Lowe and Cobb-Douglas CPIs and their Substitution Bias

Bert M. Balk*
Rotterdam School of Management
Erasmus University
Rotterdam
E-mail bbalk@rsm.nl
and
Statistics Netherlands
The Hague

Draft, April 16, 2009

Abstract

Balk and Diewert (2003) considered the substitution bias of a Lowe CPI; see also CPI Manual (2004), Chapter 17. The present paper considers a Cobb-Douglas (or Geometric Young) CPI, and compares the two price indices.

Keywords: Index number; cost-of-living index; Lowe index; Cobb-Douglas index; Geometric Young index.

JEL code: C43.

* The views expressed in this paper are those of the author and do not necessarily reflect any policy of Statistics Netherlands.

1 Introduction

Most if not all statistical agencies calculate their CPI as a weighted arithmetic mean of price relatives. Thus, formally, the aggregate price change between reference month 0 and current month t is defined as a weighted average of price relatives

$$\sum_{n=1}^{N} w_n(p_n^t/p_n^0) \text{ where } \sum_{n=1}^{N} w_n = 1. \tag{1}$$

The weights, ideally, reflect the importance of the various commodities in the expenditure of the representative consumer. Since the processing of expenditure data is a time-consuming undertaking, the most recent set of expenditure shares usually refers to some year b prior to month 0,

$$s_n^b \equiv \frac{p_n^b x_n^b}{\sum_{n=1}^{N} p_n^b x_n^b} \; (n = 1, ..., N). \tag{2}$$

In order to get rid of the year b prices, these expenditure shares are price-updated to month 0. Thus, the CPI weights are defined as

$$w_n \equiv \frac{p_n^b x_n^b (p_n^0/p_n^b)}{\sum_{n=1}^{N} p_n^b x_n^b (p_n^0/p_n^b)} \; (n = 1, ..., N). \tag{3}$$

But this implies that the functional form of the CPI becomes that of a so-called Lowe index:

$$\sum_{n=1}^{N} w_n(p_n^t/p_n^0) = \frac{\sum_{n=1}^{N} p_n^t x_n^b}{\sum_{n=1}^{N} p_n^0 x_n^b} = P^{Lo}(p^t, p^0; x^b). \tag{4}$$

Here the prices of month t are compared with those of month 0, using the quantities of some year b prior to month 0.

Alternatively, one could contemplate using the period b expenditure shares as weights. This would give a Young price index,

$$P^Y(p^t, p^0; s^b) \equiv \sum_{n=1}^{N} s_n^b (p_n^t/p_n^0). \tag{5}$$

The geometric analogue is

$$P^{CD}(p^t, p^0; s^b) \equiv \prod_{n=1}^{N} (p_n^t/p_n^0)^{s_n^b}, \tag{6}$$

which is known as a Cobb-Douglas price index.

Balk and Diewert (2003) considered the substitution bias of a Lowe CPI; see also CPI Manual (2004), Chapter 17. In this paper I consider a Cobb-Douglas (or Geometric Young) CPI, and compare the two price indices with respect to their substitution bias.

2 The basics

We consider a single consumer and assume that this consumer has a stable preference ordering[1] over a set of commodities labelled $1, ..., N$. Under suitable regularity conditions such an ordering can be represented by a utility function $U(x)$, that is a function such that $U(x') > U(x)$ if and only if the consumer prefers the quantity vector x' over x. Quantity vectors x are non-negative and it is assumed that $U(x)$ is non-decreasing in the components of x. A set $\{x|U(x) = u\}$ for $u \in \text{Range}\, U(x)$ is called a standard of living.

Suppose that the consumer faces the positive price vector p. Then, neoclassically, the consumer's decision problem can be formulated as

$$\min_x p \cdot x \,\text{subject to}\, U(x) \geq u. \tag{7}$$

This means that, given a certain utility level u and prices p, the consumer minimizes the cost of achieving this level. The (Hicksian) quantities demanded, $x(p, u)$, are obtained as solution of this minimization problem, and their cost is

$$C(p, u) \equiv p \cdot x(p, u) = \min_x \{p \cdot x \mid U(x) \geq u\}. \tag{8}$$

We call $C(p, u)$ the cost (or expenditure) function. Under suitable regularity conditions this function is a dual representation of the consumer's preference ordering.[2] The cost function is nondecreasing in u, concave in p, and linearly homogeneous in p. The last property means that

$$C(\lambda p, u) = \lambda C(p, u)\ (\lambda > 0). \tag{9}$$

We now consider the price vectors p^0 and p^1 pertaining to periods or situations 0 and 1 and we let $\bar{u}$ be some reference utility level. It is convenient to think of period 0 as an earlier and period 1 as a later period.

[1]The case of changing preferences is considered in Balk (1989).

[2]See Diewert (1993).

The Konüs price index or *cost of living index* is defined by

$$P^K(p^1, p^0; \bar{u}) \equiv \frac{C(p^1, \bar{u})}{C(p^0, \bar{u})}. \tag{10}$$

This is the minimum cost of achieving utility level $\bar{u}$ when the prices are p^1 relative to the minimum cost of achieving this level when the prices are p^0. The cost of living index $P^K(p^1, p^0; \bar{u})$ thus conditions on the standard of living given by $\bar{u}$.

We assume that the consumer acts cost-minimizing in periods 0 and 1[3]; that is, for the observed price and quantity vectors we assume that $x^0 = x(p^0, U(x^0))$ and $x^1 = x(p^1, U(x^1))$, so that

$$p^0 \cdot x^0 = C(p^0, U(x^0)) \tag{11}$$

and

$$p^1 \cdot x^1 = C(p^1, U(x^1)). \tag{12}$$

Using definition (8) we then obtain

$$C(p^1, U(x^0)) \leq p^1 \cdot x^0 \tag{13}$$

and

$$C(p^0, U(x^1)) \leq p^0 \cdot x^1. \tag{14}$$

Combining (11), (12) with (13), (14) and using definition (10), we obtain the following bounds for two cost of living index numbers

$$P^K(p^1, p^0; u^0) \leq \frac{p^1 \cdot x^0}{p^0 \cdot x^0} \equiv P^L(p^1, x^1, p^0, x^0) \tag{15}$$

$$P^K(p^1, p^0; u^1) \geq \frac{p^1 \cdot x^1}{p^0 \cdot x^1} \equiv P^P(p^1, x^1, p^0, x^0), \tag{16}$$

where $u^t \equiv U(x^t)$ $(t = 0, 1)$. These are the famous Laspeyres-Paasche bounds. Notice that they cannot be combined into a single equation, because of the different reference utility levels employed.

[3] This is frequently called 'rational behaviour'. On the limitations of this concept of rationality see Sen (1977).

If the utility function is linearly homogeneous, that is $U(\lambda x) = \lambda U(x)$ $(\lambda > 0)$, then $C(p, u) = uC(p, 1)$, and $P^K(.)$ becomes independent of the reference utility level $\bar{u}$.[4] In this case

$$P^P(p^1, x^1, p^0, x^0) \leq P^K(p^1, p^0; \bar{u}) \leq P^L(p^1, x^1, p^0, x^0) \tag{17}$$

for any $\bar{u}$. However, the assumption of linear homogeneity is a very restrictive one. It is equivalent to expenditure proportionality, which means that if we increase the consumer's budget by a certain factor, all the quantities consumed will increase by the same factor.

3 Lowe and CD price indices

Consider a third period b with observed prices p^b and quantities x^b. Again, it is convenient but not necessary to think of period b as preceding period 0. The Lowe price index, comparing period 1 to period 0, is then defined by

$$P^{Lo}(p^1, p^0; x^b) \equiv \frac{p^1 \cdot x^b}{p^0 \cdot x^b}. \tag{18}$$

Note that when $b = 0$, the Lowe index reduces to the Laspeyres index. For $b = 1$ the Lowe index reduces to the Paasche index.

The observed expenditure shares of period b are $s_n^b \equiv p_n^b x_n^b / p^b \cdot x^b$ ($n = 1, ..., N$). Then the Cobb-Douglas or Geometric Young price index, for period 1 relative to period 0, is defined by

$$P^{CD}(p^1, p^0; s^b) \equiv \prod_{n=1}^{N} (p_n^1 / p_n^0)^{s_n^b}. \tag{19}$$

Note that the expenditure shares add up to 1, $\sum_{n=1}^{N} s_n^b = 1$, so that $P^{CD}(.)$ is a weighted geometric mean of price relatives.[5] For $b = 0$, the Cobb-Douglas index reduces to the Geometric Laspeyres index. For $b = 1$ the Cobb-Douglas index reduces to the Geometric Paasche index.

Before proceeding to the discussion of their substitution bias, that is, the relation of these two price indices to some cost of living index $P^K(p^1, p^0; u)$,

[4]This requires the utility function to be positive. However, since the utility function is determined up to a monotonic transformation, there is no lack of generality to assume this to be the case.

[5]The Young price index, dating back to 1812, is the arithmetic analogue.

it is interesting to consider their mutual relation. Define the hybrid period 0 prices period b quantities expenditure shares by

$$s_n^{0b} \equiv \frac{p_n^0 x_n^b}{p^0 \cdot x^b} \; (n = 1, ..., N). \tag{20}$$

The Lowe price index can then be written as a weighted arithmetic mean

$$P^{Lo}(p^1, p^0; x^b) = \sum_{n=1}^{N} s_n^{0b} (p_n^1 / p_n^0). \tag{21}$$

Using the logarithmic mean[6], it then appears that the logarithm of the Lowe price index can be written as

$$\ln P^{Lo}(p^1, p^0; x^b) = \sum_{n=1}^{N} \frac{s_n^{0b} L(P^{Lo}(p^1, p^0; x^b), p_n^1/p_n^0)}{\sum_{n=1}^{N} s_n^{0b} L(P^{Lo}(p^1, p^0; x^b), p_n^1/p_n^0)} \ln(p_n^1/p_n^0). \tag{22}$$

The logarithm of the Cobb-Douglas price index is

$$\ln P^{CD}(p^1, p^0; s^b) = \sum_{n=1}^{N} s_n^b \ln(p_n^1/p_n^0). \tag{23}$$

Subtracting (23) from (22) and making use of the linear homogeneity property of the logmean we obtain

$$\begin{aligned} &\ln P^{Lo}(p^1, p^0; x^b) - \ln P^{CD}(p^1, p^0; s^b) = \\ &\quad \sum_{n=1}^{N} s_n^b \left(\frac{a_n^{10b}}{\sum_{n=1}^{N} s_n^b a_n^{10b}} - 1 \right) \ln \left(\frac{p_n^1/p_n^0}{P^{CD}(p^1, p^0; s^b)} \right), \end{aligned} \tag{24}$$

where $a_n^{10b} \equiv L(P^{Lo}(p^1, p^0; x^b) p_n^0/p_n^b, p_n^1/p_n^b)$ $(n = 1, ..., N)$ are price-update factors. The right-hand side of expression (24) has the structure of a weighted covariance, namely between relative price-update factors and relative price changes. Going from period b to period 1 the expenditure shares s_n^b are price-updated by the factors a_n^{10b}. The relative price-update factors are then

[6] The logarithmic mean of any two strictly positive real numbers a and b is defined by $L(a, b) \equiv (a - b)/\ln(a/b)$ if $a \neq b$ and $L(a, a) \equiv a$. It has the following properties: (1) $\min(a, b) \leq L(a, b) \leq \max(a, b)$; (2) $L(a, b)$ is continuous; (3) $L(\lambda a, \lambda b) = \lambda L(a, b)$ $(\lambda > 0)$; (4) $L(a, b) = L(b, a)$; (5) $(ab)^{1/2} \leq L(a, b) \leq (a + b)/2$; (6) $L(a, 1)$ is concave.

$a_n^{10b}/\sum_{n=1}^N s_n^b a_n^{10b}$, their weighted mean being equal to 1. The covariance is then between these factors, covering the time span from period b to 1, and relative price changes, covering the time span from period 0 to 1. When the time interval $[b, 0]$ is short relative to $[0, 1]$ this covariance is positive. Without empirical material, however, not much can be said about the sign of the covariance, and hence about the relative position of the Lowe and the Cobb-Douglas price indices.

4 Substitution bias of the Lowe price index

For defining the substitution bias of a price index we must pick the cost of living index that is to serve as our target. Since the Lowe price index $P^{Lo}(p^1, p^0; x^b)$ conditions on the quantity vector x^b it is natural to compare this index to the cost of living index $P^K(p^1, p^0; U(x^b))$ which conditions on the standard of living represented by x^b. Thus, the substitution bias of the Lowe price index is here defined as the difference

$$P^{Lo}(p^1, p^0; x^b) - P^K(p^1, p^0; u^b) \tag{25}$$

where $u^b \equiv U(x^b)$. Balk and Diewert (2003) employed second-order Taylor series approximations to explore the substitution bias. Instead of repeating this, I am using here their exact counterparts.[7] The first concerns the numerator of the cost of living index and reads

$$\begin{aligned} & C(p^1, u^b) \\ & = C(p^b, u^b) + \sum_{n=1}^N \frac{\partial C(p^b, u^b)}{\partial p_n}(p_n^1 - p_n^b) \\ & \quad +(1/2)\sum_{n=1}^N \sum_{n'=1}^N \frac{\partial^2 C(p^*, u^b)}{\partial p_n \partial p_{n'}}(p_n^1 - p_n^b)(p_{n'}^1 - p_{n'}^b) \\ & = \sum_{n=1}^N p_n^1 x_n^b + (1/2)\sum_{n=1}^N \sum_{n'=1}^N \frac{\partial^2 C(p^*, u^b)}{\partial p_n \partial p_{n'}}(p_n^1 - p_n^b)(p_{n'}^1 - p_{n'}^b), \end{aligned} \tag{26}$$

where $p^* \in [p^b, p^1]$. The expression after the second equality sign was ob-

[7] Following the suggestion of a referee of the earlier paper.

tained by using Shephard's Lemma[8] and the assumption that the consumer acts cost minimizing in period b, that is, $x^b = x(p^b, u^b)$ and therefore

$$p^b \cdot x^b = C(p^b, U(x^b)). \tag{27}$$

Likewise, the second expression concerns the denominator of the cost of living index and reads

$$\begin{aligned} & C(p^0, u^b) \\ = \;& C(p^b, u^b) + \sum_{n=1}^{N} \frac{\partial C(p^b, u^b)}{\partial p_n}(p_n^0 - p_n^b) \\ & +(1/2)\sum_{n=1}^{N}\sum_{n'=1}^{N} \frac{\partial^2 C(p^{**}, u^b)}{\partial p_n \partial p_{n'}}(p_n^0 - p_n^b)(p_{n'}^0 - p_{n'}^b) \\ = \;& \sum_{n=1}^{N} p_n^0 x_n^b + (1/2)\sum_{n=1}^{N}\sum_{n'=1}^{N} \frac{\partial^2 C(p^{**}, u^b)}{\partial p_n \partial p_{n'}}(p_n^0 - p_n^b)(p_{n'}^0 - p_{n'}^b), \end{aligned} \tag{28}$$

where $p^{**} \in [p^b, p^0]$. Substituting (26) and (28) into $P^{Lo}(p^1, p^0; x^b)$ delivers the following expression for the relative substitution bias:

$$\begin{aligned} & \frac{P^{Lo}(p^1, p^0; x^b) - P^K(p^1, p^0; u^b)}{P^K(p^1, p^0; u^b)} = \\ & \frac{1}{C(p^1, u^b)}\Bigg(P^{Lo}(p^1, p^0; x^b)(1/2)\sum_{n=1}^{N}\sum_{n'=1}^{N} \frac{\partial^2 C(p^{**}, u^b)}{\partial p_n \partial p_{n'}}(p_n^0 - p_n^b)(p_{n'}^0 - p_{n'}^b) \\ & -(1/2)\sum_{n=1}^{N}\sum_{n'=1}^{N} \frac{\partial^2 C(p^{*}, u^b)}{\partial p_n \partial p_{n'}}(p_n^1 - p_n^b)(p_{n'}^1 - p_{n'}^b)\Bigg). \end{aligned} \tag{29}$$

Thus,

$P^{Lo}(p^1, p^0; x^b) \geq P^K(p^1, p^0; u^b)$ if and only if

[8]When $C(p, u)$ is twice continuously differentiable in p Shephard's Lemma says that $\partial C(p, u)/\partial p_n = x_n(p, u)$ $(n = 1, ..., N)$. Hence, $\sum_n p_n \partial C(p, u)/\partial p_n = \sum_n p_n x_n(p, u) = C(p, u)$ and $\sum_n p_n \partial^2 C(p, u)/\partial p_n \partial p_{n'} = 0$ $(n' = 1, ..., N)$. Since $C(p, u)$ is concave in p, the square matrix of second-order partial derivatives $\partial^2 C(p, u)/\partial p_n \partial p_{n'}$ is negative semidefinite.

$$
\begin{aligned}
& -\sum_{n=1}^{N}\sum_{n'=1}^{N}\frac{\partial^2 C(p^*,u^b)}{\partial p_n \partial p_{n'}}(p_n^1 - p_n^b)(p_{n'}^1 - p_{n'}^b) \\
& \geq -P^{Lo}(p^1,p^0;x^b)\sum_{n=1}^{N}\sum_{n'=1}^{N}\frac{\partial^2 C(p^{**},u^b)}{\partial p_n \partial p_{n'}}(p_n^0 - p_n^b)(p_{n'}^0 - p_{n'}^b). \qquad (30)
\end{aligned}
$$

Notice that the concavity of $C(p,u)$ implies that $-\sum_{n=1}^{N}\sum_{n'=1}^{N}\frac{\partial^2 C(p^*,u^b)}{\partial p_n \partial p_{n'}}(p_n^1 - p_n^b)(p_{n'}^1 - p_{n'}^b)$ is positive. This expression can be interpreted as a squared distance between the price vectors p^1 and p^b; likewise, $-\sum_{n=1}^{N}\sum_{n'=1}^{N}\frac{\partial^2 C(p^{**},u^b)}{\partial p_n \partial p_{n'}}(p_n^0 - p_n^b)(p_{n'}^0 - p_{n'}^b)$ is positive and can be interpreted as a squared distance between the price vectors p^0 and p^b, using a (slightly) different metric.[9] Thus, the substitution bias of the Lowe price index is positive if and only if the distance between p^1 and p^b is greater than the distance between p^0 and p^b multiplied by the square root of the aggregate price change between the periods 0 and 1.

5 Substitution bias of the Cobb-Douglas price index

The relative substitution bias of the Cobb-Douglas price index is defined as the difference

$$
\ln P^{CD}(p^1,p^0;s^b) - \ln P^K(p^1,p^0;u^b) \qquad (31)
$$

where $u^b \equiv U(x^b)$. For the logarithm of the numerator of the cost of living index we obtain the following second-order Taylor series expression:

$$
\begin{aligned}
& \ln C(p^1,u^b) \\
& = \ln C(p^b,u^b) + \sum_{n=1}^{N}\frac{\partial \ln C(p^b,u^b)}{\partial \ln p_n}(\ln p_n^1 - \ln p_n^b) \\
& \quad +(1/2)\sum_{n=1}^{N}\sum_{n'=1}^{N}\frac{\partial^2 \ln C(p^*,u^b)}{\partial \ln p_n \partial \ln p_{n'}}(\ln p_n^1 - \ln p_n^b)(\ln p_{n'}^1 - \ln p_{n'}^b)
\end{aligned}
$$

[9]For two random column vectors a and b from the same distribution with covariance matrix S the Mahalanobis distance is defined as $[(a-b)^T S^{-1}(a-b)]^{1/2}$.

$$
\begin{aligned}
&= \ln C(p^b, u^b) + \sum_{n=1}^{N} s_n^b \ln(p_n^1/p_n^b) \\
&\quad +(1/2) \sum_{n=1}^{N} \sum_{n'=1}^{N} \frac{\partial^2 \ln C(p^*, u^b)}{\partial \ln p_n \partial \ln p_{n'}} \ln(p_n^1/p_n^b) \ln(p_{n'}^1/p_{n'}^b),
\end{aligned} \tag{32}
$$

where $p^* \in [p^b, p^1]$. The expression after the second equality sign was obtained by using Shephard's Lemma and the assumption that the consumer acts cost minimizing in period b. Similarly, for the denominator we obtain

$$
\begin{aligned}
&\ln C(p^0, u^b) \\
&= \ln C(p^b, u^b) + \sum_{n=1}^{N} \frac{\partial \ln C(p^b, u^b)}{\partial \ln p_n} (\ln p_n^0 - \ln p_n^b) \\
&\quad +(1/2) \sum_{n=1}^{N} \sum_{n'=1}^{N} \frac{\partial^2 \ln C(p^{**}, u^b)}{\partial \ln p_n \partial \ln p_{n'}} (\ln p_n^0 - \ln p_n^b)(\ln p_{n'}^0 - \ln p_{n'}^b) \\
&= \ln C(p^b, u^b) + \sum_{n=1}^{N} s_n^b \ln(p_n^0/p_n^b) \\
&\quad +(1/2) \sum_{n=1}^{N} \sum_{n'=1}^{N} \frac{\partial^2 \ln C(p^{**}, u^b)}{\partial \ln p_n \partial \ln p_{n'}} \ln(p_n^0/p_n^b) \ln(p_{n'}^0/p_{n'}^b),
\end{aligned} \tag{33}
$$

where $p^{**} \in [p^b, p^0]$. Subtracting (33) from (32) and applying the definitions of the Konüs and Cobb-Douglas price indices we obtain

$$
\begin{aligned}
&\ln P^{CD}(p^1, p^0; s^b) - \ln P^K(p^1, p^0; u^b) = \\
&\quad (1/2) \sum_{n=1}^{N} \sum_{n'=1}^{N} \frac{\partial^2 \ln C(p^{**}, u^b)}{\partial \ln p_n \partial \ln p_{n'}} \ln(p_n^0/p_n^b) \ln(p_{n'}^0/p_{n'}^b) \\
&\quad -(1/2) \sum_{n=1}^{N} \sum_{n'=1}^{N} \frac{\partial^2 \ln C(p^*, u^b)}{\partial \ln p_n \partial \ln p_{n'}} \ln(p_n^1/p_n^b) \ln(p_{n'}^1/p_{n'}^b).
\end{aligned} \tag{34}
$$

Hence,

$$
\begin{aligned}
&P^{CD}(p^1, p^0; s^b) \geq P^K(p^1, p^0; u^b) \text{ if and only if} \\
&\quad - \sum_{n=1}^{N} \sum_{n'=1}^{N} \frac{\partial^2 \ln C(p^*, u^b)}{\partial \ln p_n \partial \ln p_{n'}} \ln(p_n^1/p_n^b) \ln(p_{n'}^1/p_{n'}^b)
\end{aligned}
$$

$$\geq -\sum_{n=1}^{N}\sum_{n'=1}^{N} \frac{\partial^2 \ln C(p^{**}, u^b)}{\partial \ln p_n \partial \ln p_{n'}} \ln(p_n^0/p_n^b) \ln(p_{n'}^0/p_{n'}^b). \tag{35}$$

Expression (35) looks starkly like (30). The two factors can also be interpreted as distance measures, albeit that the matrices of second-order partial derivatives are not necessarily negative semidefinite.

6 Comparison

The relative substitution bias of the Lowe price index is given by expression (29) and that of the Cobb-Douglas price index by (34). It is not possible to compare these expressions without making some assumptions. To start with, consider the second-order partial derivatives $\partial^2 \ln C(p, u)/\partial \ln p_n \partial \ln p_{n'}$ $(n, n' = 1, ..., N)$. Straightforward computation delivers

$$\frac{\partial^2 \ln C(p, u)}{\partial \ln p_n \partial \ln p_{n'}} = \tag{36}$$
$$\frac{p_n p_{n'}}{C(p, u)} \frac{\partial^2 C(p, u)}{\partial p_n \partial p_{n'}} - s_n(p, u) s_{n'}(p, u) + \delta_{nn'} s_n(p, u) \ (n, n' = 1, ..., N),$$

where $s_n(p, u) \equiv p_n x_n(p, u)/C(p, u)$ $(n = 1, ..., N)$ and $\delta_{nn'} = 1$ if $n = n'$ and $\delta_{nn'} = 0$ otherwise. Applying this to the first factor at the right-hand side of the equality sign in (34) delivers

$$\sum_{n=1}^{N}\sum_{n'=1}^{N} \frac{\partial^2 \ln C(p^{**}, u^b)}{\partial \ln p_n \partial \ln p_{n'}} \ln(p_n^0/p_n^b) \ln(p_{n'}^0/p_{n'}^b) =$$
$$\sum_{n=1}^{N}\sum_{n'=1}^{N} \frac{p_n^{**} p_{n'}^{**}}{C(p^{**}, u^b)} \frac{\partial^2 C(p^{**}, u^b)}{\partial p_n \partial p_{n'}} \ln(p_n^0/p_n^b) \ln(p_{n'}^0/p_{n'}^b)$$
$$- \left(\sum_{n=1}^{N} s_n(p^{**}, u^b) \ln(p_n^0/p_n^b) \right)^2 + \sum_{n=1}^{N} s_n(p^{**}, u^b) \left(\ln(p_n^0/p_n^b) \right)^2$$
$$\approx \frac{1}{C(p^{**}, u^b)} \sum_{n=1}^{N}\sum_{n'=1}^{N} \frac{\partial^2 C(p^{**}, u^b)}{\partial p_n \partial p_{n'}} (p_n^0 - p_n^b)(p_{n'}^0 - p_{n'}^b)$$
$$+ \operatorname{var}(\ln(p_n^0/p_n^b))$$
$$= \frac{1}{C(p^{**}, u^b)} A + \operatorname{var}(\ln(p_n^0/p_n^b)), \tag{37}$$

where

$$
\begin{aligned}
&\operatorname{var}(\ln(p_n^0/p_n^b)) \equiv \\
&\quad \sum_{n=1}^{N} s_n(p^{**}, u^b) \left(\ln(p_n^0/p_n^b)\right)^2 - \left(\sum_{n=1}^{N} s_n(p^{**}, u^b) \ln(p_n^0/p_n^b)\right)^2
\end{aligned} \tag{38}
$$

is the variance of the relative price changes between periods b and 0, and

$$
A \equiv \sum_{n=1}^{N} \sum_{n'=1}^{N} \frac{\partial^2 C(p^{**}, u^b)}{\partial p_n \partial p_{n'}} (p_n^0 - p_n^b)(p_{n'}^0 - p_{n'}^b) \tag{39}
$$

is a temporary shorthand notation. Similarly, for the second factor at the right-hand side of the equality sign in (34) we obtain

$$
\begin{aligned}
&\sum_{n=1}^{N} \sum_{n'=1}^{N} \frac{\partial^2 \ln C(p^*, u^b)}{\partial \ln p_n \partial \ln p_{n'}} \ln(p_n^1/p_n^b) \ln(p_{n'}^1/p_{n'}^b) = \\
&\quad \approx \frac{1}{C(p^*, u^b)} \sum_{n=1}^{N} \sum_{n'=1}^{N} \frac{\partial^2 C(p^*, u^b)}{\partial p_n \partial p_{n'}} (p_n^1 - p_n^b)(p_{n'}^1 - p_{n'}^b) \\
&\quad + \operatorname{var}(\ln(p_n^1/p_n^b)) \\
&\quad = \frac{1}{C(p^*, u^b)} B + \operatorname{var}(\ln(p_n^1/p_n^b)),
\end{aligned} \tag{40}
$$

where

$$
\begin{aligned}
&\operatorname{var}(\ln(p_n^1/p_n^b)) \equiv \\
&\quad \sum_{n=1}^{N} s_n(p^*, u^b) \left(\ln(p_n^1/p_n^b)\right)^2 - \left(\sum_{n=1}^{N} s_n(p^*, u^b) \ln(p_n^1/p_n^b)\right)^2
\end{aligned} \tag{41}
$$

is the variance of the relative price changes between periods b and 1, and

$$
B \equiv \sum_{n=1}^{N} \sum_{n'=1}^{N} \frac{\partial^2 C(p^*, u^b)}{\partial p_n \partial p_{n'}} (p_n^1 - p_n^b)(p_{n'}^1 - p_{n'}^b) \tag{42}
$$

is also a temporary shorthand notation.

Substituting these two results into (34) we obtain for the relative substitution bias of the Cobb-Douglas price index the following expression

$$\ln P^{CD}(p^1,p^0;s^b) - \ln P^K(p^1,p^0;u^b) \approx \tag{43}$$
$$\frac{1/2}{C(p^*,u^b)}\left(\frac{C(p^*,u^b)}{C(p^{**},u^b)}A - B\right) - \frac{1}{2}\left(\mathrm{var}(\ln(p_n^1/p_n^b)) - \mathrm{var}(\ln(p_n^0/p_n^b))\right).$$

The relative substitution bias of the Lowe price index was given in expression (29). Substituting A and B yields

$$\frac{P^{Lo}(p^1,p^0;x^b) - P^K(p^1,p^0;u^b)}{P^K(p^1,p^0;u^b)} = \tag{44}$$
$$\frac{1/2}{C(p^1,u^b)}\left(P^{Lo}(p^1,p^0;x^b)A - B\right).$$

These two expressions make a comparison of the relative biases possible. Suppose that on average prices are monotonicly increasing between period b and 1. Then $-B$, measuring the price distance between periods b and 1, is positive and greater than $-A$ which is also positive, measuring the price distance between periods b and 0. Then

$$\frac{-A}{C(p^{**},u^b)} - \frac{-B}{C(p^*,u^b)} < \frac{-A}{p^b \cdot x^b} - \frac{-B}{C(p^1,u^b)} \tag{45}$$

because $C(p^{**},u^b) > C(p^b,u^b) = p^b \cdot x^b$ and $C(p^*,u^b) < C(p^1,u^b)$. Now

$$\frac{-A}{p^b \cdot x^b} = \frac{C(p^1,u^b)}{p^1 \cdot x^b}\frac{p^1 \cdot x^b}{p^b \cdot x^b}\frac{-A}{C(p^1,u^b)}, \tag{46}$$

of which the first right-hand side factor is less than or equal to 1 but the second is greater than $p^1 \cdot x^b/p^0 \cdot x^b$. Combining these two expressions we obtain as result

$$\frac{-A}{C(p^{**},u^b)} - \frac{-B}{C(p^*,u^b)} >< \frac{p^1 \cdot x^b}{p^0 \cdot x^b}\frac{-A}{C(p^1,u^b)} - \frac{-B}{C(p^1,u^b)}, \tag{47}$$

or

$$\frac{1/2}{C(p^*,u^b)}\left(\frac{C(p^*,u^b)}{C(p^{**},u^b)}A - B\right) >< \frac{1/2}{C(p^1,u^b)}\left(P^{Lo}(p^1,p^0;x^b)A - B\right). \tag{48}$$

Expression (43), however, contains an additional factor. Under increasing prices it is likely that $\mathrm{var}(\ln(p_n^1/p_n^b)) \geq \mathrm{var}(\ln(p_n^0/p_n^b))$, which makes this factor negative. On balance, it is likely that this factor has also more effect which implies that

$$\ln P^{CD}(p^1, p^0; s^b) - \ln P^K(p^1, p^0; u^b) \leq \frac{P^{Lo}(p^1, p^0; x^b) - P^K(p^1, p^0; u^b)}{P^K(p^1, p^0; u^b)}; \quad (49)$$

that is, the relative substitution bias of the Cobb-Douglas is less than the relative substitution bias of the Lowe price index.

7 Empirical evidence

Greenlees (1998) compared the Lowe price index[10] $P^{Lo}(p^t, p^{t-1}; x^b)$ with the Fisher price index

$$P^F(p^t, x^t, p^{t-1}, x^{t-1}) \equiv \left(\frac{p^t \cdot x^{t-1}}{p^{t-1} \cdot x^{t-1}} \frac{p^t \cdot x^t}{p^{t-1} \cdot x^t} \right)^{1/2}. \quad (50)$$

Assuming linear homogeneity of the utility function, this price index differentially approximates the cost of living index $P^K(p^t, p^{t-1}; \bar{u})$ to the second order, for any utility level $\bar{u}$; thus in particular for u^b.

Greenlees ran a simple regression

$$\ln P^{Lo}(p^t, p^{t-1}; x^b) - \ln P^F(p^t, x^t, p^{t-1}, x^{t-1}) = \alpha + \beta(t-b) + \epsilon(b \leq t-1) \quad (51)$$

on two slightly different sets of detailed annual price index and expenditure data for 1986-1995 and 1982-1995. The constant α appeared to be significantly positive, and of the order of 0.1 percent. The evidence on β, however, turned out to be inconclusive. Greenlees' conclusion was that "There is no definitive evidence that substitution bias increases over time with the age of the market basket."

My second piece of evidence uses a set of building blocks for the official Danish CPI.[11] The data set concerns 444 elementary aggregates. I used annual price

[10] Greenlees talks about Laspeyres index numbers, but he considers ratios of Laspeyres index numbers, which are Lowe index numbers.

[11] The data are by courtesy of Carsten Boldsen Hansen and the computations by courtesy of Jan de Haan.

index numbers (= arithmetic means of monthly index numbers) from 1996 to 2006, and expenditure shares for the years 1996, 1999, and 2003. Lowe and Cobb-Douglas price index numbers for year $t = 2000, ..., 2006$ relative to year $0 = 1999$ are calculated for three weight-reference years $b = 1996, 1999, 2003$. The results are contained in Table 1. Recall that the Lowe and Cobb-Douglas index with 1999 as weight-reference year are identical to the Laspeyres and the Geometric Laspeyres index respectively. That the entries in the column 'Lo(99)' are greater than those in the column 'CD(99)' should therefore come as no surprise.

Table 1: Lowe and Cobb-Douglas index numbers (1999 = 100)

Year	Lo(96)	CD(96)	Lo(99)	CD(99)	Lo(03)	CD(03)
1999	100.00	100.00	100.00	100.00	100.00	100.00
2000	102.99	102.74	103.00	102.85	102.54	102.81
2001	105.50	105.02	105.53	105.23	104.78	105.27
2002	108.16	107.43	108.37	107.83	107.30	107.97
2003	110.51	109.55	110.74	109.96	109.40	110.16
2004	112.15	110.68	112.23	111.05	110.62	111.25
2005	114.48	112.35	114.47	112.75	112.66	113.01
2006	116.83	114.18	116.79	114.58	114.79	114.85

It turns out that also for the weight-reference year 1996 the Lowe price index numbers are greater than their Cobb-Douglas counterparts. For the weight-reference year 2003, however, the order appears to be reversed. Further, a more recent weight-reference year leads to greater Cobb-Douglas index numbers. For the Lowe index numbers the evidence is mixed.

References

[1] Balk, B. M., 1989, "Changing Consumer Preferences and the Cost-of-Living Index: Theory and Nonparametric Expressions", Journal of Economics / Zeitschrift für Nationalökonomie 50, 157-169.

[2] Balk, B. M. and W. E. Diewert, 2003, The Lowe Consumer Price Index and its Substitution Bias, Discussion Paper No. 04-07 (Department of Economics, University of British Columbia, Vancouver).

[3] CPI Manual, 2004, Consumer Price Index Manual: Theory and Practice (Published for ILO, IMF, OECD, UN, Eurostat, The World Bank by ILO, Geneva).

[4] Diewert, W. E., 1993, "Duality Approaches to Microeconomic Theory", in Essays in Index Number Theory, Volume 1, edited by W. E. Diewert and A. O. Nakamura (North-Holland, Amsterdam).

[5] Greenlees, J. S., 1998, "Expenditure Weight Updates and Measured Inflation", in Proceedings of the Third Meeting of the International Working Group on Price Indices, edited by B. M. Balk (Statistics Netherlands, Voorburg). Downloadable from www.ottawagroup.org.

[6] Sen, A. K., 1977, "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory", Philosophy and Public Affairs 6, 317-344.