INDEX NUMBER THEORY AND MEASUREMENT ECONOMICS

By W.E. Diewert, March 1, 2009.

CHAPTER 7: The Use of Annual Weights in a Monthly Index

1. The Lowe Index with Monthly Prices and Annual Base Year Quantities

It is now necessary to discuss a major practical problem with the theory of bilateral indexes that we have been discussing in earlier chapters. Recall that the *Lowe* (1823) *index* was defined by equation (19) in chapter 1 as follows:

$$(1)\ P_{Lo}(p^0,p^1,q) \equiv p^1 \bullet q / p^0 \bullet q.$$

The Lowe index can be written in expenditure share form as follows:

$$(2)\ P_{Lo}(p^0,p^1,q) \equiv \sum_{n=1}^N p_n^1 q_n / \sum_{n=1}^N p_n^0 q_n = \sum_{n=1}^N (p_n^1/p_n^0)\, s_n$$

where the (hypothetical) *hybrid expenditure shares* s_n corresponding to the quantity weights vector q are defined by:[1]

$$(3)\ s_n \equiv p_n^0 q_n / \sum_{n=1}^N p_n^0 q_n \qquad \text{for } n = 1,\ldots,N.$$

Up to now, it has been assumed that the quantity vector $q \equiv (q_1,q_2,\ldots,q_N)$ that appeared in the definition of the Lowe index, $P_{Lo}(p^0,p^1,q)$, is either the base period quantity vector q^0 or the current period quantity vector q^1 or an average of these two quantity vectors. In fact, in terms of actual statistical agency practice, the quantity vector q is usually taken to be an annual quantity vector that refers to a *base year*, b say, that is prior to the base period for the prices, period 0. Typically, a statistical agency will produce a Consumer Price Index at a monthly or quarterly frequency but for the sake of definiteness, a monthly frequency will be assumed in what follows. Thus a typical price index will have the form $P_{Lo}(p^0,p^t,q^b)$, where p^0 is the price vector pertaining to the base period month for prices, month 0, p^t is the price vector pertaining to the current period month for prices, month t say, and q^b is a reference basket quantity vector that refers to the base year b, which is equal to or prior to month 0.[2] Note that this Lowe index $P_{Lo}(p^0,p^t,q^b)$ is *not* a true Laspeyres index (because the annual quantity vector q^b is not equal to the monthly quantity vector q^0 in general).[3]

[1] Fisher (1922; 53) used the terminology "weighted by a hybrid value" while Walsh (1932; 657) used the term "hybrid weights".
[2] Month 0 is called the price reference period and year b is called the weight reference period.
[3] Triplett (1981; 12) defined the Lowe index, calling it a Laspeyres index, and calling the index that has the weight reference period equal to the price reference period, a pure Laspeyres index. However, Balk (1980; 69) asserted that although the Lowe index is of the fixed base type, it is not a Laspeyres price index. Triplett also noted the hybrid share representation for the Lowe index defined by (2) and (3) above. Triplett noted that the ratio of two Lowe indices using the same quantity weights was also a Lowe index. Baldwin (1990; 255) called the Lowe index an *annual basket index*.

The question is: why do statistical agencies *not* pick the reference quantity vector q in the Lowe formula to be the monthly quantity vector q^0 that pertains to transactions in month 0 (so that the index would reduce to an ordinary Laspeyres price index)? There are two main reasons why this is not done:

- Most economies are subject to seasonal fluctuations and so picking the quantity vector of month 0 as the reference quantity vector for all months of the year would not be representative of transactions made throughout the year.
- Monthly household quantity or expenditure weights are usually collected by the statistical agency using a household expenditure survey with a relatively small sample. Hence the resulting weights are usually subject to very large sampling errors and so standard practice is to average these monthly expenditure or quantity weights over an entire year (or in some cases, over several years), in an attempt to reduce these sampling errors.

The index number problems that are caused by seasonal monthly weights will be studied in more detail in a later chapter. For now, it can be argued that the use of annual weights in a monthly index number formula is simply a method for dealing with the seasonality problem.[4]

One problem with using annual weights corresponding to a perhaps distant year in the context of a monthly Consumer Price Index must be noted at this point: if there are systematic (but divergent) trends in commodity prices and households increase their purchases of commodities that decline (relatively) in price and decrease their purchases of commodities that increase (relatively) in price, then the use of distant quantity weights will tend to lead to an upward bias in this Lowe index compared to one that used more current weights, as will be shown below. This observation suggests that statistical agencies should strive to get up to date weights on an ongoing basis.

It is useful to explain how the annual quantity vector q^b could be obtained from monthly expenditures on each commodity during the chosen base year b. Let the month m expenditure of the reference population in the base year b for commodity i be $v_i^{b,m}$ and let the corresponding price and quantity be $p_i^{b,m}$ and $q_i^{b,m}$ respectively. Of course, value, price and quantity for each commodity are related by the following equations:

(4) $v_n^{b,m} = p_n^{b,m} q_n^{b,m}$; $\quad n = 1,\ldots,N$; $m = 1,\ldots,12$.

For each commodity n, the annual total, q_n^b can be obtained by price deflating monthly values and summing over months in the base year b as follows:

(5) $q_n^b = \sum_{m=1}^{12} v_n^{b,m} / p_n^{b,m} = \sum_{m=1}^{12} q_n^{b,m}$; $\quad n = 1,\ldots,N$

[4] In fact, the use of the Lowe index $P_{Lo}(p^0,p^t,q^b)$ in the context of seasonal commodities corresponds to Bean and Stine's (1924; 31) Type A index number formula. Bean and Stine made 3 additional suggestions for price indexes in the context of seasonal commodities. Their contributions will be evaluated in a later chapter.

where (4) was used to derive the second equation in (5). In practice, the above equations will be evaluated using aggregate expenditures over closely related commodities and the price $p_n^{b,m}$ will be the month m price index for this elementary commodity group n in year b relative to the first month of year b.

For some purposes, it is also useful to have annual prices by commodity to match up with the annual quantities defined by (5). Following national income accounting conventions, a reasonable[5] price p_n^b to match up with the annual quantity q_n^b is the value of total consumption of commodity n in year b divided by q_n^b. Thus we have:

$$(6)\ p_n^b \equiv \sum_{m=1}^{12} v_n^{b,m} / q_n^b \qquad n = 1,\ldots,N$$
$$= \sum_{m=1}^{12} v_n^{b,m} / [\sum_{m=1}^{12} v_n^{b,m} / p_n^{b,m}] \qquad \text{using (5)}$$
$$= [\sum_{m=1}^{12} s_n^{b,m} (p_n^{b,m})^{-1}]^{-1}$$

where the share of annual expenditure on commodity n in month m of the base year is

$$(7)\ s_n^{b,m} \equiv v_n^{b,m} / \sum_{k=1}^{12} v_n^{b,k}\ ; \qquad n = 1,\ldots,N\ ;\ m = 1,\ldots,12.$$

Thus the annual base year price for commodity n, p_n^b, turns out to be a monthly expenditure weighted *harmonic mean* of the monthly prices for commodity n in the base year, $p_n^{b,1}, p_n^{b,2},\ldots, p_n^{b,12}$.

Using the annual commodity prices for the base year defined by (6), a vector of these prices can be defined as $p^b \equiv [p_1^b,\ldots,p_N^b]$. Using this definition, the Lowe index $P_{Lo}(p^0,p^t,q^b)$ can be expressed as a ratio of two Laspeyres indexes where the price vector p^b plays the role of base period prices in each of the two Laspeyres indexes:

$$(8)\ P_{Lo}(p^0,p^t,q^b) \equiv p^t \bullet q^b / p^0 \bullet q^b$$
$$= [p^t \bullet q^b / p^b \bullet q^b]/[p^0 \bullet q^b / p^b \bullet q^b]$$
$$= P_L(p^b,p^t,q^b)/P_L(p^b,p^0,q^b)$$
$$= \sum_{n=1}^{N} (p_n^t/p_n^b)\, s_n^b / \sum_{n=1}^{N} (p_n^0/p_n^b)\, s_n^b$$

where the Laspeyres formula P_L was defined in Chapter 1. Thus the above equation shows that the Lowe monthly price index comparing the prices of month 0 to those of month t using the quantities of base year b as weights, $P_{Lo}(p^0,p^t,q^b)$, is equal to the Laspeyres index that compares the prices of month t to those of year b, $P_L(p^b,p^t,q^b)$, divided by the Laspeyres index that compares the prices of month 0 to those of year b, $P_L(p^b,p^0,q^b)$. Note that the Laspeyres index in the numerator can be calculated if the base

[5] Hence these annual commodity prices are essentially unit value prices. Under conditions of high inflation, the annual prices defined by (6) may no longer be "reasonable" or representative of prices during the entire base year because the expenditures in the final months of the high inflation year will be somewhat artificially blown up by general inflation. Under these conditions, the annual prices and annual commodity expenditure shares should be interpreted with caution. For more on dealing with situations where there is high inflation within a year, see Hill (1996).

year commodity expenditure shares, s_n^b, are known along with the price ratios that compare the prices of commodity n in month t, p_n^t, with the corresponding annual average prices in the base year b, p_n^b. The Laspeyres index in the denominator can be calculated if the base year commodity expenditure shares, s_n^b, are known along with the price ratios that compare the prices of commodity n in month 0, p_n^0, with the corresponding annual average prices in the base year b, p_n^b.

There is another convenient formula for evaluating the Lowe index, $P_{Lo}(p^0,p^t,q^b)$, and that is to use the hybrid weights formula, (2). In the present context, the formula becomes:

$$(9)\ P_{Lo}(p^0,p^t,q^b) \equiv \sum_{n=1}^{N} p_n^t q_n^b / \sum_{n=1}^{N} p_n^0 q_n^b$$
$$= \sum_{n=1}^{N} (p_n^t/p_n^0)\, s_n^{0b}$$

where the (hypothetical) *hybrid expenditure shares* s_n^{0b} corresponding to the quantity weights vector q are defined by:[6]

$$(10)\ s_n^{0b} \equiv p_n^0 q_n^b / \sum_{n=1}^{N} p_n^0 q_n^b \qquad \text{for } n = 1,\ldots,N$$
$$= p_n^b q_n^b (p_n^0/p_n^b) / \sum_{j=1}^{N} p_j^b q_j^b (p_j^0/p_j^b).$$

The second equation in (10) shows how the base year expenditures on commodity n, $p_n^b q_n^b$, can be multiplied by the commodity price indexes, p_n^0/p_n^b, in order to calculate the hybrid shares.

There is one additional formula for the Lowe index, $P_{Lo}(p^0,p^t,q^b)$, that will be exhibited. Note that the Laspeyres decomposition of the Lowe index defined by the fourth line in (8) involves the very long term price relatives, p_n^t/p_n^b, which compare the prices in month t, p_n^t, with the possibly distant base year prices, p_n^b, and that the hybrid share decomposition of the Lowe index defined by the second line in (9) involves the long term monthly price relatives, p_i^t/p_i^0, which compare the prices in month t, p_i^t, with the base month prices, p_i^0. Both of these formulae are not satisfactory in practice due to the problem of sample attrition: each month, a substantial fraction of commodities disappears from the marketplace and thus it is useful to have a formula for updating the previous month's price index using just month over month price relatives. In other words, long term price relatives disappear at a rate that is too large in practice to base an index number formula on their use. The Lowe index for month t+1, $P_{Lo}(p^0,p^{t+1},q^b)$, can be written in terms of the Lowe index for month t, $P_{Lo}(p^0,p^t,q^b)$, and an updating factor as follows:

$$(11)\ P_{Lo}(p^0,p^{t+1},q^b) \equiv p^{t+1}\bullet q^b / p^0\bullet q^b$$
$$= [p^t\bullet q^b / p^0\bullet q^b][p^{t+1}\bullet q^b / p^t\bullet q^b]$$
$$= P_{Lo}(p^0,p^t,q^b)[p^{t+1}\bullet q^b / p^t\bullet q^b]$$
$$= P_{Lo}(p^0,p^t,q^b)[\sum_{n=1}^{N} s_n^{tb}(p_n^{t+1}/p_n^t)]$$

[6] Fisher (1922; 53) used the terminology "weighted by a hybrid value" while Walsh (1932; 657) used the term "hybrid weights".

where the hybrid weights s_n^{tb} are defined by

(12) $s_n^{tb} \equiv p_n^t q_n^b / \sum_{k=1}^N p_k^t q_k^b$; $\qquad n = 1,\ldots,N.$

Thus the required updating factor, going from month t to month t+1, is the chain link index $\sum_{n=1}^N s_n^{tb} (p_n^{t+1}/p_n^t)$, which uses the hybrid share weights s_n^{tb} corresponding to month t and base year b.

It should be noted that the month t hybrid shares, can be constructed from the previous month's hybrid shares, $s_n^{t-1,b} \equiv p_n^{t-1} q_n^b / \sum_{k=1}^N p_k^{t-1} q_k^b$, by using the following updating formula:

$$
\begin{aligned}
(13)\ s_n^{tb} &\equiv p_n^t q_n^b / \sum_{k=1}^N p_k^t q_k^b ; \qquad n = 1,\ldots,N \\
&= p_n^{t-1} q_n^b (p_n^t/p_n^{t-1}) / \sum_{j=1}^N p_j^t q_j^b \\
&= [p_n^{t-1} q_n^b (p_n^t/p_n^{t-1})/p^{t-1}\bullet q^b] / [\sum_{j=1}^N p_j^t q_j^b / p^{t-1}\bullet q^b] \\
&= s_n^{t-1,b}(p_n^t/p_n^{t-1}) / [\sum_{j=1}^N p_j^t q_j^b / p^{t-1}\bullet q^b] \\
&= s_n^{t-1,b}(p_n^t/p_n^{t-1}) / [\sum_{j=1}^N (p_j^t/p_j^{t-1}) p_j^{t-1} q_j^b / p^{t-1}\bullet q^b] \\
&= s_n^{t-1,b}(p_n^t/p_n^{t-1}) / [\sum_{j=1}^N (p_j^t/p_j^{t-1}) s_j^{t-1,b}] \\
&= (p_n^t/p_n^{t-1}) s_n^{t-1,b} / [\sum_{j=1}^N (p_j^t/p_j^{t-1}) s_j^{t-1,b}].
\end{aligned}
$$

Formula (13) can be used recursively until we get to t = 1, when (13) becomes:

$$
\begin{aligned}
(14)\ s_n^{1,b} &\equiv p_n^1 q_n^b / \sum_{k=1}^N p_k^1 q_k^b ; \qquad n = 1,\ldots,N \\
&= (p_n^1/p_n^0) s_n^{0,b} / [\sum_{j=1}^N (p_j^1/p_j^0) s_j^{0,b}].
\end{aligned}
$$

The hybrid shares, $s_n^{0,b}$, that use the components of the base year quantity vector q^b and the base month price vector p^0, can be constructed from base year expenditures, $p_n^b q_n^b$, and the "mixed" month to year price relatives, (p_n^0/p_n^b), using formula (10) above. Thus we have developed a complete set of "practical" updating formulae.

The Lowe index $P_{Lo}(p^0,p^t,q^b)$ can be regarded as an *approximation* to the ordinary Laspeyres index, $P_L(p^0,p^t,q^0)$, that compares the prices of the base month 0, p^0, to those of month t, p^t, using the quantity vector of month 0, q^0, as weights. It turns out that there is a relatively simple formula that relates these two indexes. However, before we present this formula, we digress momentarily and develop a relationship between the Paasche and Laspeyres price indexes. It turns out that we can adapt this methodology to the problem of relating the Lowe index to the Laspeyres index.

2. The Bortkiewicz Decomposition between the Paasche and Laspeyres Indexes

In this section, we will develop a relationship between the ordinary Paasche and Laspeyres price indexes.[7] In order to explain this formula, it is first necessary to make a few definitions. Define the *nth price relative* between month 0 and month t as

[7] This relationship was originally discovered by Bortkiewicz (1923; 374-375).

(15) $r_n \equiv p_n^t/p_n^0$; $\quad n = 1,\ldots,N.$

The ordinary *Laspeyres price index*, relating the prices of month 0 to those of month t, can be defined as a weighted average of these price relatives as follows:

$$
\begin{aligned}
(16)\ P_L(p^0,p^t,q^0) &\equiv \sum_{n=1}^N s_n^0\,(p_n^t/p_n^0) \\
&= \sum_{n=1}^N s_n^0\, r_n && \text{using (15)} \\
&\equiv r^*
\end{aligned}
$$

where the *month 0 expenditure shares* s_n^0 are defined as follows:

(17) $s_n^0 \equiv p_n^0q_n^0/\sum_{k=1}^N p_k^0q_k^0$; $\quad n = 1,\ldots,N.$

Define the *nth quantity relative* t_n as the ratio of the quantity of commodity n used in the month t, q_n^t, to the quantity used in month 0, q_n^0, as follows:

(18) $t_n \equiv q_n^t/q_n^0$; $\quad n = 1,\ldots,N.$

The *Laspeyres quantity index*, $Q_L(q^0,q^t,p^0)$, that compares quantities in month t, q^t, to the corresponding quantities in month 0, q^0, using the prices of month 0, p^0, as weights can be defined as a weighted average of the quantity ratios t_n as follows:

$$
\begin{aligned}
(19)\ Q_L(q^0,q^t,p^0) &\equiv p^0\bullet q^t/p^0\bullet q^0 \\
&= \sum_{n=1}^N s_n^0\, t_n && \text{using (17) and (18)} \\
&\equiv t^*.
\end{aligned}
$$

Before we compare the Paasche and Laspeyres price indexes, we need to undertake a preliminary computation using the above definitions of r_n and t_n:

$$
\begin{aligned}
(20)\ Cov(r,t,s^0) &\equiv \sum_{n=1}^N (r_n - r^*)(t_n - t^*)s_n^0 \\
&= \sum_{n=1}^N r_nt_ns_n^0 - \sum_{n=1}^N r_nt^*s_n^0 - \sum_{n=1}^N r^*t_ns_n^0 + \sum_{n=1}^N r^*t^*s_n^0 \\
&= \sum_{n=1}^N r_nt_ns_n^0 - t^*\sum_{n=1}^N r_ns_n^0 - r^*\sum_{n=1}^N t_ns_n^0 + r^*t^*\sum_{n=1}^N s_n^0 \\
&= \sum_{n=1}^N r_nt_ns_n^0 - t^*\sum_{n=1}^N r_ns_n^0 - r^*\sum_{n=1}^N t_ns_n^0 + r^*t^* && \text{using } \sum_{n=1}^N s_n^0 = 1 \\
&= \sum_{n=1}^N r_nt_ns_n^0 - t^*r^* - r^*t^* + r^*t^* && \text{using (16) and (19)} \\
&= \sum_{n=1}^N r_nt_ns_n^0 - t^*r^*.
\end{aligned}
$$

Note that $Cov(r,t,s^0)$ can be interpreted as a *weighted covariance* between the vector of price relatives, $r \equiv [r_1,\ldots,r_N]$, and the vector of quantity relatives, $t \equiv [t_1,\ldots,t_N]$, using the base period vector of expenditure shares, $s^0 \equiv [s_1^0,\ldots,s_N^0]$, as weights. More explicitly, let r and t be discrete random variables that take on the N values r_n and t_n respectively. Let s_n^0 be the joint probability that $r = r_n$ and $t = t_n$ for $n = 1,\ldots,N$ and let the joint probability be 0 if $r = r_i$ and $t = t_j$ where $i \neq j$. It can be verified that $Cov(r,t,s^0)$ defined in the first line of (20) is the covariance between the price relatives r_n and the corresponding quantity relatives t_n. This covariance can be converted into a correlation coefficient.[8]

Now we are ready to exhibit von Bortkiewicz's formula relating the Paasche and Laspeyres indexes. The *Paasche index*, $P_P(p^0,p^t,q^0)$, that compares the prices of the base month 0, p^0, to those of month t, p^t, using the quantity vector of month t, q^t, as a weighting vector is defined as follows:

$$
\begin{aligned}
(21)\ P_P(p^0,p^t,q^t) &\equiv \textstyle\sum_{n=1}^N p_n^t q_n^t / \sum_{n=1}^N p_n^0 q_n^t \\
&= \textstyle\sum_{n=1}^N r_n t_n\, p_n^0 q_n^0 / \sum_{n=1}^N t_n\, p_n^0 q_n^0 && \text{using definitions (15) and (18)} \\
&= [\textstyle\sum_{n=1}^N r_n t_n\, p_n^0 q_n^0/p^0\bullet q^0]/[\sum_{n=1}^N t_n\, p_n^0 q_n^0/p^0\bullet q^0] \\
&= \textstyle\sum_{n=1}^N r_n t_n\, s_n^0 / \sum_{n=1}^N t_n\, s_n^0 && \text{using definitions (17)} \\
&= \textstyle\sum_{n=1}^N r_n t_n\, s_n^0 / t^* && \text{using definition (19)} \\
&= [\{\textstyle\sum_{n=1}^N (r_n - r^*)(t_n - t^*)s_n^0\} + r^*t^*]/t^* && \text{using (20)} \\
&= [\textstyle\sum_{n=1}^N (r_n - r^*)(t_n - t^*)s_n^0/t^*] + r^* \\
&= [\textstyle\sum_{n=1}^N (r_n - r^*)(t_n - t^*)s_n^0/Q_L(q^0,q^t,p^0)] + P_L(p^0,p^t,q^0) \\
& && \text{using (16) and (19).}
\end{aligned}
$$

Subtracting $P_L(p^0,p^t,q^0)$ from both sides of (21) leads to the following relationship between the Paasche and Laspeyres price indexes:

$$
\begin{aligned}
(22)\ P_P(p^0,p^t,q^t) - P_L(p^0,p^t,q^0) &= \textstyle\sum_{n=1}^N (r_n - r^*)(t_n - t^*)s_n^0/Q_L(q^0,q^t,p^0) \\
&= \mathrm{Cov}(r,t,s^0)/Q_L(q^0,q^t,p^0).
\end{aligned}
$$

Thus the difference between the Paasche and Laspeyres price indexes relating the prices of period 0 to those of period t is equal to the covariance between the relative price and relative quantity vectors, $Cov(r,t,s^0)$, divided by the Laspeyres quantity index, $Q_L(q^0,q^t,p^0)$. Usually, this covariance will be negative for most value aggregates[9] so that usually the Paasche index will be less than the corresponding Laspeyres index.

In the following section, we will develop a similar relationship between the Lowe and Laspeyres indexes using the same technique as was used by Bortkiewicz.

3. The Relationship between the Lowe, Laspeyres and Paasche Indexes

We shall use the same notation for the long term monthly price relatives $r_n \equiv p_n^t/p_n^0$ that was used in the previous section so that (15)-(17) are still used in the present section. However, we shall change the definition of the t_n in the previous section in order to relate the base year annual quantities q_n^b to the base month quantities q_n^0:

$$(23)\ t_n \equiv q_n^b/q_n^0\ ; \qquad n = 1,\ldots,N.$$

[8] See Bortkiewicz (1923; 374-375) for the first application of this correlation coefficient decomposition technique.

[9] As we shall see later, this corresponds to the situation where demander substitution effects outweigh supplier substitution effects.

We also define a new *Laspeyres quantity index* $Q_L(q^0,q^b,p^0)$, which compares the base year quantity vector q^b to the base month quantity vector q^0, using the price weights of the base month p^0, as follows:

$$
\begin{aligned}
(24)\ Q_L(q^0,q^b,p^0) &\equiv p^0\bullet q^b/p^0\bullet q^0 \\
&= \textstyle\sum_{n=1}^N p_n^0 q_n^b / \sum_{n=1}^N p_n^0 q_n^0 \\
&= \textstyle\sum_{n=1}^N p_n^0 q_n^0 (q_n^b/q_n^0) / \sum_{n=1}^N p_n^0 q_n^0 \\
&= \textstyle\sum_{n=1}^N s_n^0 (q_n^b/q_n^0) && \text{using definitions (17)} \\
&= \textstyle\sum_{n=1}^N s_n^0 (t_n) && \text{using definitions (23)} \\
&\equiv t^*.
\end{aligned}
$$

Using definition (9), the *Lowe index* comparing the prices in month t to those of month 0, using the quantity weights of the base year b, is equal to:

$$
\begin{aligned}
(25)\ P_{Lo}(p^0,p^t,q^b) &\equiv \textstyle\sum_{n=1}^N p_n^t q_n^b / \sum_{n=1}^N p_n^0 q_n^b \\
&= \textstyle\sum_{n=1}^N p_n^t t_n q_n^0 / \sum_{n=1}^N p_n^0 t_n q_n^0 && \text{using definitions (23)} \\
&= \textstyle\sum_{n=1}^N r_n p_n^0 t_n q_n^0 / \sum_{n=1}^N p_n^0 t_n q_n^0 && \text{using definitions (15)} \\
&= [\textstyle\sum_{n=1}^N r_n t_n p_n^0 q_n^0/p^0\bullet q^0] / [\sum_{n=1}^N t_n p_n^0 q_n^0/p^0\bullet q^0] \\
&= \textstyle\sum_{n=1}^N r_n t_n s_n^0 / \sum_{n=1}^N t_n s_n^0 && \text{using definitions (17)} \\
&= \textstyle\sum_{n=1}^N r_n t_n s_n^0 / t^* && \text{using (24)} \\
&= [\{\textstyle\sum_{n=1}^N (r_n - r^*)(t_n - t^*)s_n^0\} + t^*r^*]/t^* && \text{using the identity (20)} \\
&= [\textstyle\sum_{n=1}^N (r_n - r^*)(t_n - t^*)s_n^0/t^*] + r^* \\
&= [\textstyle\sum_{n=1}^N (r_n - r^*)(t_n - t^*)s_n^0/t^*] + P_L(p^0,p^t,q^0) && \text{using definition (16)} \\
&= [\mathrm{Cov}(r,t,s^0)/Q_L(q^0,q^b,p^0)] + P_L(p^0,p^t,q^0)
\end{aligned}
$$

where the last equality follows using definitions (20) and (24). Subtracting the Laspeyres price index relating the prices of month t to those of month 0, $P_L(p^0,p^t,q^0)$, from both sides of (25) leads to the following relationship of this monthly Laspeyres price index to its Lowe counterpart:

$$
\begin{aligned}
(26)\ P_{Lo}(p^0,p^t,q^b) - P_L(p^0,p^t,q^0) &= \textstyle\sum_{n=1}^N (r_n - r^*)(t_n - t^*)s_n^0/Q_L(q^0,q^b,p^0) \\
&= \mathrm{Cov}(r,t,s^0)/Q_L(q^0,q^b,p^0).
\end{aligned}
$$

Thus the difference between the Lowe and Laspeyres price indexes relating the prices of period 0 to those of period t is equal to the covariance between the relative price and relative quantity vectors, $Cov(r,t,s^0)$, divided by the Laspeyres quantity index, $Q_L(q^0,q^b,p^0)$. Thus (26) tells us that the Lowe price index using the quantities of year b as weights, $P_{Lo}(p^0,p^t,q^b)$, is equal to the usual Laspeyres index using the quantities of month 0 as weights, $P_L(p^0,p^t,q^0)$, plus a covariance term $\sum_{n=1}^N (r_n - r^*)(t_n - t^*)s_n^0$ between the long term monthly price relatives $r_n \equiv p_n^t/p_n^0$ and the quantity relatives $t_n \equiv q_n^b/q_n^0$ (which are equal to the base year quantities q_n^b divided by the base month quantities q_n^0), divided by the Laspeyres quantity index $Q_L(q^0,q^b,p^0)$ between month 0 and base year b.

Formula (26) shows that the Lowe price index will coincide with the Laspeyres price index if the covariance or correlation between the month 0 to t price relatives $r_n \equiv p_n^t/p_n^0$ and the month 0 to year b quantity relatives $t_n \equiv q_n^b/q_n^0$ is zero. Note that this covariance will be zero under three different sets of conditions:

- If the month t prices are proportional to the month 0 prices so that all r_n equal r^*;
- If the base year b quantities are proportional to the month 0 quantities so that all t_n equal t^*;
- If the distribution of the relative prices r_n is independent of the distribution of the relative quantities t_n.

The first two conditions are unlikely to hold empirically but the third is possible, at least approximately, if consumers do not systematically change their purchasing habits in response to changes in relative prices.

If the covariance in (26) is negative, then the Lowe index will be less than the Laspeyres and finally, if the covariance is positive, then the Lowe index will be greater than the Laspeyres index. Although the sign and magnitude of the covariance term, $\sum_{n=1}^N (r_n - r^*)(t_n - t^*)s_n^0$, is ultimately an empirical matter, it is possible to make some reasonable conjectures about its likely sign. If the base year b precedes the price reference month 0 *and there are long term trends in prices*, then it is likely that this covariance is *positive* and hence *the Lowe index will exceed the corresponding Laspeyres price index*[10]; i.e.,

$$(27)\ P_{Lo}(p^0,p^t,q^b) > P_L(p^0,p^t,q^0).$$

To see why this covariance is likely to be positive, suppose that there is a long term *upward* trend in the price of commodity n so that $r_n - r^* \equiv (p_n^t/p_n^0) - r^*$ is positive. With normal consumer substitution responses[11], q_n^t/q_n^0 less an average quantity change of this type is likely to be negative, or, upon taking reciprocals, q_n^0/q_n^t less an average quantity change of this (reciprocal) type is likely to be positive. But if the long term upward trend in prices has persisted back to the base year b, then $t_n - t^* \equiv (q_n^b/q_n^0) - t^*$ is also likely to be positive. Hence, the covariance will be positive under these circumstances. Moreover, the more distant is the base year b from the base month 0, the bigger the residuals $t_n - t^*$ will likely be and the bigger will be the positive covariance. Similarly, the more distant is the current period month t from the base period month 0, the bigger the residuals $r_n - r^*$ will likely be and the bigger will be the positive covariance. *Thus under the assumptions that there are long term trends in prices and normal consumer*

---

[10] It is also necessary to assume that households have normal substitution effects in response to these long term trends in prices; i.e., if a commodity increases (relatively) in price, its consumption will decline (relatively) and if a commodity decreases relatively in price, its consumption will increase relatively.

[11] Walsh (1901; 281-282) was well aware of consumer substitution effects as can be seen in the following comment which noted the basic problem with a fixed basket index that uses the quantity weights of a single period: "The argument made by the arithmetic averagist supposes that we buy the same quantities of every class at both periods in spite of the variation in their prices, which we rarely, if ever, do. As a rough proposition, we –a community –generally spend more on articles that have risen in price and get less of them, and spend less on articles that have fallen in price and get more of them."

substitution responses, the Lowe index will usually be greater than the corresponding Laspeyres index.

Recall relationship (22) in the previous section, which related the difference between the Paasche and Laspeyres price indexes, $P_P(p^0,p^t,q^0)$ and $P_L(p^0,p^t,q^0)$, to the covariance term, $\sum_{n=1}^{N} (r_n - r^*)(t_n - t^*)s_n^0$, where the quantity relatives $t_n \equiv q_n^t/q_n^0$ were defined by (18). Although the sign and magnitude of the covariance term in (22), $\sum_{n=1}^{N} (r_n - r^*)(t_n - t^*)s_n^0$, is again an empirical matter, it is possible to make a reasonable conjecture about its likely sign. If *there are long term trends in prices and consumers respond normally to price changes in their purchases*, then it is likely that that this covariance is *negative* and hence the Paasche index will be less than the corresponding Laspeyres price index; i.e.,

(28) $P_P(p^0,p^t,q^t) < P_L(p^0,p^t,q^0)$.

To see why this covariance is likely to be negative, suppose that there is a long term upward trend in the price of commodity n[12] so that $r_n - r^* \equiv (p_n^t/p_n^0) - r^*$ is positive. With normal consumer substitution responses, q_n^t/q_n^0 less an average quantity change of this type is likely to be negative. Hence $t_n - t^* \equiv (q_n^t/q_n^0) - t^*$ is likely to be negative. Thus, the covariance will be negative under these circumstances. Moreover, the more distant is the base month 0 from the current month t, the bigger in magnitude the residuals $t_n - t^*$ will likely be and the bigger in magnitude will be the negative covariance.[13] Similarly, the more distant is the current period month t from the base period month 0, the bigger the residuals $r_n - r^*$ will likely be and the bigger in magnitude will be the covariance. *Thus under the assumptions that there are long term trends in prices and normal consumer substitution responses, the Laspeyres index will be greater than the corresponding Paasche index*, with the divergence likely growing as month t becomes more distant from month 0.

Putting the arguments in the three previous paragraphs together, it can be seen that under the assumptions that there are long term trends in prices and normal consumer substitution responses, the Lowe price index between months 0 and t will exceed the corresponding Laspeyres price index which in turn will exceed the corresponding Paasche price index; i.e., under these hypotheses,

(29) $P_{Lo}(p^0,p^t,q^b) > P_L(p^0,p^t,q^0) > P_P(p^0,p^t,q^t)$.

Thus if the long run target price index is an average of the Laspeyres and Paasche indexes, it can be seen that the Laspeyres index will have an *upward bias* relative to this target index and the Paasche index will have a *downward bias*. In addition, *if the base year b is prior to the price reference month, month 0, then the Lowe index will also have*

---

[12] The reader can carry through the argument if there is a long term relative decline in the price of the ith commodity. The argument required to obtain a negative covariance requires that there be some differences in the long term trends in prices; i.e., if all prices grow (or fall) at the same rate, we have price proportionality and the covariance will be zero.

[13] However, $Q_L = t^*$ may also be growing in magnitude so the net effect on the divergence between P_L and P_P is ambiguous.

an upward bias relative to the Laspeyres index and hence also to the target index. The previous sentence is not good news for statistical agencies that base their consumer price index on the Lowe index that uses base year quantities for a distant year as weights.

5. The Lowe Index and Midyear Indexes

The discussion in the previous sections assumed that the base year b for quantities preceded the base month for prices, month 0. However, if the current period month t is quite distant from the base month 0, then it is possible to collect expenditure information for a base year b *that lies between months 0 and t*. If the year b does fall between months 0 and t, then the Lowe index becomes a *midyear index*.[14] It turns out that if the base year is between monthly periods 0 and t, then the Lowe midyear index no longer has the upward biases indicated by the inequalities in (29) under the assumption of long term trends in prices and normal substitution responses by quantities.

We now assume that the base year quantity vector q^b corresponds to a year that lies between months 0 and t. Under the assumption of long term trends in prices and normal substitution effects so that there are also long term trends in quantities (in the opposite direction to the trends in prices so that if the nth commodity price is trending up, then the corresponding nth quantity is trending down), it is likely that the intermediate year quantity vector will lie between the monthly quantity vectors q^0 and q^t. The midyear Lowe index, $P_{Lo}(p^0,p^t,q^b)$, and the Laspeyres index going from month 0 to t, $P_L(p^0,p^t,q^0)$, will still satisfy the exact relationship given by equation (26) above. Thus $P_{Lo}(p^0,p^t,q^b)$ will equal $P_L(p^0,p^t,q^0)$ plus the covariance term $[\sum_{n=1}^N (r_n - r^*)(t_n - t^*)s_n^0\,]/Q_L(q^0,q^b,p^0)$, where $Q_L(q^0,q^b,p^0)$ is the Laspeyres quantity index going from month 0 to base year b. This covariance term is likely to be *negative* so that

$$(30)\quad P_{Lo}(p^0,p^t,q^b) < P_L(p^0,p^t,q^0).$$

To see why this covariance is likely to be negative, suppose that there is a long term upward trend in the price of commodity n so that $r_n - r^* \equiv (p_n^t/p_n^0) - r^*$ is positive. With

[14] This concept can be traced to Peter Hill (1998; 46): "When inflation has to be measured over a specified sequence of years, such as a decade, a pragmatic solution to the problems raised above would be to take the middle year as the base year. This can be justified on the grounds that the basket of goods and services purchased in the middle year is likely to be much more representative of the pattern of consumption over the decade as a whole than baskets purchased in either the first or the last years. Moreover, choosing a more representative basket will also tend to reduce, or even eliminate, any bias in the rate of inflation over the decade as a whole as compared with the increase in the CoL index." Thus in addition to introducing the concept of a midyear index, Hill also introduced the terminology *representativity bias*. Baldwin (1990; 255-256) also introduced the term *representativeness*: "Here representativeness [in an index number formula] requires that the weights used in any comparison of price levels are related to the volume of purchases in the periods of comparison." However, this basic idea dates back to Walsh (1901; 104) (1921a; 90). Baldwin (1990; 255) also noted that his concept of representativeness was the same as Drechsler's (1973; 19) concept of *characteristicity*. For additional material on *midyear indexes*, see Schultz (1999) and Okamoto (2001). Note that the midyear index concept could be viewed as a close competitor to Walsh's (1901; 431) multiyear fixed basket index where the quantity vector was chosen to be an arithmetic or geometric average of the quantity vectors in the span of periods under consideration.

normal consumer substitution responses, q_n will tend to decrease relatively over time and since q_n^b is assumed to be between q_n^0 and q_n^t, q_n^b/q_n^0 less an average quantity change of this type is likely to be negative. Hence $t_n - t^* \equiv (q_n^b/q_n^0) - t^*$ is likely to be negative. Thus, the covariance is likely to be negative under these circumstances. *Thus under the assumptions that the quantity base year falls between months 0 and t and that there are long term trends in prices and normal consumer substitution responses, the Laspeyres index will normally be larger than the corresponding Lowe midyear index*, with the divergence likely growing as month t becomes more distant from month 0.

It can also be seen that under the above assumptions, the midyear Lowe index is likely to be greater than the Paasche index between months 0 and t; i.e.,

(31) $P_{Lo}(p^0,p^t,q^b) > P_P(p^0,p^t,q^t)$.

To see why the above inequality is likely to hold, think of q^b starting at the month 0 quantity vector q^0 and then trending smoothly to the month t quantity vector q^t. When $q^b = q^0$, the Lowe index $P_{Lo}(p^0,p^t,q^b)$ becomes the Laspeyres index $P_L(p^0,p^t,q^0)$. When $q^b = q^t$, the Lowe index $P_{Lo}(p^0,p^t,q^b)$ becomes the Paasche index $P_P(p^0,p^t,q^t)$. Under the assumption of trending prices and normal substitution responses to these trending prices, it was shown earlier that the Paasche index will be less than the corresponding Laspeyres price index; i.e., that $P_P(p^0,p^t,q^t)$ was less than $P_L(p^0,p^t,q^0)$; recall (22). Thus under the assumption of smoothly trending prices and quantities between months 0 and t, and *assuming that q^b is between q^0 and q^t*, we will have

(32) $P_P(p^0,p^t,q^t) < P_{Lo}(p^0,p^t,q^b) < P_L(p^0,p^t,q^0)$.

Thus if the base year for the Lowe index is chosen to be in between the base month for the prices, month 0, and the current month for prices, month t, and there are trends in prices with corresponding trends in quantities that correspond to normal consumer substitution effects, then the resulting Lowe index is likely to lie between the Paasche and Laspeyres indexes going from months 0 to t. If the trends in prices and quantities are smooth, then choosing the base year half way between periods 0 and t should give a Lowe index that is approximately half way between the Paasche and Laspeyres indexes and hence will be very close to an ideal target index between months 0 and t. This basic idea has been implemented by Okamoto (2001) using Japanese consumer data and he found that the resulting midyear indexes approximated the corresponding Fisher ideal indexes very closely. However, the assumption of smooth trends in prices and quantities is necessary to get this close approximation.

All of the inequalities derived in this chapter rest on the assumption of long term trends in prices (and corresponding economic responses in quantities). If there are no systematic long run trends in prices, but only random fluctuations around a common trend in all prices, then the above inequalities are not valid and the Lowe index using a prior base year will probably provide a perfectly adequate approximation to both the Paasche and Laspeyres indices. However, there are some reasons for believing that there are some long run trends in prices. Some of these reasons are:

- The computer chip revolution of the past 40 years has led to strong downward trends in the prices of products that use these chips intensively. As new uses for chips have been developed over the years, the share of products that are chip intensive has grown and this implies that what used to be a relatively minor problem has become a more major problem.
- Other major scientific advances have had similar effects. For example, the invention of fiber optic cable (and lasers) has led to a downward trend in telecommunications prices as obsolete technologies based on copper wire are gradually replaced.
- Since the end of World War II, there have been a series of international trade agreements that have dramatically reduced tariffs around the world. These reductions, combined with improvements in transportation technologies, have led to a very rapid growth of international trade and remarkable improvements in international specialization. Manufacturing activities in the more developed economies have gradually been outsourced to lower wage countries, leading to deflation in goods prices in most countries around the world. However, many services cannot be readily outsourced[15] and so on average, the price of services trends upwards while the price of goods trends downwards.
- At the microeconomic level, there are tremendous differences in growth rates of firms. Successful firms expand their scale, lower their costs, and cause less successful competitors to wither away with their higher prices and lower volumes. This leads to a systematic negative correlation between changes in item prices and the corresponding changes in item volumes that can be very large indeed.

Thus there is some a priori basis for assuming long run divergent trends in prices and hence there is some basis for concern that a Lowe index that uses a distant base year for quantity weights that is prior to the base month for prices may be upward biased, compared to a more ideal target index.

6. The Young Index

Recall the definitions for the base year quantities, q_n^b, and the base year prices, p_n^b, (5) and (6) above. The *base year expenditure shares* s_n^b can be defined in the usual way as follows:

$$(33)\quad s_n^b \equiv p_n^b q_n^b / \sum_{k=1}^N p_k^b q_k^b \,; \qquad n = 1,\ldots,N.$$

Define the vector of base year expenditure shares in the usual way as $s^b \equiv [s_1^b,\ldots,s_N^b]$. These base year expenditure shares were used to provide an alternative formula for the base year b Lowe price index going from month 0 to t defined in (8) as $P_{Lo}(p^0,p^t,q^b) = \sum_{n=1}^N (p_n^t/p_n^b)\, s_n^b / \sum_{n=1}^N (p_n^0/p_n^b)\, s_n^b$. Rather than using this index as their short run target index, many statistical agencies use the following closely related *Young price index*:

[15] However some services can be internationally outsourced; e.g., call centers, computer programming, airline maintenance, etc.

(34) $P_Y(p^0,p^t,s^b) \equiv \sum_{n=1}^N (p_n^t/p_n^0)\, s_n^b$.

This type of index was first defined by the English economist, Arthur Young (1812).[16] Note that there is a change in focus when the Young index is used compared to the other indexes proposed earlier in this chapter. Up to this point, the indexes proposed have been of the fixed basket type (or averages of such indexes) where a *commodity basket* that is somehow representative for the two periods being compared is chosen and then "purchased" at the prices of the two periods and the index is taken to be the ratio of these two costs. On the other hand, for the Young index, one instead chooses *representative expenditure shares* that pertain to the two periods under consideration and then uses these shares to calculate the overall index as a share weighted average of the individual price ratios, p_n^t/p_n^0. Note that this share weighted average of price ratios view of index number theory is a bit different from the view taken at the beginning of this chapter, which viewed the index number problem as the problem of decomposing a value ratio into the product of two terms, one of which expresses the amount of price change between the two periods and the other which expresses the amount of quantity change.[17]

Statistical agencies sometimes regard the Young index defined above as an approximation to the Laspeyres price index $P_L(p^0,p^t,q^0)$. Hence, it is of interest to see how the two indexes compare. Defining the long term monthly price relatives going from month 0 to t as $r_n \equiv p_n^t/p_n^0$ and using definitions (34) and (16):

$$
\begin{aligned}
(35)\ P_Y(p^0,p^t,s^b) - P_L(p^0,p^t,s^0) &= \textstyle\sum_{n=1}^N (p_n^t/p_n^0)\, s_n^b - \sum_{n=1}^N (p_n^t/p_n^0)\, s_n^0 \\
&= \textstyle\sum_{n=1}^N (p_n^t/p_n^0)[s_n^b - s_n^0] \\
&= \textstyle\sum_{n=1}^N r_n[s_n^b - s_n^0] \qquad \text{using definitions (15)} \\
&= \textstyle\sum_{n=1}^N [r_n - r^*][s_n^b - s_n^0] + r^*\sum_{n=1}^N [s_n^b - s_n^0] \\
&= \textstyle\sum_{n=1}^N [r_n - r^*][s_n^b - s_n^0]
\end{aligned}
$$

since $\sum_{n=1}^N s_n^b = \sum_{n=1}^N s_n^0 = 1$ and defining $r^* \equiv \sum_{n=1}^N s_n^0 r_n = P_L(p^0,p^t,q^0)$. Thus the Young index $P_Y(p^0,p^t,s^b)$ is equal to the Laspeyres index $P_L(p^0,p^t,q^0)$ plus the *pseudo*

[16] The attribution of this formula to Young is due to Walsh (1901; 536) (1932; 657).

[17] Fisher's 1922 book is famous for developing the value ratio decomposition approach to index number theory but his introductory chapters took the share weighted average point of view: "An index number of prices, then shows the *average percentage change* of prices from one point of time to another." Irving Fisher (1922; 3). Fisher went on to note the importance of economic weighting: "The preceding calculation treats all the commodities as equally important; consequently, the average was called 'simple'. If one commodity is more important than another, we may treat the more important as though it were two or three commodities, thus giving it two or three times as much 'weight' as the other commodity." Irving Fisher (1922; 6). Walsh (1901; 430-431) considered both approaches: "We can either (1) draw some average of the total money values of the classes during an epoch of years, and with weighting so determined employ the geometric average of the price variations [ratios]; or (2) draw some average of the mass quantities of the classes during the epoch, and apply to them Scrope's method." Scrope's method is the same as using the Lowe index. Walsh (1901; 88-90) consistently stressed the importance of weighting price ratios by their economic importance (rather than using equally weighted averages of price relatives). Both the value ratio decomposition approach and the share weighted average approach to index number theory were studied from the axiomatic perspective in Chapter 3.

covariance between the difference in the annual shares pertaining to year b and the month 0 shares, $s_n^b - s_n^0$, and the deviations of the relative prices from their mean, $r_n - r^*$.[18]

It is no longer possible to guess at what the likely sign of the covariance term is. The question is no longer whether the *quantity* demanded goes down as the price of commodity n goes up (the answer to this question is usually yes) but the new question is: does the *share* of expenditure on commodity n go down as the price of commodity n goes up? The answer to this question depends on the elasticity of demand for the product. However, let us provisionally assume that there are long run trends in commodity prices and if the trend in prices for commodity n is above the mean, then the expenditure share for the commodity trends *down* (and vice versa). Thus we are assuming high elasticities or very strong substitution effects. Assuming also that the base year b is prior to month 0, then under these conditions, suppose that there is a long term upward trend in the price of commodity n so that $r_n - r^* \equiv (p_n^t/p_n^0) - r^*$ is positive. With the assumed very elastic consumer substitution responses, s_n will tend to decrease relatively over time and since s_n^b is assumed to be prior to s_n^0, s_n^0 is expected to be less than s_n^b or $s_n^b - s_n^0$ will likely be positive. Thus, the covariance is likely to be *positive* under these circumstances. *Hence with long run trends in prices and very elastic responses of consumers to price changes, the Young index is likely to be greater than the corresponding Laspeyres index.*

Assume that there are long run trends in commodity prices. Suppose the trend in price for commodity n is above the mean, and suppose that the expenditure share for the commodity trends *up* (and vice versa). Thus we are assuming low elasticities or very weak substitution effects. Assume also that the base year b is prior to month 0 and since we have supposed that there is a long term upward trend in the price of commodity n, then $r_n - r^* \equiv (p_n^t/p_n^0) - r^*$ will be positive. With the assumed very inelastic consumer substitution responses, s_n will tend to increase relatively over time and since s_n^b is assumed to be prior to s_n^0, it will be the case that s_n^0 is greater than s_n^b or $s_n^b - s_n^0$ is negative. Thus, the covariance is likely to be *negative* under these circumstances. *Hence with long run trends in prices and very inelastic responses of consumers to price changes, the Young index is likely to be less than the corresponding Laspeyres index.*

The previous two paragraphs indicate that a priori, it is not known what the likely difference between the Young index and the corresponding Laspeyres index will be. If elasticities of substitution are close to one, then the two sets of expenditure shares, s_i^b and s_i^0, will be close to each other and the difference between the two indices will be close to zero. However, if monthly expenditure shares have strong seasonal components, then the annual shares s_i^b could differ substantially from the monthly shares s_i^0.

It is useful to have a formula for updating the previous month's Young price index using just month over month price relatives. The Young index for month t+1, $P_Y(p^0,p^{t+1},s^b)$, can be written in terms of the Young index for month t, $P_Y(p^0,p^t,s^b)$, and an updating factor as follows:

[18] Strictly speaking, the covariance between the vectors r and $[s^b-s^0]$ is $(1/N)[r-r^*1_N]\bullet[s^b-s^0]$; i.e., the weighting factor (1/N) is missing in (35).

$$(36)\ P_Y(p^0,p^{t+1},s^b) \equiv \sum_{n=1}^N (p_n^{t+1}/p_n^0)\, s_n^b$$
$$= P_Y(p^0,p^t,s^b) \sum_{n=1}^N (p_n^{t+1}/p_n^0)\, s_n^b / \sum_{n=1}^N (p_n^t/p_n^0)\, s_n^b$$
$$= P_Y(p^0,p^t,s^b) \sum_{n=1}^N (p_n^{t+1}/p_n^t)(p_n^t/p_n^0)\, s_n^b / \sum_{n=1}^N (p_n^t/p_n^0)\, s_n^b$$
$$= P_Y(p^0,p^t,s^b) \sum_{n=1}^N (p_n^{t+1}/p_n^t)\, s_n^{b0t}$$

where the *hybrid weights* s_n^{b0t} are defined as

$$(37)\ s_n^{b0t} \equiv (p_n^t/p_n^0)\, s_n^b / \sum_{n=1}^N (p_n^t/p_n^0)\, s_n^b\ ; \qquad n = 1,..,N.$$

Thus the hybrid weights s_n^{b0t} can be obtained from the base year expenditure shares s_n^b by updating them; i.e., by multiplying them by the price relatives, (or indexes at higher levels of aggregation), p_n^t/p_n^0. Thus the required updating factor, going from month t to month t+1, is the chain link index, $\sum_{n=1}^N s_n^{b0t} (p_n^{t+1}/p_n^t)$, which uses the hybrid share weights s_n^{b0t} defined by (37).

Even if the Young index provides a close approximation to the corresponding Laspeyres index, it is difficult to recommend the use of the Young index as a final estimate of the change in prices going from period 0 to t, just as it was difficult to recommend the use of the Laspeyres index as the *final* estimate of inflation going from period 0 to t. Recall that the problem with the Laspeyres index was its lack of symmetry in the treatment of the two periods under consideration; i.e., using the justification for the Laspeyres index as a good fixed basket index, there was an identical justification for the use of the Paasche index as an equally good fixed basket index to compare periods 0 and t. The Young index suffers from a similar lack of symmetry with respect to the treatment of the base period. The problem can be explained as follows. The Young index, $P_Y(p^0,p^t,s^b)$ defined by (34) calculates the price change between months 0 and t treating month 0 as the base. But there is no particular reason to treat month 0 as the base month other than convention. Hence, if we treat month t as the base and use the same formula to measure the price change from month t back to month 0, the index $P_Y(p^t,p^0,s^b) = \sum_{n=1}^N s_n^b(p_n^0/p_n^t)$ would be appropriate. This estimate of price change can then be made comparable to the original Young index by taking its reciprocal, leading to the following *rebased Young index*[19], $P_Y{}^*(p^0,p^t,s^b)$, defined as follows:

$$(38)\ P_Y{}^*(p^0,p^t,s^b) \equiv 1/ \sum_{n=1}^N (p_n^0/p_n^t)\, s_n^b$$
$$= [\sum_{n=1}^N s_n^b\, (p_n^t/p_n^0)^{-1}]^{-1}.$$

Thus the rebased Young index, $P_Y{}^*(p^0,p^t,s^b)$, that uses the current month as the initial base period is a *share weighted harmonic mean* of the price relatives going from month 0 to month t, whereas the original Young index, $P_Y(p^0,p^t,s^b)$, is a *share weighted arithmetic mean* of the same price relatives.

Fisher argued as follows that an index number formula should give the same answer no matter which period was chosen as the base:

---

[19] Using Fisher's (1922; 118) terminology, $P_Y{}^*(p^0,p^t,s^b) \equiv 1/[P_Y(p^t,p^0,s^b)]$ is the *time antithesis* of the original Young index, $P_Y(p^0,p^t,s^b)$.

"Either one of the two times may be taken as the 'base'. Will it make a difference which is chosen? Certainly, it *ought* not and our Test 1 demands that it shall not. More fully expressed, the test is that the formula for calculating an index number should be such that it will give the same ratio between one point of comparison and the other point, *no matter which of the two is taken as the base*." Irving Fisher (1922; 64).

Problem

1. Show that the Young index and its rebased counterpart satisfy the following inequality:

(39) $P_Y^*(p^0,p^t,s^b) \leq P_Y(p^0,p^t,s^b)$

with a strict inequality provided that the period t price vector p^t is not proportional to the period 0 price vector p^0.[20] Thus a statistical agency that uses the direct Young index $P_Y(p^0,p^t,s^b)$ will generally show a higher inflation rate than a statistical agency that uses the same raw data but uses the rebased Young index, $P_Y^*(p^0,p^t,s^b)$.

The inequality (39) does not tell us by how much the Young index will exceed its rebased time antithesis. However, it can be shown that to the accuracy of a certain second order Taylor series approximation, the following relationship holds between the direct Young index and its time antithesis:

(40) $P_Y(p^0,p^t,s^b) \approx P_Y^*(p^0,p^t,s^b) + P_Y(p^0,p^t,s^b)\ \mathrm{Var}\ e$

where Var e is defined as

(41) $\mathrm{Var}\ e \equiv \sum_{n=1}^N s_n^b[e_n - e^*]^2$.

The *deviations* e_n are defined by

(42) $1+e_n \equiv r_n/r^*$; $\quad n = 1,\ldots,N$

where the r_n and their weighted mean r^* are defined by (43) and (44) below:

(43) $r_n \equiv p_n^t/p_n^0$; $\quad n = 1,\ldots,N$;

(44) $r^* \equiv \sum_{n=1}^N s_n^b r_n$

[20] Walsh (1901; 330-332) explicitly noted the inequality (39) and also noted that the corresponding geometric average would fall between the harmonic and arithmetic averages. Walsh (1901; 432) computed some numerical examples of the Young index and found big differences between it and his "best" indexes, even using weights that were representative for the periods being compared. Recall that the Lowe index becomes the Walsh index when geometric mean quantity weights are chosen and so the Lowe index can perform well when representative weights are used. This is not necessarily the case for the Young index, even using representative weights. Walsh (1901; 433) summed up his numerical experiments with the Young index as follows: "In fact, Young's method, in every form, has been found to be bad."

which turns out to equal the direct Young index, $P_Y(p^0,p^t,s^b)$. The weighted mean of the e_n is defined as

(45) $e^* \equiv \sum_{n=1}^{N} s_n^{\,b}\, e_n$.

Problem

2. Show that $e^* = 0$.

Looking at (40), we see that *the more dispersion there is in the price relatives* p_n^t/p_n^0*, to the accuracy of a second order approximation, the more the direct Young index will exceed its counterpart that uses month t as the initial base period rather than month 0.*

We indicate how the result (40) can be established.

The direct Young index, $P_Y(p^0,p^t,s^b)$, and its time antithesis, $P_Y^*(p^0,p^t,s^b)$, can be written as functions of r^*, the weights $s_n^{\,b}$ and the deviations of the price relatives e_n as follows:

(46) $P_Y(p^0,p^t,s^b) = \sum_{n=1}^{N} s_n^{\,b} r_n = r^*$;

(47) $P_Y^*(p^0,p^t,s^b) \equiv [\sum_{n=1}^{N} s_n^{\,b}\, (p_n^t/p_n^0)^{-1}]^{-1}$

$= [\sum_{n=1}^{N} s_n^{\,b}\, (r_n)^{-1}]^{-1}$ using (43)

$= r^*[\sum_{n=1}^{N} s_n^{\,b}\, (1+e_n)^{-1}]^{-1}$ using (42)

$\equiv r^*\, f(e_1,e_2,\ldots,e_N)$.

Problem

3. Calculate the second order Taylor series approximation to f(e) defined in (47) around the point $e = 0_N$ and show that it simplifies to $1 - \mathrm{Var}\ e$. Hence we obtain the approximate equality $P_Y^*(p^0,p^t,s^b) \approx r^*(1-\mathrm{Var}\ e)$, which is equivalent to the approximate equality (40).[21]

Given two a priori equally plausible index number formula that give different answers, such as the Young index and its time antithesis, Fisher (1922; 136) generally suggested taking the geometric average of the two indexes[22] and a benefit of this averaging is that the resulting formula will satisfy the time reversal test. Thus rather than using *either* the base period 0 Young index, $P_Y(p^0,p^t,s^b)$, *or* the current period t Young index,

[21] This type of second order approximation is due to Dalén (1992; 143) for the case $r^* = 1$ and to Diewert (1995; 29) for the case of a general r^*.

[22] "We now come to a third use of these tests, namely, to 'rectify' formulae, i.e., to derive from any given formula which does not satisfy a test another formula which does satisfy it; …. This is easily done by 'crossing', that is, by averaging antitheses. If a given formula fails to satisfy Test 1 [the time reversal test], its time antithesis will also fail to satisfy it; but the two will fail, as it were, in opposite ways, so that a cross between them (obtained by *geometrical* averaging) will give the golden mean which does satisfy." Irving Fisher (1922; 136). Actually the basic idea behind Fisher's rectification procedure was suggested by Walsh, who was a discussant for Fisher (1921), where Fisher gave a preview of his 1922 book: "We merely have to take any index number, find its antithesis in the way prescribed by Professor Fisher, and then draw the geometric mean between the two." Correa Moylan Walsh (1921b; 542).

$P_Y*(p^0,p^t,s^b)$, which is always below the base period 0 Young index if there is any dispersion in relative prices, it seems preferable to use the following index, which is the *geometric average* of the two alternatively based Young indexes:[23]

(48) $P_Y**(p^t,p^0,s^b) \equiv [P_Y(p^t,p^0,s^b)\ P_Y*(p^t,p^0,s^b)]^{1/2}$.

If the base year shares s_n^b happen to coincide with both the month 0 and month t shares, s_n^0 and s_n^t respectively, it can be seen that the time rectified Young index $P_Y**(p^0,p^t,s^b)$ defined by (48) will coincide with the Fisher ideal price index between months 0 and t, $P_F(p^0,p^t,q^0,q^t)$ (which will also equal the Laspeyres and Paasche indexes under these conditions). Note also that the index P_Y** defined by (48) can be produced on a timely basis by a statistical agency.

7. An Economic Approach to the Monthly Lowe Index with Annual Weights[24]

Recall the definition of the Lowe index, $P_{Lo}(p^0,p^1,q)$ defined by (1) above. We noted earlier that this formula is frequently used by statistical agencies as a target index for a CPI. We also noted that while the price vectors p^0 (the base period price vector) and p^1 (the current period price vector) were *monthly or quarterly* price vectors, the quantity vector $q \equiv (q_1,q_2,\ldots,q_N)$ which appeared in this basket type formula was usually taken to be an *annual* quantity vector that referred to a *base year*, b say, that is prior to the base period for the prices, month 0. Thus, typically, a statistical agency will produce a Consumer Price Index at a monthly frequency that has the form $P_{Lo}(p^0,p^t,q^b)$, where p^0 is the price vector pertaining to the base period month for prices, month 0, p^t is the price vector pertaining to the current period month for prices, month t say, and q^b is a reference basket quantity vector that refers to the base year b, which is equal to or prior to month 0.[25] The question to be addressed in the present section is: can this index be related to one based on the economic approach to index number theory?

Assume that the consumer has preferences defined over consumption vectors $q \equiv [q_1,\ldots,q_N]$ that can be represented by the continuous increasing utility function $f(q)$. Thus if $f(q^1) > f(q^0)$, then the consumer prefers the consumption vector q^1 to q^0. Let q^b be the annual consumption vector for the consumer in the base year b. Define the base year utility level u^b as the utility level that corresponds to $f(q)$ evaluated at q^b:

(48) $u^b \equiv f(q^b)$.

For any vector of positive commodity prices $p \equiv [p_1,\ldots,p_N]$ and for any feasible utility level u, the consumer's *cost function*, $C(u,p)$,can be defined in the usual way as the minimum expenditure required to achieve the utility level u when facing the prices p:

---

[23] This index is a base year weighted counterpart to an equally weighted index proposed by Carruthers, Sellwood and Ward (1980; 25) and Dalén (1992; 140) in the context of elementary index number formulae. See Chapter 20 for further discussion of this unweighted index.

[24] The material in this section is based on joint work with Bert Balk.

[25] As noted earlier, month 0 is called the price reference period and year b is called the weight reference period.

(49) $C(u,p) \equiv \min_q \{\sum_{n=1}^N p_nq_n : f(q_1,\ldots,q_N) = u\}$.

Let $p^b \equiv [p_1^b,\ldots,p_N^b]$ be the vector of annual prices that the consumer faced in the base year b. Assume that the observed base year consumption vector $q^b \equiv [q_1^b,\ldots,q_N^b]$ solves the following base year cost minimization problem:

(50) $C(u^b,p^b) \equiv \min_q \{\sum_{n=1}^N p_n^bq_n : f(q_1,\ldots,q_N) = u^b\} = \sum_{n=1}^N p_n^bq_n^b$.

The cost function will be used below in order to define the consumer's *cost of living price index*.

Let p^0 and p^t be the *monthly* price vectors that the consumer faces in months 0 and t. Then the *Konüs* (1924) *true cost of living index*, $P_K(p^0,p^t,q^b)$, between months 0 and t, using the base year utility level $u^b = f(q^b)$ as the reference standard of living, is defined as the following ratio of minimum monthly costs of achieving the utility level u^b:

(51) $P_K(p^0,p^t,q^b) \equiv C[f(q^b),p^t)]/C[f(q^b),p^0)]$.

Using the definition of the monthly cost minimization problem that corresponds to the cost $C[f(q^b),p^t]$, it can be seen that the following inequality holds:

(52) $C[f(q^b),p^t)] \equiv \min_q \{\sum_{n=1}^N p_n^tq_n : f(q_1,\ldots,q_N) = f(q_1^b,\ldots,q_N^b)\}$
$\leq \sum_{n=1}^N p_n^tq_n^b$

since the base year quantity vector q^b is feasible for the cost minimization problem. Similarly, using the definition of the monthly cost minimization problem that corresponds to the month 0 cost $C(f(q^b),p^0)$, it can be seen that the following inequality holds:

(53) $C[f(q^b),p^0)] \equiv \min_q \{\sum_{n=1}^N p_n^0q_n : f(q_1,\ldots,q_N) = f(q_1^b,\ldots,q_N^b)\}$
$\leq \sum_{n=1}^N p_n^0q_n^b$

since the base year quantity vector q^b is feasible for the cost minimization problem.

It will prove useful to rewrite the two inequalities (52) and (53) as equalities. This can be done if nonnegative *substitution bias terms*, e^t and e^0, are subtracted from the right hand sides of these two inequalities. Thus (52) and (53) can be rewritten as follows:

(54) $C(u^b,p^t) = \sum_{n=1}^N p_n^tq_n^b - e^t$;
(55) $C(u^b,p^0) = \sum_{n=1}^N p_n^0q_n^b - e^0$

where $e^0 \geq 0$ and $e^t \geq 0$. Using (54) and (55) and the definition (8) for the Lowe index, the following approximate equality for the Lowe index results:

(56) $P_{Lo}(p^0,p^t,q^b) \equiv \sum_{n=1}^N p_n^tq_n^b/\sum_{n=1}^N p_n^0q_n^b$
$= [C(u^b,p^t) + e^t]/[C(u^b,p^0) + e^0]$

$$\approx C(u^b,p^t)/C(u^b,p^0)$$
$$\equiv P_K(p^0,p^t,q^b) \qquad \text{using definition (51).}$$

Thus if the nonnegative substitution bias terms e^0 and e^t are small, then the Lowe index between months 0 and t, $P_{Lo}(p^0,p^t.q^b)$, will be an adequate approximation to the true cost of living index between months 0 and t, $P_K(p^0,p^t,q^b)$.[26]

A bit of algebraic manipulation shows that the Lowe index will be *exactly* equal to its cost of living counterpart if the substitution bias terms satisfy the following relationship:[27]

$$(57)\ e^t/e^0 = C(u^b,p^t)/C(u^b,p^0) = P_K(p^0,p^t,q^b).$$

Equations (56) and (57) can be interpreted as follows: if the rate of growth in the amount of substitution bias between months 0 and t is equal to the rate of growth in the minimum cost of achieving the base year utility level u^b between months 0 and t, then the observable Lowe index, $P_{Lo}(p^0,p^t.q^b)$, will be exactly equal to its true cost of living index counterpart, $P_K(p^0,p^t,q^b)$.[28]

It is difficult to know whether condition (57) will hold or whether the substitution bias terms e^0 and e^t will be small. Thus in the following two sections, first and second order Taylor series approximations to these substitution bias terms will be developed.

8. A First Order Approximation to the Bias of the Lowe Index

The true cost of living index between months 0 and t, using the base year utility level u^b as the reference utility level, is the ratio of two unobservable costs, $C(u^b,p^t)/C(u^b,p^0)$. However, both of these hypothetical costs can be approximated by first order Taylor series approximations that can be evaluated using observable information on prices and base year quantities. The first order Taylor series approximation to $C(u^b,p^t)$ around the annual base year price vector p^b is given by the following approximate equation:[29]

$$(58)\ C(u^b,p^t) \approx C(u^b,p^b) + \textstyle\sum_{n=1}^N [\partial C(u^b,p^b)/\partial p_n][p_n^t - p_n^b]$$
$$= \textstyle\sum_{n=1}^N p_n^b q_n^b + \sum_{n=1}^N [\partial C(u^b,p^b)/\partial p_n][p_n^t - p_n^b] \qquad \text{using (50)}$$
$$= \textstyle\sum_{n=1}^N p_n^b q_n^b + \sum_{n=1}^N q_n^b[p_n^t - p_n^b] \qquad \text{using Shephard's Lemma}$$
$$= \textstyle\sum_{n=1}^N p_n^t q_n^b .$$

---

[26] Although $P_K(p^0,p^t,q^b)$ is a true cost of living index, it may not be a very relevant one if the base year consumption vector q^b is rather far from the consumption vectors that pertain to months 0 and t. This limitation of the analysis must be kept in mind.
[27] This assumes that e^0 is greater than zero. If e^0 is equal to zero, then to have equality of P_K and P_{Lo}, it must be the case that e^t is also equal to zero.
[28] It can be seen that when month t is set equal to month 0, $e^t = e^0$ and $C(u^b,p^t) = C(u^b,p^0)$ and thus (57) is satisfied and $P_{Lo} = P_K$. This is not surprising since both indexes are equal to unity when t = 0.
[29] This type of Taylor series approximation was used in Schultze and Mackie (2002; 91) in the cost of living index context but it essentially dates back to Hicks (1941-42; 134) in the consumer surplus context. See also Diewert (1992; 568) and Hausman (2002; 8).

Similarly, the first order Taylor series approximation to $C(u^b,p^0)$ around the annual base year price vector p^b is given by the following approximate equation:

$$(59)\ C(u^b,p^0) \approx C(u^b,p^b) + \sum_{n=1}^{N} [\partial C(u^b,p^b)/\partial p_n][p_n^0 - p_n^b]$$
$$= \sum_{n=1}^{N} p_n^b q_n^b + \sum_{n=1}^{N} [\partial C(u^b,p^b)/\partial p_n][p_n^0 - p_n^b] \quad \text{using (50)}$$
$$= \sum_{n=1}^{N} p_n^b q_n^b + \sum_{n=1}^{N} q_n^b[p_n^0 - p_n^b] \quad \text{using Shephard's Lemma}$$
$$= \sum_{n=1}^{N} p_n^0 q_n^b .$$

Comparing (58) with (54) and comparing (59) with (55), it can be seen that to the accuracy of the first order approximations used in (58) and (59), the substitution bias terms e^t and e^0 will be zero. Using these results to reinterpret (56), it can be seen that if the month 0 and month t price vectors, p^0 and p^t, are not too different from the base year vector of prices p^b, then the Lowe index $P_{Lo}(p^0,p^t,q^b)$ will approximate the true cost of living index $P_K(p^0,p^t,q^b)$ to the accuracy of a first order approximation. *This result is quite useful, since it indicates that if the monthly price vectors p^0 and p^t are just randomly fluctuating around the base year prices p^b (with modest variances), then the Lowe index will serve as an adequate approximation to a theoretical cost of living index.* However, if there are systematic long term trends in prices and month t is fairly distant from month 0 (or the end of year b is quite distant from month 0), then the first order approximations given by (58) and (59) may no longer be adequate and the Lowe index may have a considerable bias relative to its cost of living counterpart. The hypothesis of long run trends in prices will be explored in the following section.

9. A Second Order Approximation to the Substitution Bias of the Lowe Index

A second order Taylor series approximation to $C(u^b,p^t)$ around the base year price vector p^b is given by the following approximate equation:

$$(60)\ C(u^b,p^t) \approx C(u^b,p^b) + \sum_{n=1}^{N} [\partial C(u^b,p^b)/\partial p_n][p_n^t - p_n^b]$$
$$+ (1/2) \sum_{n=1}^{N} \sum_{k=1}^{N} [\partial^2 C(u^b,p^b)/\partial p_n \partial p_k][p_n^t - p_n^b][p_k^t - p_k^b]$$
$$= \sum_{n=1}^{N} p_n^t q_n^b + (1/2) \sum_{n=1}^{N} \sum_{k=1}^{N} [\partial^2 C(u^b,p^b)/\partial p_n \partial p_k][p_n^t - p_n^b][p_k^t - p_k^b]$$

where the last equality follows using (50) and Shephard's Lemma.[30] Similarly, a second order Taylor series approximation to $C(u^b,p^0)$ around the base year price vector p^b is given by the following approximate equation:

$$(61)\ C(u^b,p^0) \approx C(u^b,p^b) + \sum_{n=1}^{N} [\partial C(u^b,p^b)/\partial p_n][p_n^0 - p_n^b]$$
$$+ (1/2) \sum_{n=1}^{N} \sum_{k=1}^{N} [\partial^2 C(u^b,p^b)/\partial p_n \partial p_k][p_n^0 - p_n^b][p_k^0 - p_k^b]$$
$$= \sum_{n=1}^{N} p_n^0 q_n^b + (1/2) \sum_{n=1}^{N} \sum_{k=1}^{N} [\partial^2 C(u^b,p^b)/\partial p_n \partial p_k][p_n^0 - p_n^b][p_k^0 - p_k^b]$$

where the last equality follows using (50) and Shephard's Lemma.

[30] This type of second order approximation is due to Hicks (1941-42; 133-134) (1946; 331). See also Diewert (1992; 568), Hausman (2002; 18) and Schultze and Mackie (2002; 91). For alternative approaches to modeling substitution bias, see Diewert (1998) (2002; 598-603) and Hausman (2002).

Comparing (60) to (54), and (61) to (55), it can be seen that to the accuracy of a second order approximation, the month 0 and month t substitution bias terms, e^0 and e^t, will be equal to the following expressions involving the second order partial derivatives of the consumer's cost function $\partial^2C(u^b,p^b)/\partial p_n\partial p_k$ evaluated at the base year standard of living u^b and at the base year prices p^b:

(62) $e^0 \approx -(1/2) \sum_{n=1}^N \sum_{k=1}^N [\partial^2C(u^b,p^b)/\partial p_n\partial p_k][p_n^0 - p_n^b][p_k^0 - p_k^b]$;

(63) $e^t \approx -(1/2) \sum_{n=1}^N \sum_{k=1}^N [\partial^2C(u^b,p^b)/\partial p_n\partial p_k][p_n^t - p_n^b][p_k^t - p_k^b]$.

Since the consumer's cost function C(u,p) is a concave function in the components of the price vector p[31], we know[32] that the N by N (symmetric) matrix of second order partial derivatives $[\partial^2C(u^b,p^b)/\partial p_i\partial p_j]$ is negative semidefinite.[33] Hence, for arbitrary price vectors p^b, p^0 and p^t, the right hand sides of (62) and (63) will be nonnegative. Thus to the accuracy of a second order approximation, the substitution bias terms e^0 and e^t will be nonnegative.

Now assume that there are *long run systematic trends in prices*. Assume that the last month of the base year for quantities occurs M months prior to month 0, the base month for prices, and assume that prices trend *linearly* with time, starting with the last month of the base year for quantities. Thus assume the existence of constants α_j for j = 1,…,N such that the price of commodity j in month t is given by:

(64) $p_j^t = p_j^b + \alpha_j(M + t)$; $j = 1,\ldots,N$; $t = 0,1,2,\ldots,T$.

Substituting (64) into (62) and (63) leads to the following second order approximations to the two substitution bias terms, e^0 and e^t:[34]

(65) $e^0 \approx \gamma M^2$;
(66) $e^t \approx \gamma (M+t)^2$

where γ is defined as follows:

(67) $\gamma \equiv -(1/2) \sum_{n=1}^N \sum_{k=1}^N [\partial^2C(u^b,p^b)/\partial p_n\partial p_k] \alpha_n\alpha_k \geq 0$.

It should be noted that the parameter γ will be zero under two sets of conditions:[35]

[31] See for example Diewert (1993; 109-110).
[32] See for example Diewert (1993; 149).
[33] A symmetric N by N matrix A with ijth element equal to a_{ij} is negative semidefinite if and only if for every vector $z \equiv [z_1,\ldots,z_N]$, it is the case that $\sum_{i=1}^N \sum_{j=1}^N a_{ij} z_iz_j \leq 0$.
[34] Note that the period 0 approximate bias defined by the right hand side of (65) is fixed while the period t approximate bias defined by the right hand side of (66) increases quadratically with time t. Hence the period t approximate bias term will eventually overwhelm the period 0 approximate bias in this linear time trends case if t is allowed to become large enough.

- All of the second order partial derivatives of the consumer's cost function $\partial^2 C(u^b,p^b)/\partial p_n \partial p_k$ are equal to zero;
- Each commodity price change parameter α_j is proportional to the corresponding commodity j base year price p_j^b.[36]

The first condition is empirically unlikely since it implies that the consumer will not substitute away from commodities whose relative price has increased. The second condition is also empirically unlikely, since it implies that the structure of relative prices remains unchanged over time. Thus in what follows, it will be assumed that γ is a positive number.

In order to simplify the notation in what follows, define the denominator and numerator of the month t Lowe index, $P_{Lo}(p^0,p^t,q^b)$, as a and b respectively; i.e.; define:

(68) $a \equiv \sum_{n=1}^N p_n^0 q_n^b$;
(69) $b \equiv \sum_{n=1}^N p_n^t q_n^b$.

Using equations (64) to eliminate the month 0 prices p_n^0 from (68) and the month t prices p_i^t from (69) leads to the following expressions for a and b:

(70) $a = \sum_{n=1}^N p_n^b q_n^b + \sum_{n=1}^N \alpha_n q_n^b M$;
(71) $b = \sum_{n=1}^N p_n^b q_n^b + \sum_{n=1}^N \alpha_n q_n^b (M+t) = a + \sum_{n=1}^N \alpha_n q_n^b t$.

It is assumed that a and b are positive[37] and that

(72) $\sum_{n=1}^N \alpha_n q_n^b \geq 0$.

Assumption (72) rules out a general deflation in prices.

Define the *bias* in the month t Lowe index, B^t, as the difference between the true cost of living index $P_K(p^0,p^t,q^b)$ defined by (51) and the corresponding Lowe index $P_{Lo}(p^0,p^t,q^b)$:

(73) $B^t \equiv P_K(p^0,p^t,q^b) - P_{Lo}(p^0,p^t,q^b)$
$= [C[f(q^b),p^t)]/C[f(q^b),p^0)] - [b/a]$ using (51), (56), (70) and (71)
$= [(b - e^t)/(a - e^0)] - [b/a]$ using (54), (55), (70) and (71)
$\approx [(b - \gamma (M+t)^2)/(a - \gamma M^2)] - [b/a]$ using the approximations (65) and (66)

---

[35] A more general condition that ensures the positivity of γ is that the vector $[\alpha_1,\ldots,\alpha_N]$ is not an eigenvector of the matrix of second order partial derivatives $\partial^2 C(u^b,p^b)/\partial p_i \partial p_j$ that corresponds to a zero eigenvalue.

[36] We know that C(u,p) is linearly homogeneous in the components of the price vector p; see Diewert (1993; 109) for example. Hence using Euler's Theorem on homogeneous functions, it can be shown that p^b is an eigenvector of the matrix of second order partial derivatives $\partial^2 C(u,p)/\partial p_i \partial p_j$ that corresponds to a zero eigenvalue and thus $\sum_{i=1}^N \sum_{j=1}^N [\partial^2 C(u,p)/\partial p_i \partial p_j]\, p_i^b p_j^b = 0$; see Diewert (1993; 149) for a detailed proof of this result.

[37] We also assume that $a - \gamma M^2$ is positive.

$= [a\{b - \gamma (M^2 + 2Mt + t^2)\} - b\{a - \gamma M^2\}]/[a(a - \gamma M^2)]$
$= \gamma\{(b - a)M^2 - 2aMt - at^2\}/[a(a - \gamma M^2)]$

$= \gamma\{(\sum_{n=1}^N \alpha_n q_n^b)tM^2 - 2aMt - at^2\}/[a(a - \gamma M^2)]$ using (71)
$= -\gamma\{2aMt - (\sum_{n=1}^N \alpha_n q_n^b)tM^2 + at^2\}/[a(a - \gamma M^2)]$ rearranging terms
$= -\gamma\{2[\sum_{n=1}^N p_n^b q_n^b + \sum_{n=1}^N \alpha_n q_n^b M]Mt - (\sum_{n=1}^N \alpha_n q_n^b)tM^2 + at^2\}/[a(a - \gamma M^2)]$
using definition (70)
$= -\gamma\{2\sum_{n=1}^N p_n^b q_n^b tM + (\sum_{n=1}^N \alpha_n q_n^b)tM^2 + at^2\}/[a(a - \gamma M^2)]$
< 0

where the inequality follows from $\gamma > 0$, $a - \gamma M^2 > 0$, $\sum_{n=1}^N p_n^b q_n^b > 0$, $\sum_{n=1}^N \alpha_n q_n^b \geq 0$, $a > 0$ and $t \geq 1$. Thus for $t \geq 1$, *the Lowe index will have an upward bias* (to the accuracy of a second order Taylor series approximation) compared to the corresponding true cost of living index $P_K(p^0,p^t,q^b)$, since the approximate bias defined by the last expression in (73) is the sum of one nonpositive and two negative terms. *Moreover this approximate bias will grow quadratically in time t.*[38]

In order to give the reader some idea of the magnitude of the approximate bias B^t defined by the last line of (73), a simple special case will be considered at this point. Suppose there are only 2 commodities and at the base year, all prices and quantities are equal to 1. Thus $p_i^b = q_i^b = 1$ for $i = 1,2$ and $\sum_{i=1}^N p_i^b q_i^b = 2$. Assume that $M = 24$ so that the base year data on quantities take 2 years to process before the Lowe index can be implemented. Assume that the monthly rate of growth in price for commodity 1 is $\alpha_1 = 0.002$ so that after 1 year, the price of commodity 1 rises 0.024 or 2.4 %. Assume that commodity 2 falls in price each month with $\alpha_2 = -0.002$ so that the price of commodity 2 falls 2.4 % in the first year after the base year for quantities. Thus the relative price of the two commodities is steadily diverging by about 5 percent per year. Finally, assume that $\partial^2 C(u^b,p^b)/\partial p_1 \partial p_1 = \partial^2 C(u^b,p^b)/\partial p_2 \partial p_2 = -1$ and $\partial^2 C(u^b,p^b)/\partial p_1 \partial p_2 = \partial^2 C(u^b,p^b)/\partial p_2 \partial p_1 = 1$. These assumptions imply that the own elasticity of demand for each commodity is -1 at the base year consumer equilibrium. Making all of these assumptions means that:

(74) $2 = \sum_{n=1}^2 p_n^b q_n^b = a = b$; $\sum_{n=1}^2 \alpha_n q_n^b = 0$; $M = 24$; $\gamma = 0.000008$.

Substituting the parameter values defined in (74) into (73) leads to the following formula for the approximate amount that the Lowe index will exceed the corresponding true cost of living index at month t:

(75) $-B^t = 0.000008\,(96t + 2t^2)/\{2(2 - 0.004608)\}$.

Evaluating (75) at $t = 12$, $t = 24$, $t = 36$, $t = 48$ and $t = 60$ leads to the following estimates for $-B^t$: 0.0029 (the approximate bias in the Lowe index at the end of the first year of

[38] If M is large relative to t, then it can be seen that the first two terms in the last equation of (73) can dominate the last term, which is the quadratic in t term.

operation for the index); 0.0069 (the bias after 2 years); 0.0121 (3 years); 0.0185 (4 years); 0.0260 (5 years). Thus at the end of the first year of the operation of the Lowe index, it will only be above the corresponding true cost of living index by approximately a third of a percentage point but by the end of the fifth year of operation, it will exceed the corresponding cost of living index by about 2.6 percentage points, which is no longer a negligible amount.[39]

The numerical results in the previous paragraph are only indicative of the approximate magnitude of the difference between a cost of living index and the corresponding Lowe index. The important point to note is that to the accuracy of a second order approximation, the Lowe index will generally exceed its cost of living counterpart. However, the results also indicate that this difference can be reduced to a negligible amount if:

- the lag in obtaining the base year quantity weights is minimized and
- the base year is changed as frequently as possible.

It also should be noted that the numerical results depend on the assumption that long run trends in prices exist, which may not be true,[40] and on elasticity assumptions that may not be justified.[41] Thus statistical agencies should prepare their own carefully constructed estimates of the differences between a Lowe index and a cost of living index in the light of their own particular circumstances.

References

Baldwin, A. (1990), "Seasonal Baskets in Consumer Price Indexes", *Journal of Official Statistics* 6:3, 251-273.

Balk, B.M. (1980), "A Method for Constructing Price Indices for Seasonal Commodities", *The Journal of the Royal Statistical Society Series A* 143, 68-75.

Bean, L. H. and O. C. Stine (1924), "Four Types of Index Numbers of Farm Prices", *Journal of the American Statistical Association* 19, 30-35.

Bortkiewicz, L.v. (1923), "Zweck und Struktur einer Preisindexzahl", *Nordisk Statistisk Tidsskrift* 2, 369-408.

[39] Note that the relatively large magnitude of M compared to t leads to a bias that grows approximately linearly with t rather than quadratically.

[40] For mathematical convenience, the trends in prices were assumed to be linear rather than the more natural assumption of geometric trends in prices.

[41] Another key assumption that was used to derive the numerical results is the magnitude of the divergent trends in prices. If the price divergence vector is doubled to $\alpha_1 = 0.004$ and $\alpha_2 = -0.004$, then the parameter γ quadruples and the approximate bias will also quadruple.

Carruthers, A. G., D. J. Sellwood and P. W. Ward (1980), "Recent Developments in the Retail Prices Index", *The Statistician* 29, 1-32.

Dalén, J. (1992), "Computing Elementary Aggregates in the Swedish Consumer Price Index," *Journal of Official Statistics* 8, 129-147.

Diewert, W.E. (1992), "Exact and Superlative Welfare Change Indicators", *Economic Inquiry* 30, 565-582.

Diewert, W.E. (1993), "Duality Approaches To Microeconomic Theory", pp. 105-175 in *Essays in Index Number Theory*, Volume 1, Contributions to Economic Analysis 217, W.E. Diewert and A.O. Nakamura (eds.), Amsterdam: North-Holland.

Diewert, W. E. (1995), "Axiomatic and Economic Approaches to Elementary Price Indexes", Discussion Paper No. 95-01, Department of Economics, University of British Columbia, Vancouver, Canada.

Diewert, W.E. (1998), "Index Number Issues in the Consumer Price Index", *The Journal of Economic Perspectives* 12:1, 47-58.

Diewert, W.E. (2002), "Harmonized Indexes of Consumer Prices: Their Conceptual Foundations", *Swiss Journal of Economics and Statistics* 138, 547-637.

Drechsler, L. (1973), "Weighting of Index Numbers in Multilateral International Comparisons", *Review of Income and Wealth* 19, 17-34.

Fisher, I. (1921), "The Best Form of Index Number", *Journal of the American Statistical Association* 17, 533-537.

Fisher, I. (1922), *The Making of Index Numbers*, Houghton-Mifflin, Boston.

Hausman, J. (2002), "Sources of Bias and Solutions to Bias in the CPI", NBER Working Paper 9298, Cambridge, MA: National Bureau of Economic Research.

Hicks, J.R. (1941-42), "Consumers' Surplus and Index Numbers", *The Review of Economic Studies* 9, 126-137.

Hicks, J.R. (1946), *Value and Capital*, Second Edition, Oxford: Clarendon Press.

Hill, T. P. (1996), *Inflation Accounting: A Manual on National Accounting under Conditions of High Inflation*, Paris: OECD.

Hill, T. P. (1998), "The Measurement of Inflation and Changes in the Cost of Living", *Statistical Journal of the United Nations ECE* 15, 37-51.

Konüs, A.A. (1924), "The Problem of the True Index of the Cost of Living", translated in *Econometrica* 7, (1939), 10-29.

Lowe, J. (1823), *The Present State of England in Regard to Agriculture, Trade and Finance*, second edition, London: Longman, Hurst, Rees, Orme and Brown.

Okamoto, M. (2001), "Mid-Year Basket Index as a Practical Approximation to a Superlative Index", Paper presented at the Sixth Meeting of the International Working Group on Price Indices, Canberra, April 2-6, 2001.

Schultz, B. (1999), "Effects of Using Various Macro-Index Formulae in Longitudinal Price and Comparisons: Empirical Studies", pp. 236-249 in *Proceedings of the Fourth Meeting of the International Working Group on Price Indices*, W. Lane (ed.), Bureau of Labor Statistics, U.S. Department of Labor, Washington D.C.

Schultze, C.L. and C. Mackie (eds.) (2002), *At What Price? Conceptualizing and Measuring Cost-of Living and Price Indices*, Washington DC: National Academy Press.

Triplett, J. E. (1981), "Reconciling the CPI and the PCE Deflator", *Monthly Labor Review* (September), 3-15.

Walsh, C. M. (1901), *The Measurement of General Exchange Value*, New York: Macmillan and Co.

Walsh, C. M. (1921a), *The Problem of Estimation*, London: P.S. King & Son.

Walsh, C. M. (1921b), "Discussion", *Journal of the American Statistical Association* 17, 537-544.

Walsh, C. M. (1932), "Index Numbers", pp. 652-658 in *Encyclopedia of the Social Sciences*, Volume 7, E. R. A. Seligman (editor in chief), New York : The Macmillan Co.

Young, A. (1812), *An Inquiry into the Progressive Value of Money in England as Marked by the Price of Agricultural Products*, London.