be adopted after the war when consumer opportunities for substitution would be restored by the abolition of food rationing.