Laspeyres criticized this formula by showing that the index generally changed even if all prices remained constant (i.e. P_D does not satisfy an identity test to use modern terminology). An even more effective criticism of P_D is that it is not invariant to changes in the units of measurement (whereas P_L is invariant). Laspeyres did not write any further papers on index number theory. He wrote papers on economic history, the history of economic thought and on topical economic issues of his time; see Rinne [1981].

Selected Works

Laspeyres, E., 1863. Geschichte der Volkswirtschaftlichen Anschauungen der Niederländer und ihrer Literatur zur Zeit der Republik, Leipzig.

Laspeyres, E., 1871. "Die Berechnung einer mittleren Waarenpreissteigerung," *Jahrbücher für Nationalökonomie und Statistik* 16, 296-315.

References for Chapter 4

Drobisch, M.W., 1871. "Uber die Berechnung der Veranderungen der Waarenpreise und des Geldswerths," *Jahrbucher für Nationalökonomie und Statistik* 16, 143–156.

Rinne, H., 1981. "Ernst Louis Etienne Laspeyres 1834–1913," *Jahrbucher für Nationalökonomie und Statistik* 198, 194–216.

Chapter 5

INDEX NUMBERS*

W.E. Diewert

The index number problem may be phrased as follows. Suppose we have price data $p^i \equiv (p_1^i, \ldots, p_N^i)$ and quantity data $x^i \equiv (x_1^i, \ldots, x_N^i)$ on N commodities that pertain to economic unit i or that pertain to the same economic unit at time period i for $i = 1, 2, \ldots, I$. The *index number problem* is to find I numbers P^i and I numbers X^i such that

$$(1) \qquad P^i X^i = p^i \cdot x^i \equiv \sum_{n=1}^{N} p_n^i x_n^i \quad \text{for } i = 1, \ldots, I.$$

P^i is the *price index* for period i (or unit i) and X^i is the corresponding *quantity index*. P^i is supposed to be representative of all of the prices p_n^i, $n = 1, \ldots, N$ in some sense, while X^i is to be similarly representative of the quantities x_n^i, $n = 1, \ldots, N$. In what precise sense P^i and X^i represent the individual prices and quantities is not immediately evident and it is this ambiguity which leads to different approaches to index number theory. Note that we require that the product of the price and quantity indexes, $P^i X^i$, equals the actual period (or unit) i net expenditures on the N commodities, $p^i \cdot x^i$. Thus if the P^i are determined, then the X^i may be implicitly determined using equations (1), or vice versa.

Each individual consumes the services of thousands of commodities over a year and most producers utilize and/or produce thousands of individual products and services. Index numbers are used to reduce and summarize this overwhelming abundance of microeconomic information. Hence index numbers intrude themselves on virtually every empirical investigation in economics.

Index number theory splits naturally into two divisions, depending on the size of I. If $I = 2$, so that there are data for only two time periods or two economic units, then we are in the realm of *bilateral index number theory* while if $I > 2$, then we are in the realm of *multilateral indexes*. Bilateral approaches are considered in Sections 1–5 below and multilateral approaches are considered in Sections 6–10.

*First published in *The New Palgrave: A Dictionary of Economics*, Vol. 2, J. Eatwell, M. Milgate and P. Newman (eds.), The Macmillan Press, 1987, pp. 767–780.

The four main approaches to index number theory are: (i) *statistical* (Section 1), (ii) *test or axiomatic* (Sections 2 and 9), (iii) *microeconomic* which relies on the assumption of maximizing or minimizing behavior (Sections 3, 4 and 5), and (iv) *neostatistical* (Section 10).

1. Statistical Approaches

Let $I = 2$ and consider the following formula for P^2/P^1 due originally to Dutot [1738]:

$$P^2/P^1 = \left(\sum_{n=1}^{N} p_n^2/N\right) \Big/ \left(\sum_{n=1}^{N} p_n^1/N\right). \tag{2}$$

Thus the average level of prices in say period 2 relative to period 1 is set equal to the arithmetic average of the period 2 prices divided by the arithmetic average of the period 1 prices. The right hand side of (2) is called an *index number formula*.

Given an index number formula, we may solve the aggregation problem (1) as follows: set P^2 equal to the index number formula and determine P^1, X^1 and X^2 by:

$$P^1 = 1, \quad X^1 = p^1 \cdot x^1 \quad \text{and} \quad X^2 = p^2 \cdot x^2/P^2. \tag{3}$$

Setting $P^1 = 1$ is regarded as an arbitrary normalization; any other convenient normalization such as $P^1 = 100$ could be chosen, in which case $X^1 \equiv p^1 \cdot x^1/P^1$, $P^2 \equiv (P^2/P^1)P^1$ and $X^2 \equiv p^2 \cdot x^2/P^2$, where P^2/P^1 is the index number formula.

Rather than taking P^2/P^1 to be a ratio of average prices, Carli [1764] suggested taking an average of the price ratios as follows:

$$P^2/P^1 = \sum_{n=1}^{N} (p_n^2/p_n^1)/N. \tag{4}$$

The average of the price ratios in (4) is an arithmetic average. Jevons [1865] suggested using a geometric average:

$$P^2/P^1 = \prod_{n=1}^{N} (p_n^2/p_n^1)^{1/N}. \tag{5}$$

Once the ratio P^2/P^1 has been determined by (4) or (5), P^1, X^1 and X^2 may be determined using (3). But which of the three alternative formulae for P^2/P^1 should we use?

Walsh [1901] criticized the use of the Dutot formula (2) on the grounds that the index was not invariant to changes in the units of measurement. This criticism was a telling one, and virtually nobody uses formula (2) at present. However, formulae (4) and (5) are invariant to changes in the units of measurement, so we must still discriminate between them.

Jevons argued that changes in the quantity of money between the two periods would lead to proportional changes in all prices except for random errors. In particular, Jevons argued that the price ratios, p_n^2/p_n^1, would be independently and symmetrically distributed around a common mean. If this distribution happened to be the normal distribution, then the maximum likelihood estimator for the common mean leads to the index number formula (4). If the ratios p_n^2/p_n^1 happened to be log normally distributed, then statistical considerations would lead us to the index number formula (5).

Bowley [1928] attacked the use of both (4) and (5) on two grounds. First, from an empirical point of view, he showed that price ratios were not symmetrically distributed about a common mean and their logarithms also failed to be symmetrically distributed. Secondly, from a theoretical point of view, he argued that it was unlikely that prices or price ratios were independently distributed. Keynes [1930] developed Bowley's second objection in more detail; he argued that changes in the money supply would not affect all prices at the same time. Moreover, real disturbances in the economy could cause one set of prices to differ in a systematic way from other prices, depending on various elasticities of substitution and complementarity. In other words, prices are not randomly distributed, but are systematically related to each other through the general equilibrium of the economy.

The above criticisms led to a movement away from the use of unweighted averages of price ratios to represent price movements independently of quantity movements. Walsh [1901] and others suggested that the quantity observations x_n^i that were associated with the individual price observations p_n^i should be used as weights in the price index formula.

Scrope [1833] suggested the following formula:

$$P^2/P^1 = p^2 \cdot x/p^1 \cdot x \tag{6}$$

where $x \equiv (x_1, \ldots, x_N)$ was a somewhat vaguely specified quantity vector which was used to weight the price vectors p^1 and p^2 as in (6).

Laspeyres [1871] recommended that x be set equal to x^1, the period 1 quantity vector, while Paasche [1874] suggested that x be set equal to x^2, the period 2 quantity vector. This led to the following formulae:

$$P^2/P^1 = p^2 \cdot x^1/p^1 \cdot x^1 \equiv P_L(p^1, p^2, x^1, x^2); \tag{7}$$

$$P^2/P^1 = p^2 \cdot x^2/p^1 \cdot x^2 \equiv P_P(p^1, p^2, x^1, x^2). \tag{8}$$

Note that we are now following custom in the index number literature by defining an index number formula to be a function $P(p^1, p^2, x^1, x^2)$ of the price and quantity vectors that pertain to the two observations or time periods under consideration: P_L defines the Laspeyres price index while P_P defines the Paasche price index.

Pigou [1912] and Irving Fisher [1922] advocated taking a geometric mean of the Paasche and Laspeyres indexes and the resulting formula has come to be known as the Fisher ideal price index P_F:

$$(9) \qquad P^2/P^1 = [P_L P_P]^{1/2} \equiv P_F(p^1, p^2, x^1, x^2).$$

Rather than taking a geometric average of (7) and (8), Walsh [1901] [1921] advocated using formula (6) where the weight vector x was chosen to be the vector of geometric means of the two quantity vectors:

$$(10) \qquad \begin{aligned} P^2/P^1 &= \sum_{n=1}^{N} p_n^2 (x_n^1 x_n^2)^{1/2} \Big/ \sum_{n=1}^{N} p_n^1 (x_n^1 x_n^2)^{1/2} \\ &\equiv P_W(p^1, p^2, x^1, x^2). \end{aligned}$$

Törnqvist [1936] advocated a weighted geometric mean of the price ratios of the following form:

$$(11) \qquad P^2/P^1 = \prod_{n=1}^{N} (p_n^2/p_n^1)^{s_n} \equiv P_T(p^1, p^2, x^1, x^2)$$

where $s_n \equiv (1/2)(p_n^1 x_n^1/p^1 \cdot x^1) + (1/2)(p_n^2 x_n^2/p^2 \cdot x^2)$ is the average expenditure share on good n for $n = 1, \ldots, N$.

It turns out that formulae (7), (8), (9) and (11) are the most widely used formulae for a price index. However, at this point, we have no way of justifying their popularity. Walsh [1901] and Fisher [1922] present scores of functional forms for price indexes — on what basis are we to choose one as being better than the other?

This question leads us to discuss the test or axiomatic approach to index number theory.

2. The Test Approach to Bilateral Indexes

Consider $P(p^1, p^2, x^1, x^2)$, a function of the N period i prices, $p^i \equiv (p_1^i, \ldots, p_N^i)$, and the N period i quantities, $x^i \equiv (x_1^i, \ldots, x_N^i)$ for $i = 1, 2$. The price index P is supposed to represent the level of prices in period 2 relative to period 1. What

properties or *tests* should such an index number formula satisfy? The following nine tests (or closely related variants) have been considered in the literature.

BT1: *Identity Test:* $P(p^1, p^2, \alpha x^1, \beta x^2) = 1$ for all numbers $\alpha > 0, \beta > 0$ if $p^1 = p^2$ and $x^1 = x^2$.

BT2: *Proportionality Test:* $P(p^1, \alpha p^2, x^1, x^2) = \alpha P(p^1, p^2, x^1, x^2)$ for $\alpha > 0$.

BT3: *Invariance to Changes in Scale Test:* $P(\alpha p^1, \alpha p^2, \beta x^1, \gamma x^2) = P(p^1, p^2, x^1, x^2)$ for all $\alpha > 0$, $\beta > 0$ and $\gamma > 0$.

BT4: *Invariance to Changes in Units (Commensurability) Test:*

$$\begin{aligned} &P(\alpha_1 p_1^1, \ldots, \alpha_N p_N^1;\ \alpha_1 p_1^2, \ldots, \alpha_N p_N^2; \\ &\qquad \alpha_1^{-1} x_1^1, \cdots, \alpha_N^{-1} x_N^1;\ \alpha_1^{-1} x_1^2, \ldots, \alpha_N^{-1} x_N^2) \\ &\qquad\qquad = P(p^1, p^2, x^1, x^2) \quad \text{for } \alpha_1 > 0, \ldots, \alpha_N > 0. \end{aligned}$$

BT5: *Symmetric Treatment of Countries or Time (Country or Time Reversal) Test:* $P(p^2, p^1, x^2, x^1) = 1/P(p^1, p^2, x^1, x^2)$.

BT6: *Symmetric Treatment of Commodities (Commodity Reversal) Test:* $P(\tilde{p}^1, \tilde{p}^2, \tilde{x}^1, \tilde{x}^2) = P(p^1, p^2, x^1, x^2)$ where $\tilde{p}^i$ denotes a permutation of the elements of the vector p^i and $\tilde{x}^i$ denotes the same permutation of the elements of x^i, $i = 1, 2$.

BT7: *Monotonicity Test:* $P(p^1, p^2, x^1, x^2) \le P(p^1, p^3, x^1, x^2)$ if $p^2 \le p^3$; i.e., if $p_n^2 \le p_n^3$ for $n = 1, \ldots, N$.

BT8: *Mean Value Test:* $\min_n\{p_n^2/p_n^1\} \le P(p^1, p^2, x^1, x^2) \le \max_n\{p_n^2/p_n^1\}$.

BT9: *Circularity Test:* $P(p^1, p^2, x^1, x^2) P(p^2, p^3, x^2, x^3) = P(p^1, p^3, x^1, x^3)$.

Tests BT1 and BT3 may be found in Vartia [1985] who calls BT3 the *strong monetary unit test*. Test BT2 may be found in Walsh [1901], tests BT2 and BT4 are in Fisher [1911], tests BT5 and BT6 are in Fisher [1922], tests BT8 and a stronger version of BT7 are in Eichhorn and Voeller [1976] and test BT9 may be traced back to Westergaard [1890].

BT1 may be interpreted as follows: if prices and quantities are all equal in the two periods (or for the two regions under consideration), then the price index should be unity. This equality should still hold even if all quantities in period 1 are multiplied by the same number α and all quantities in period 2 are multiplied by the same β.

BT2 means if all period 2 prices are multiplied by α, then the new price index should equal α times the old price index.

Tests BT3–BT6 are invariance or symmetry tests. BT3 says that the price index should remain unchanged when each price in both periods is multiplied by the same number α and when quantities in period 1(2) are all multiplied by $\alpha(\beta)$. BT4 says that the index should remain unchanged if each good is measured in different units. BT5 says that if we interchange the role of periods 1 and 2 in our price index, then the new price index should equal the reciprocal of the original index. BT6 says that the index number formula should treat all commodities in an evenhanded way: no commodity can be singled out to play an asymmetric role. For example, suppose $P(p^1, p^2, x^1, x^2) \equiv p_1^2/p_1^1$. Then for $N \geq 2$, this formula, which equals the price ratio for commodity 1 only, fails BT6.

BT7 says that if period 2 prices increase in any manner, then the price index cannot decrease.

BT8 says that the price index should lie between the smallest and largest price ratios over all commodities.

BT9 is a transitivity test which looks beyond the case of only two periods or countries. BT9 says that if we have price and quantity data for three time periods, then the product of the price index going from period 1 to period 2 times the price index going from period 2 to 3 should equal the price index going from period 1 to 3 directly.

All of the above tests seem to be reasonable and desirable.

If $N = 1$ so that there is only one commodity, then BT1 and BT2 imply that $P(p^1, p^2, x^1, x^2)$ must equal p_1^2/p_1^1 and, of course, this index formula will satisfy all of the remaining tests.

In the general N commodity case (assuming that all prices and quantities are positive), which tests are satisfied by the index number formulae defined in the previous section?

It can be shown that the Dutot index (2) satisfies all tests except BT4 (but this is a fatal flaw), the Carli index (4) fails only BT5 and BT9, the Jevons geometric index (5) satisfies all tests, the Laspeyres and Paasche indexes defined by (7) and (8) fail BT5 and BT9, the Fisher and Walsh indexes (9) and (10) fail only BT9, and the Törnqvist index fails BT7 and BT9. Thus from the viewpoint of the test approach, it would appear that the geometric index is best.

The above conclusion is not warranted since our list of desirable tests is incomplete. We shall consider an additional two tests where the geometric index receives a failing grade.

Consider an index number formula that utilizes positive price and quantity information for N commodities, $P^N(p^1, p^2, x^1, x^2)$ say. Now consider the same functional form that uses information on only the first $N-1$ commodities, P^{N-1} say. Then we may want the index number formula to satisfy the following property:

$$(12) \quad \lim_{x_N^1 \to 0, x_N^2 \to 0} P^N(p^1, p^2, x^1, x^2) = P^{N-1}(p_1^1, \ldots, p_{N-1}^1, p_1^2, \ldots, p_{N-1}^2, x_1^1, \ldots, x_{N-1}^1, x_1^2, \ldots, x_{N-1}^2).$$

Thus as the quantity of commodity N tends to 0 in both periods (and thus commodity N becomes irrelevant), the N commodity price index tends to the $N-1$ commodity price index, where the prices and quantities of good N have been deleted from the formula. This might be called the *irrelevance of tiny commodities test*, test BT10. Obviously, this test only makes sense if the basic index number formula P satisfies BT6, so that all commodities are treated symmetrically. The geometric price index (5) fails test BT10 as do the other unweighted formulae (2) and (4). However, the quantity weighted price indexes, (7)–(11), all pass this test. Thus from the viewpoint of passing tests, the Fisher and Walsh indexes, (9) and (10), now look just as good as the geometric index (5).

There is another reason for not preferring the geometric price index. Recall our basic aggregation problem (1). It is clear that we can interchange the role of prices and quantities in the two periods and define a *quantity index* $Q(p^1, p^2, x^1, x^2)$ in much the same way that we defined the price index $P(p^1, p^2, x^1, x^2)$. We set $P^2/P^1 = P(p^1, p^2, x^1, x^2)$ and we may set $X^2/X^1 = Q(p^1, p^2, x^1, x^2)$. From (1), we deduce that the product of the price and quantity indexes should equal the value ratio for the two periods; i.e. P and Q should satisfy the following *product test* due to Fisher [1911] but named by Frisch [1930]:

$$(13) \quad P(p^1, p^2, x^1, x^2)Q(p^1, p^2, x^1, x^2) = p^2 \cdot x^2/p^1 \cdot x^1.$$

However, rather than defining Q independently of P, (13) may be rearranged to yield a *definition* of Q in terms of P. The resulting quantity index is called the *implicit quantity index* that corresponds to P. If we define the implicit quantity index, Q_G say, which corresponds to the geometric price index (5) and consider the quantity counterparts to tests BT1–BT9 above (BT10 does not have a sensible quantity counterpart), we find that Q_G fails the quantity counterpart to BT8, the mean value test. Hence we have another reason for failing the geometric price index and its corresponding implicit quantity index. On the other hand, the implicit Fisher quantity index satisfies tests BT1–BT8 adapted for the quantity context.

The implicit Fisher quantity index Q_F is:

$$\begin{aligned} & Q_F(p^1, p^2, x^1, x^2) \\ & \quad \equiv p^2 \cdot x^2/p^1 \cdot x^1 P_F(p^1, p^2, x^1, x^2) \\ & \quad = (p^1 \cdot x^2 p^2 \cdot x^2/p^1 \cdot x^1 p^2 \cdot x^1)^{1/2} \quad \text{using (9)} \\ (14) \quad & \quad = P_F(x^1, x^2, p^1, p^2). \end{aligned}$$

Thus Q_F has the same functional form as P_F except that the role of prices and quantities has been interchanged. We have shown that P_F and Q_F satisfy Fisher's [1922] *factor reversal test* which may be stated as follows:

$$(15) \quad P(p^1, p^2, x^1, x^2)P(x^1, x^2, p^1, p^2) = p^2 \cdot x^2/p^1 \cdot x^1.$$

Of the price indexes defined in the previous section, only the Fisher price index satisfies (15). However, the Fisher price index is by no means the only index number formula that satisfies the factor reversal test. Walsh [1921] showed how to generate hundreds of formulae that would satisfy the test: take a price index $P(p^1, p^2, x^1, x^2)$, and define its factor antithesis by $p^2 \cdot x^2 / \; p^1 \cdot x^1 P(x^1, x^2, p^1, p^2)$. Define a new price index by taking the square root of the product of the original index and its factor antithesis. This new index will automatically satisfy the factor reversal test.

The consistency and independence of various bilateral index number tests was studied in some detail by Eichhorn and Voeller [1976]. Our conclusion at this point echoes that of Frisch [1936]: the test approach to index number theory, while extremely useful, does not lead to a single unique index number formula. Thus we turn to economic approaches to index number theory to see if we are led to a more definite conclusion.

3. Microeconomic Approaches to Price Indexes

Before a definition of a microeconomic price index is presented, it is necessary to make a few preliminary definitions.

Let $F(x)$ be a function of N variables, $x \equiv (x_1, \ldots, x_N)$. In the consumer context, F represents a consumer's preferences; i.e. if $F(x^2) > F(x^1)$, then the consumer prefers the commodity vector x^2 over x^1. In this context, F is called a *utility function*. In the producer context, $F(x)$ might represent the output that could be produced using the input vector x. In this context, F is called a *production function*. In order to cover both contexts, we follow the example of Diewert [1976a] and call F an *aggregator function*.

Suppose the consumer or producer faces prices $p \equiv (p_1, \ldots, p_N)$ for the N commodities. Then the economic agent will generally find it is useful to minimize the cost of achieving at least a given utility or output level u; we define the *cost function* or *expenditure function* C as the solution to this minimization problem:

$$C(u, p) \equiv \underset{x}{\text{minimum}}\{p \cdot x : F(x) \geq u\} \tag{16}$$

where $p \cdot x \equiv \sum_{n=1}^{N} p_n x_n$ is the inner product of the price vector p and quantity vector x.

Note that the cost function depends on $1 + N$ variables; the utility or output level u and the N commodity prices in the vector p. Moreover, the functional form for the aggregator function F completely determines the functional form for C.

We say that an aggregator function is *neoclassical* if F is: (i) continuous, (ii) positive; i.e. $F(x) > 0$ if $x \gg 0_N$ (which means each component of x is positive), and (iii) linearly homogeneous; i.e. $F(\lambda x) = \lambda F(x)$ if $\lambda > 0$. If F is neoclassical, then the corresponding cost function $C(u, p)$ equals u times the unit cost function, $c(p) \equiv C(1, p)$, where $c(p)$ is the minimum cost of producing one unit of utility or output; i.e.,

$$C(u, p) = uC(1, p) \equiv uc(p). \tag{17}$$

Shephard [1953] formally defined an aggregator function F to be *homothetic* if there exists an increasing continuous function of one variable g such that $g[F(x)]$ is neoclassical. However, the concept of homotheticity was well known to Frisch [1936] who termed it expenditure proportionality. If F is homothetic, then its cost function C has the following decomposition:

$$\begin{aligned} C(u, p) &\equiv \min_x\{p \cdot x : F(x) \geq u\} \\ &= \min_x\{p \cdot x : g[F(x)] \geq g(u)\} \\ &= g(u)c(p) \end{aligned} \tag{18}$$

where $c(p)$ is the unit cost function that corresponds to $g[F(x)]$.

Let $p^1 \gg 0_N$ and $p^2 \gg 0_N$ be positive price vectors pertaining to periods or observations 1 and 2. Let $x > 0_N$ be a nonnegative, nonzero reference quantity vector. Then the Konüs [1924] *price index* or *cost of living index* is defined as:

$$P_K(p^1, p^2, x) \equiv C[F(x), p^2]/C[F(x), p^1]. \tag{19}$$

In the consumer (producer) context, P_K may be interpreted as follows. Pick a reference utility (output) level $u \equiv F(x)$. Then $P_K(p^1, p^2, x)$ is the minimum cost of achieving the utility (output) level u when the economic agent faces prices p^2 relative to the minimum cost of achieving the same u when the agent faces prices p^1. If $N = 1$ so that there is only one consumer good (or input), then it is easy to show that $P_K(p_1^1, p_1^2, x_1) = p_1^2 x_1 / p_1^1 x_1 = p_1^2 / p_1^1$.

Using the fact that a cost function is linearly homogeneous in its price arguments, it can be shown that P_K has the following homogeneity property: $P_K(p^1, \lambda p^2, x) = \lambda P_K(p^1, p^2, x)$ for $\lambda > 0$ which is analogous to the proportionality test BT2 in the previous section. P_K also satisfies $P_K(p^2, p^1, x) = 1/P_K(p^1, p^2, x)$ which is analogous to the time reversal test, BT5.

Note that the functional form for P_K is completely determined by the functional form for the aggregator function F which determines the functional form for the cost function C.

In general, P_K depends not only on the two price vectors p^1 and p^2, but also on the reference vector x. Malmquist [1953], Pollak [1971a] and Samuelson and Swamy [1974] have shown that P_K is independent of x and is equal to a ratio of unit cost functions, $c(p^2)/c(p^1)$, if and only if the aggregator function F is homothetic.

If we knew the consumer's preferences or the producer's technology, then we would know F and we could construct the cost function C and the Konüs price index P_K. However, we generally do not know F or C and thus it is useful to develop *bounds* that depend on observable price and quantity data but do not depend on the specific functional form for F or C.

Samuelson [1947] and Pollak [1971a] established the following bounds on P_K. Let $p^1 \gg 0_N$, and $p^2 \gg 0_N$. Then for every reference quantity vector $x > 0_N$, we have

$$\min_n\{p_n^2/p_n^1\} \le P_K(p^1, p^2, x) \le \max_n\{p_n^2/p_n^1\}; \tag{20}$$

i.e., P_K lies between the smallest and largest price ratios. Unfortunately, these bounds are usually too wide to be of much practical use.

To obtain closer bounds, we now assume that the observed quantity vectors for the two periods, $x^i \equiv (x_1^i, \ldots, x_N^i)$, $i = 1, 2$, are solutions to the producer's or consumer's cost minimization problems; i.e., we assume:

$$p^i \cdot x^i = C(F(x^i), p^i), \qquad p^i \gg 0_N, \quad x^i > 0_N, \qquad i = 1, 2. \tag{21}$$

Given the above assumptions, we now have two natural choices for the reference quantity vector x that occurs in the definition of $P_K(p^1, p^2, x) : x^1$ or x^2. The *Laspeyres–Konüs price index* is defined as $P_K(p^1, p^2, x^1)$ and the *Paasche–Konüs price index* is defined as $P_K(p^1, p^2, x^2)$.

Under the assumption of cost minimizing behavior (21), Konüs [1924] established the following bounds:

$$P_K(p^1, p^2, x^1) \le p^2 \cdot x^1/p^1 \cdot x^1 \equiv P_L(p^1, p^2, x^1, x^2); \tag{22}$$

$$P_K(p^1, p^2, x^2) \ge p^2 \cdot x^2/p^1 \cdot x^2 \equiv P_P(p^1, p^2, x^1, x^2), \tag{23}$$

where P_L and P_P are the Laspeyres and Paasche price indexes defined earlier by (7) and (8). If in addition, the aggregator function is homothetic, then Frisch [1936] showed that for any reference vector $x > 0_N$,

$$P_P \equiv p^2 \cdot x^2/p^1 \cdot x^2 \le P_K(p^1, p^2, x) \le p^2 \cdot x^1/p^1 \cdot x^1 \equiv P_L. \tag{24}$$

In the consumer context, it is unlikely that preferences will be homothetic; hence the bounds (24) cannot be justified in general. However, Konüs [1924] showed that bounds similar to (24) would hold even in the general nonhomothetic case, provided that we choose a reference vector $x \equiv \lambda x^1 + (1 - \lambda)x^2$ which is a $\lambda, (1 - \lambda)$ weighted average of the two observed quantity points.

Specifically, Konüs showed that there exists a λ between 0 and 1 such that if $P_P \le P_L$, then

$$P_P \le P_K[p^1, p^2, \lambda x^1 + (1 - \lambda)x^2] \le P_L \tag{25}$$

or if $P_P > P_L$, then

$$P_L \le P_K[p^1, p^2, \lambda x^1 + (1 - \lambda)x^2] \le P_P. \tag{26}$$

The bounds on the microeconomic price index P_K given by (20) and (22)–(26) are the best bounds that we can obtain without making further assumptions on F. In the time series context, the bounds given by (25) or (26) are quite satisfactory: the Paasche and Laspeyres price indexes for consecutive time periods will usually differ by less than 1 percent. However, in the cross section context where the observations represent, for example, production data for two producers in the same industry but in different regions, the bounds are often not very useful since P_L and P_P can differ by 50 percent or more in the cross sectional context; see Ruggles [1967].

In Section 5 below, we will make additional assumptions on the aggregator function F or its cost function dual C that will enable us to determine P_K exactly. Before we do this, in the next section, we will define various quantity indexes that have their origins in microeconomic theory.

4. Microeconomic Approaches to Quantity Indexes

In the one commodity case, a natural definition for a quantity index is x_1^2/x_1^1, the ratio of the single quantity in period 2 to the corresponding quantity in period 1. This ratio is also equal to the expenditure ratio, $p_1^2 x_1^2/p_1^1 x_1^1$, divided by the price ratio, p_1^2/p_1^1. This suggests that in the N commodity case, a reasonable definition for a quantity index would be the expenditure ratio divided by the Konüs price index, P_K. This course of action was suggested by Pollak [1971a]. Thus we define the *Konüs–Pollak quantity index*, Q_K, by:

$$\begin{aligned} Q_K(p^1, p^2, x^1, x^2, x) &\equiv p^2 \cdot x^2/p^1 \cdot x^1 P_K(p^1, p^2, x) \\ &= \Big[C[F(x^2), p^2]/C[F(x), p^2]\Big] \Big/ \Big[C[F(x^1), p^1]/C[F(x), p^1]\Big] \end{aligned} \tag{27}$$

where the second line follows from the definition of P_K, (19), and the assumption of cost minimizing behavior in the two periods, (21).

The definition of Q_K depends on the reference vector x which appears in the definition of P_K. The general definition of Q_K simplifies considerably if we choose x to be x^1 or x^2. Thus define the *Laspeyres–Konüs quantity index* as

$$Q_K(p^1, p^2, x^1, x^2, x^1) \equiv C[F(x^2), p^2]/C[F(x^1), p^2] \tag{28}$$

and the *Paasche–Konüs quantity index* as

$$Q_K(p^1, p^2, x^1, x^2, x^2) \equiv C[F(x^2), p^1]/C[F(x^1), p^1]. \tag{29}$$

It turns out that the indexes defined by (28) and (29) are special cases of another class of quantity indexes. For any reference price vector $p \gg 0_N$, define the *Allen* [1949] *quantity index* by

$$Q_A(x^1, x^2, p) \equiv C[F(x^2), p]/C[F(x^1), p]. \tag{30}$$

If p is chosen to be p^1, (30) becomes (29) and if $p = p^2$, then (30) becomes (28).

Using the properties of cost functions, it can be shown that if $F(x^2) \geq F(x^1)$, then $Q_A(x^1, x^2, p) \geq 1$ while if $F(x^2) \leq F(x^1)$, then $Q_A(x^1, x^2, p) \leq 1$. Thus the Allen quantity index correctly indicates whether the commodity vector x^2 is larger or smaller than x^1. It can also be seen that Q_A satisfies a counterpart to the time reversal test; i.e., $Q_A(x^2, x^1, p) = 1/Q_A(x^1, x^2, p)$.

Just as the price index P_K depended on the unobservable aggregator function, so also do the quantity indexes Q_K and Q_A. Thus it is useful to develop bounds for the quantity indexes that do not depend on the particular functional form for F.

Samuelson [1947] and Allen [1949] established the following bounds for (28) and (29):

$$Q_A(x^1, x^2, p^1) = Q_K(p^1, p^2, x^1, x^2, x^2) \leq p^1 \cdot x^2/p^1 \cdot x^1 \equiv Q_L; \tag{31}$$

$$Q_A(x^1, x^2, p^2) = Q_K(p^1, p^2, x^1, x^2, x^1) \geq p^2 \cdot x^2/p^2 \cdot x^1 \equiv Q_P. \tag{32}$$

Note that the observable *Laspeyres* and *Paasche quantity indexes*, Q_L and Q_P, appear on the right hand sides of (31) and (32).

Diewert [1981a], utilizing some results of Pollak [1971a] and Samuelson and Swamy [1974], established the following results: if the underlying aggregator function F is neoclassical and (21) holds, then for all $p \gg 0_N$ and $x \gg 0_N$,

$$Q_P \leq Q_A(x^1, x^2, p) = Q_K(p^1, p^2, x^1, x^2, x) = F(x^2)/F(x^1) \leq Q_L. \tag{33}$$

Thus if the aggregator function F is neoclassical, then the Allen quantity index for all reference vectors p equals the Konüs quantity index for all reference quantity vectors x which in turn equals the ratio of aggregates $(F(x^2)/F(x^1))$. Moreover, Q_A and Q_K are bounded from below by the Paasche quantity index Q_P, and bounded from above by the Laspeyres quantity index Q_L in the neoclassical case.

In the general nonhomothetic case, Diewert [1981a] showed that there exists a λ between 0 and 1 such that $Q_K[p^1, p^2, x^1, x^2, \lambda x^1 + (1-\lambda)x^2]$ lies between

Q_P and Q_L and there exists a λ^* between 0 and 1 such that $Q_A[x^1, x^2, \lambda^* p^1 + (1-\lambda^*)p^2]$ also lies between Q_P and Q_L. Thus the observable Paasche and Laspeyres quantity indexes bound both the Konüs quantity index and the Allen quantity index, provided that we choose appropriate reference vectors between x^1 and x^2 and p^1 and p^2 respectively.

Using the linear homogeneity property of the cost function in its price arguments, we can show that the Konüs price index has the desirable homogeneity property, $P_K(p^1, \lambda p^1, x) = \lambda$ for all $\lambda > 0$; i.e., if period 2 prices are proportional to period 1 prices, then P_K equals this common proportionality factor. It would be desirable for an analogous homogeneity property to hold for quantity indexes. Unfortunately, it is not in general true that $Q_K(x^1, \lambda x^1, p^1, p^2, x) = \lambda$ or that $Q_A(x^1, \lambda x^1, p) = \lambda$. Thus we turn to a third microeconomic approach to defining a quantity index which does have the desirable quantity proportionality property.

Let x^1 and x^2 be the observable quantity vectors in the two situations as usual, let $F(x)$ be an increasing, continuous aggregator function, and let $x \gg 0$ be a reference quantity vector. Then the *Malmquist* [1953] *quantity index* Q_M is defined as:

$$Q_M(x^1, x^2, x) \equiv D[F(x), x^2]/D[F(x), x^1] \tag{34}$$

where $D(u, x^i) \equiv \max_k\{k : F(x^i/k) \geq u,\ k > 0\}$ is the *deflation* or *distance function* which corresponds to F. Thus $D[F(x), x^2]$ is the biggest number which will just deflate the quantity vector x^2 onto the boundary of the utility (or production) possibilities set $\{z : F(z) \geq F(x)\}$ indexed by the reference quantity vector x while $D[F(x), x^1]$ is the biggest number which will just deflate the quantity vector x^1 onto the set $\{z : F(z) \geq F(x)\}$ and Q_M is the ratio of these two deflation factors.

Q_M depends on the unobservable aggregator function F and as usual, we are interested in bounds for Q_M.

Diewert [1981a] showed that Q_M satisfied bounds analogous to (20); i.e.,

$$\min_n\{x_n^2/x_n^1\} \leq Q_M(x^1, x^2, x) \leq \max_n\{x_n^2/x_n^1\}. \tag{35}$$

It should be noted that we do not require the assumption of cost minimizing behavior in order to define the Malmquist quantity index or to establish the bounds (35). However, in order to establish the following bounds due to Malmquist [1953] for Q_M, we do need the assumption of cost minimizing behavior (21) and we require the reference vector x to be x^1 or x^2:

$$Q_M(x^1, x^2, x^1) \leq p^1 \cdot x^2/p^1 \cdot x^1 \equiv Q_L; \tag{36}$$

$$Q_M(x^1, x^2, x^2) \geq p^2 \cdot x^2/p^2 \cdot x^1 \equiv Q_P. \tag{37}$$

Diewert [1981a] showed that under the hypothesis of cost minimizing behavior, there exists a λ between 0 and 1 such that $Q_M[x^1, x^2, \lambda x^1 + (1-\lambda)x^2]$ lies between Q_P and Q_L. Thus the Paasche and Laspeyres quantity indexes provide bounds for a Malmquist quantity index for some reference indifference or product surface indexed by a quantity vector which is a $\lambda, (1-\lambda)$ weighted average of the two observable quantity vectors, x^1 and x^2.

Pollak [1971a] showed that if F is neoclassical, then we can extend the string of equalities in (33) to include the Malmquist quantity index $Q_M(x^1, x^2, x)$, for any reference quantity vector x. Thus in the case of a linearly homogeneous aggregator function, all three theoretical quantity indexes coincide and this common theoretical index is bounded from below by the Paasche quantity index Q_P and bounded from above by the Laspeyres quantity index Q_L.

In the general case of a nonhomothetic aggregator function, our best theoretical quantity index, the Malmquist index, is also bounded by the Paasche and Laspeyres indexes, provided that we choose a suitable reference quantity vector.

We noted in the price index context that the Paasche and Laspeyres price indexes were usually quite close in the time series context. A similar remark also applies to the Paasche and Laspeyres quantity indexes. Thus taking an average of the Paasche and Laspeyres indexes, such as the Fisher price and quantity indexes, will generally approximate underlying microeconomic price and quantity indexes sufficiently accurately for most practical purposes. However, this observation does not apply to the cross sectional context, where the Paasche and Laspeyres indexes can differ widely. In the following section, we offer another microeconomic justification for using the Fisher indexes that also applies in the context of making interregional and cross country comparisons.

5. Exact and Superlative Indexes

Assume that the producer or consumer is maximizing a neoclassical aggregator function f subject to a budget constraint during the two periods. Under these conditions, it can be shown that the economic agent is also minimizing cost subject to a utility or output constraint. Moreover, the cost function C that corresponds to f can be written as $C[f(x), p] = f(x)c(p)$ where c is the unit cost function (recall (17) above).

Suppose a price index $P(p^1, p^2, x^1, x^2)$ and a quantity index $Q(p^1, p^2, x^1, x^2)$ of the type considered in Sections 1 and 2 are given. The quantity index Q is defined to be *exact* for a neoclassical aggregator function f with unit cost dual c if for every $p^1 \gg 0_N$, $p^2 \gg 0_N$ and $x^i \gg 0_N$ which is a solution to the aggregator maximization problem $\max_x\{f(x) : p^i \cdot x \leq p^i \cdot x^i\} = f(x^i) > 0$ for

$i = 1, 2$, we have

$$Q(p^1, p^2, x^1, x^2) = f(x^2)/f(x^1). \tag{38}$$

Under the same hypothesis, the price index P is *exact* for f and c if we have

$$P(p^1, p^2, x^1, x^2) = c(p^2)/c(p^1). \tag{39}$$

In (38) and (39), the price and quantity vectors are not regarded as being independent. The p^i can be independent, but the x^i are solutions to the corresponding aggregator maximization problem involving p^i, for $i = 1, 2$. Note that if Q is exact for a neoclassical f, then Q can be interpreted as a Konüs, Allen or Malmquist quantity index and the corresponding P defined implicitly by (13) can be interpreted as a Konüs price index.

The concept of exactness is due to Konüs and Byushgens [1926]. Below, we shall give some examples of exact index number formulae. Additional examples may be found in Afriat [1972b], Pollak [1971a] and Samuelson and Swamy [1974].

Konüs and Byushgens [1926] showed that Irving Fisher's ideal quantity index Q_F defined by (14) and the corresponding price index P_F defined by (9) are exact for the homogeneous quadratic aggregator function f defined by

$$f(x_1, \ldots, x_N) \equiv \left(\sum_{n=1}^N \sum_{m=1}^N a_{nm} x_n x_m\right)^{1/2} \equiv (x^T A x)^{1/2} \tag{40}$$

where $A \equiv [a_{nm}]$ is a symmetric $N \times N$ matrix of constants. Thus under the assumption of maximizing behavior, we can calculate $f(x^2)/f(x^1) = Q_F$ and $c(p^2)/c(p^1) = P_F$ where f is defined by (40) and c is the unit cost function that corresponds to f. The important thing to note is that f depends on $N(N+1)/2$ unknown a_{nm} parameters but we do not need to know these parameters in order to evaluate $f(x^2)/f(x^1)$ and $c(p^2)/c(p^1)$.

Diewert [1976a] showed that the Törnqvist price index P_T defined by (11) is exact for the unit cost function $c(p)$ defined by:

$$\ln c(p) \equiv \alpha_0 + \sum_{n=1}^N \alpha_n \ln p_n + (1/2) \sum_{m=1}^N \sum_{n=1}^N \alpha_{mn} \ln p_m \ln p_n \tag{41}$$

where the parameters α_n and α_{mn} satisfy the following restrictions:

$$\sum_{n=1}^N \alpha_n = 1, \quad \sum_{n=1}^N \alpha_{mn} = 0 \quad \text{for } m = 1, \ldots, N \quad \text{and} \\ \alpha_{mn} = \alpha_{nm} \quad \text{for all } m, n. \tag{42}$$

Thus we may calculate $c(p^2)/c(p^1) = P_T$ and $f(x^2)/f(x^1) = p^2 \cdot x^2 / p^1 \cdot x^1 P_T \equiv \widetilde{Q}_T$ where c is the unit cost function defined by (41), f is the aggregator function which corresponds to this c, and $\widetilde{Q}_T$ is the implicit Törnqvist quantity index. Note that we do not have to know the parameters α_n and α_{mn} in order to evaluate $c(p^2)/c(p^1)$ and $f(x^2)/f(x^1)$.

The unit cost function defined by (41) is the *translog* unit cost function defined by Christensen, Jorgenson and Lau [1971]. Since P_T is exact for this translog functional form, P_T is sometimes called the *translog price index*.

Before we present our final example of an exact index number formula, we need to define a family of quantity indexes Q_r that depend on a number, $r \neq 0$:

(43)
$$Q_r(p^1,p^2,x^1,x^2) \equiv \left[\sum_{n=1}^N s_n^1(x_n^2/x_n^1)^{r/2}\right]^{1/r} \times \left[\sum_{m=1}^N s_m^2(x_m^2/x_m^1)^{-r/2}\right]^{-1/r}$$

where $s_n^i \equiv p_n^i x_n^i/p^i \cdot x^i$ is the period i expenditure share for good n. For each $r \neq 0$, define the corresponding implicit price index by:

(44)
$$\widetilde{P}_r(p^1,p^2,x^1,x^2) \equiv p^2 \cdot x^2/p^1 \cdot x^1 Q_r(p^1,p^2,x^1,x^2).$$

A bit of algebra will show that when $r = 2$, $\widetilde{P}_2 = P_F$, the Fisher price index defined by (9) and when $r = 1$, $\widetilde{P}_1 = P_W$, the Walsh price index defined by (10).

Diewert [1976a] showed that Q_r and $\widetilde{P}_r$ are exact for the *quadratic mean of order r aggregator function* f_r defined by:

(45)
$$f_r(x_1,\ldots,x_N) \equiv \left(\sum_{m=1}^N \sum_{n=1}^N a_{mn} x_m^{r/2} x_n^{r/2}\right)^{1/r}$$

where $A \equiv [a_{mn}]$ is a symmetric matrix of constants. Thus the Walsh price index P_W is exact for $f_1(x)$ defined by (45) when $r = 1$.

Diewert [1974a] defined a linearly homogeneous function f of N variables to be *flexible* if it could provide a second order approximation to an arbitrary twice continuously differentiable linearly homogeneous function. It can be shown that f defined by (40), c defined by (41) and (42) and f_r defined by (45) for each $r \neq 0$ are all examples of flexible functional forms.

Let the price and quantity indexes P and Q satisfy the product test equality, (13). Then Diewert [1976a] defined P and Q to be *superlative indexes* if either P is exact for a flexible unit cost function c or Q is exact for a flexible aggregator function f. Thus P_F, P_W, P_T and $\widetilde{P}_r$ are all superlative price indexes.

At this point, it may seem that we are in the same position that we were at the end of Section 2 where we could not find an index number formula that satisfied all reasonable tests; hence we could not single out any formula as being *the* best. In the present context, we have a similar problem: how are we to discriminate between P_F, P_W and P_T? Fortunately, it does not matter very much which of these formulae we choose to use in applications; they will all give the same answer to a reasonably high degree of approximation. Diewert [1978b] showed that all known superlative index number formulae approximate each other to the second order when each index is evaluated at an equal price and quantity point. This means the P_F, P_W, P_T and each $\widetilde{P}_r$ have the same first and second order partial derivatives with respect to all $4N$ arguments when the derivatives are evaluated at a point where $p^1 = p^2$ and $x^1 = x^2$. A similar string of equalities also holds for the corresponding implicit quantity indexes defined using the product test (13). Empirically, it has been found that superlative indexes typically approximate each other to something less than 0.2 percent in the time series context and to about 2 percent in the cross section context; see Fisher [1922] and Ruggles [1967].

Diewert [1978b] also showed that the Paasche and Laspeyres indexes approximate the superlative indexes to the first order at an equal price and quantity point. In the time series context, for adjacent periods, the Paasche and Laspeyres price indexes typically differ by less than 0.5 percent; hence these indexes may also provide acceptable approximations to a superlative index.

Having considered the case of two observations at great length, we now turn our attention to the I observation case. In the next section, we consider the case of I consecutive time series observations of prices and quantities on the same producer or consumer. We shall consider the general case of I observations separated by space (and possibly by time) in Section 9 below.

6. The Fixed Base Versus the Chain Principle

Consider the case of I consecutive observations through time on the prices and quantities of N goods utilized by an economic unit, $p^i \equiv (p_1^i,\ldots,p_N^i)$ and $x^i \equiv (x_1^i,\ldots,x_N^i)$, $i = 1,\ldots,I$.

Suppose that we have decided on a price index P and a quantity index Q where P and Q satisfy (13). This allows us to compare prices and quantities for any two observations. Based on the test approach, we would probably choose P and Q to be the Fisher indexes, P_F and Q_F, since they satisfy the most tests. Since the Fisher indexes also lie between the Paasche and Laspeyres bounds derived in Sections 3 and 4 and moreover are superlative indexes, the choice of the Fisher indexes also fits in well with the microeconomic approach to index number theory.

However, if we choose any superlative price index P in order to make bilateral comparisons of prices, then we are faced with a problem when $I \geq 3$

because the circular test BT9 will not be satisfied. Thus choose observation 1 as the base and make all price comparisons relative to period 1, so that the relative price level in period i is $P(p^1, p^i, x^1, x^i)$. Now choose another observation, say I, as the base so that the price level in period i relative to I is $P(p^I, p^i, x^I, x^i)$. In order to make the price level in period 1 equal to unity, divide each $P(p^I, p^i, x^I, x^i)$ by $P(p^I, p^1, x^I, x^1)$. If the circular test held, then the two price series would coincide; i.e., we would have

$$(46) \qquad P(p^1, p^i, x^1, x^i) = P(p^I, p^i, x^I, x^i)/P(p^I, p^1, x^I, x^1), \quad i = 1, \ldots, I.$$

However, since the circular test does not hold in general, we are faced with a problem: which period should we choose as the base?

There are a number of alternative strategies that could be followed, including: (i) choose the first observation as the base, (ii) take an average over all possible choices of the base period, (iii) abandon the use of a bilateral formula and develop an entirely new multilateral approach, or (iv) use the *chain principle*, which will be explained below.

Alternative (i) seems rather arbitrary but it has simplicity to recommend it. Virtually all official time series indexes are constructed using fixed base Paasche or Laspeyres indexes with the base year being changed every 5 to 15 years.

Alternative (ii) seems attractive at first glance since it treats each period in a symmetric fashion. The problem with the method is that economic history has to be rewritten every time a new observation is added to the initial I observations.

Alternative (iii) may also seem attractive. For example, in making a price comparison between i and j, we may want to utilize the quantity information for all periods, so that the bilateral index number formula $P(p^i, p^j, x^i, x^j)$ could be replaced by $P^*(p^i, p^j, x^1, \ldots, x^I)$. For example, we could use the Scrope index (6) where x could be taken to be the average quantities over all periods, $\sum_{i=1}^I x^i/I$, or we could use the Törnqvist index (11) where s_n could be set equal to the average over all periods for the commodity n expenditure share, $\sum_{i=1}^I (1/I) p_n^i x_n^i / p^i \cdot x^i$. Both of these examples lead to indexes that satisfy the circularity property for the original I observations. However, as was the case with alternative (ii), these new multilateral indexes would have to be recomputed as new time series observations become available.

Alternative (iv) is to use the *chain principle*, originally suggested by Marshall [1887], although the term is due to Fisher [1911]. This principle makes use of the natural order provided by the march of time. One first chooses a bilateral index number formula P. The period 1 price level is set equal to unity and the period 2 price level is set equal to $P(p^1, p^2, x^1, x^2)$. The period 3 price level is set equal to $P(p^1, p^2, x^1, x^2)P(p^2, p^3, x^2, x^3)$. The period 4 price level is set equal to the period 3 price level times $P(p^3, p^4, x^3, x^4)$, and so on. Thus the

period i price level is not obtained by the direct comparison of period i prices with period 1 prices, $P(p^1, p^i, x^1, x^i)$, but rather as the product of the period by period relative price levels; i.e., by travelling along the links of a chain.

The chain principle has one substantial disadvantage: if $p^i = p^j$ and $x^i = x^j$ and periods i and j are not adjacent, then it is not necessarily the case that the measured price level in period i will coincide with the measured price level in period j.

However, the chain principle has a number of advantages: (i) no single period is singled out to play an asymmetric role, (ii) the price levels for I periods are not changed as additional periods are added to the data set, (iii) if a good disappears or a new good is introduced so that N, the number of commodities, changes, then the chain price indexes (and the corresponding implicit quantity indexes) will still be comparable for all periods before and after the change in N and (iv) all superlative indexes will closely approximate each other if the chain principle is used, since changes in prices and quantities tend to be small for adjacent time periods.

The above mentioned disadvantage of the chain method is due to the lack of circularity in the bilateral formula $P(p^1, p^2, x^1, x^2)$. However, experience has shown (e.g., see Fisher [1922]) that deviations from circularity for superlative index number formulae are small in the time series context. In fact, one can show that if the bilateral index P is superlative (or equal to the Paasche or Laspeyres index), then P satisfies the circular test to the first order; i.e., the first order derivatives of $P(p^1, p^3, x^1, x^3)$ and of $P(p^1, p^2, x^1, x^2)P(p^2, p^3, x^2, x^3)$ with respect to the components of p^1, p^2, p^3, x^1, x^2 and x^3 coincide when evaluated at an equal price and quantity point where $p^1 = p^2 = p^3$ and $x^1 = x^2 = x^3$. (This is a new result.)

Our conclusion at this point is that alternatives (ii) and (iii) do not look very attractive in the time series context: the use of a fixed base or the chain principle seems preferable. However, in the context of cross section data where there is no natural way of ordering the data sequentially, alternatives (ii) and (iii) become much more attractive. We consider genuine multilateral index number formulae in Sections 9 and 10 below.

7. Aggregation Over Consumers

Thus far, our discussion of the microeconomic approach to indexes has been limited to the one consumer or producer case (or the case of two consumers or producers who have identical preferences or technologies). In this section, we relax these restrictions and discuss aggregate consumer or household indexes. In the following section, we discuss aggregate output price and quantity indexes.

Consider the case of two countries or regions (or two time periods for the same country). We suppose that there are H_i households in country i and the preferences of household h in country i are represented by a continuous utility function, $F^{ih}(x, z^{ih})$, where $x \equiv (x_1, \ldots, x_N) \geq 0_N$ is a nonnegative vector of market goods, $p^i \equiv (p_1^i, \ldots, p_N^i) \gg 0_N$ is the corresponding vector of positive prices that each household in region i faces and z^{ih} is a vector of demographic variables or consumption of public goods by household h in country i.

The *restricted cost* or *expenditure function* of household h in country i is defined as:

$$(47) \quad C^{ih}(u^{ih}, p^i, z^{ih}) \equiv \min_x \{p^i \cdot x : F^{ih}(x, z^{ih}) \geq u^{ih}\}, \qquad i = 1, 2;\ h = 1, \ldots, H_i,$$

where u^{ih} is a utility or welfare level. We assume that $x^{ih} > 0_N$, the observable consumption vector for household h in the country i, solves (47) with $u^{ih} = F^{ih}(x^{ih}, z^{ih})$ for $i = 1, 2$ and $h = 1, \ldots, H_i$.

We may use the preferences and observed choices of household h in country i to define a *Konüs price index* P^{ih} for the level of prices in country 2 relative to country 1:

$$(48) \quad P^{ih}(p^1, p^2) \equiv C^{ih}(u^{ih}, p^2, z^{ih}) / C^{ih}(u^{ih}, p^1, z^{ih}), \qquad i = 1, 2;\ h = 1, \ldots, H_i.$$

Suppose that each expenditure function C^{ih} has a translog functional form; i.e., $\ln C^{ih}(u, p, z)$ is a quadratic function in the logarithms of its variables, similar to the right hand side of (41) except that now (u, p, z) replaces p. We note that we can approximate arbitrary preferences to the second order using this functional form; i.e., the translog restricted expenditure function is a flexible functional form. Caves, Christensen and Diewert [1982b] establish the following result: if household h in country 1 and household k in country 2 have translog expenditure functions with identical coefficients on the second order terms in commodity prices (this forces some similarity in preferences), then the geometric mean of the two theoretical Konüs price indexes, P^{1h} and P^{2k}, equals the observable Törnqvist or translog price index $P_T(p^1, p^2, x^{1h}, x^{2k})$; i.e., we have

$$(49) \quad [P^{1h}(p^1, p^2) P^{2k}(p^1, p^2)]^{1/2} = P_T(p^1, p^2, x^{1h}, x^{2k}), \qquad h = 1, \ldots, H_1;\ k = 1, \ldots, H_2.$$

Rather than deal with the individual household indexes p^{ih} defined by (48), one can define an average index, where the average is taken over all households in the country. Thus for a nonnegative weights vector $\alpha^i \equiv (\alpha_1^i, \ldots, \alpha_{H_i}^i)$

such that $\sum_{h=1}^{H_i} \alpha_h^i = 1$, define the *country i average price index* P^i as an α^i weighted geometric mean of the individual indexes:

$$(50) \qquad P^i(p^1, p^2, \alpha^i) \equiv \prod_{h=1}^{H_i} P^{ih}(p^1, p^2)^{\alpha_h^i}, \quad i = 1, 2.$$

The theoretical index P^i defined in (50) utilizes the price vectors p^1 and p^2 in both countries but utilizes only the preferences of households in country i. As a final bit of averaging, we take the geometric mean of P^1 and P^2 to obtain a final theoretical price index that treats each household in each country in a symmetric fashion:

$$(51) \qquad P(p^1, p^2, \alpha^1, \alpha^2) \equiv [P^1(p^1, p^2, \alpha^1) P^2(p^1, p^2, \alpha^2)]^{1/2}.$$

The natural choices for the household weighting vectors α^1 and α^2 are: (i) *democratic weights* and (ii) *plutocratic weights*. In case (i), each household in each country is given an equal weight; i.e., $\alpha_h^i \equiv 1/H_i$, $i = 1, 2$, and $h = 1, \ldots, H_i$. In case (ii), each household gets a weight that is proportional to its share of consumption in its own country; i.e., $\alpha_h^i \equiv p^i \cdot x^{ih} / p^i \cdot x^i \equiv s_h^i$ for $i = 1, 2$ and $h = 1, \ldots, H_i$, where $x_i \equiv \sum_{h=1}^{H_i} x^{ih}$ is country i's aggregate consumption vector. If each household has a translog restricted expenditure function with identical coefficients on the second order terms in commodity prices, then making repeated use of (49), we can deduce that the aggregate price index (51) in case (i) is:

$$(52) \quad P(p^1, p^2, 1/H_1, \ldots, 1/H_1,\ 1/H_2, \ldots, 1/H_2) = \prod_{h=1}^{H_1} \prod_{k=1}^{H_2} P_T(p^1, p^2, x^{1h}, x^{2k})^{1/H_1 H_2}$$

while in case (ii), the aggregate price index is

$$(53) \qquad P(p^1, p^2, s^1, s^2) = P_T(p^1, p^2, x^1, x^2)$$

where $s^i \equiv (s_1^i, \ldots, s_{H_i}^i)$ is the country i expenditure share vector for households in country i and $x^i \equiv \sum_{h=1}^{H_i} x^{ih}$ is the aggregate country i consumption vector. Thus if individual household consumption data x^{ih} are available, then the *democratic aggregate price index* defined by (52) can be evaluated as a geometric mean of individual household translog price indexes. If only aggregate country consumption data x^i are available, then the *plutocratic aggregate price index* defined by (53) can be evaluated as a translog price index (recall (11)), using the aggregate consumption vectors x^1 and x^2 as quantity weights.

The terms democratic and plutocratic are due to Prais [1959]. For other approaches to aggregate consumer price indexes, see Prais [1959], Pollak [1981] and Diewert [1983a].

We turn now to the construction of an aggregate quantity index for the two periods. Let the preferences of household h in country i be represented by the continuous, increasing utility function $F^{ih}(x)$ where we have absorbed the vector of demographic variables into the function F^{ih}. For $i = 1, 2$, and $h = 1, \ldots, H_i$, define the household h in country i *deflation function* $D^{ih}(u, x)$ for $x > 0_N$ and u in the range of F by

$$(54) \qquad D^{ih}(u, x) \equiv \max_{\delta}\{\delta : F^{ih}(x/\delta) \geq u,\ \delta \geq 0\}.$$

As in Section 5, define the Malmquist quantity index for $x^* > 0_N$ relative to $x > 0_N$ using the preferences of household h in country i by:

$$(55) \qquad Q^{ih}(x, x^*, u) \equiv D^{ih}(u, x^*)/D^{ih}(u, x), \quad i = 1, 2;\ h = 1, \ldots, H_i.$$

Let the observed consumption vector of household h in country i be $x^{ih} > 0_N$ and define the corresponding utility level by $u^{ih} \equiv F^{ih}(x^{ih})$.

Define the *index of average household consumption* in country j relative to i by:

$$(56) \qquad Q^{ij} \equiv \sum_{k=1}^{H_j} Q^{jk}\left(\sum_{h=1}^{H_i} x^{ih}/H_i, x^{jk}, u^{jk}\right)\Big/H_i; \quad i, j = 1, 2,\ i \neq j.$$

To explain the meaning of (56), define the country i average or per capita consumption vector by

$$(57) \qquad \overline{x}^i \equiv \sum_{h=1}^{H_i} x^{ih}/H_i, \qquad i = 1, 2.$$

Then Q^{12} is an average of the individual country 2 Malmquist indexes $Q^{2k}(\overline{x}^1, x^{2k}, u^{2k})$ which in turn compares the observed consumption vector of household k in country 2 with the average consumption vector $\overline{x}^1$ for country 1, using the indifference surface through x^{2k} of household k in country 2 as the reference indifference surface. Similarly, Q^{21} is an average of the individual country 1 Malmquist indexes $Q^{1k}(\overline{x}^2, x^{1k}, u^{1k})$.

Diewert [1986] showed that under the assumption of expenditure minimizing behavior on the part of consumers in both countries, Q^{12} has the lower bound $p^2 \cdot \overline{x}^2/p^2 \cdot \overline{x}^1$, a Paasche quantity index in per capita quantities, and $[Q^{21}]^{-1}$ has the upper bound $p^1 \cdot \overline{x}^2/p^1 \cdot \overline{x}^1$, a Laspeyres quantity index in per capita quantities. Diewert also shows that there exists a $0 \leq \lambda \leq 1$ such that $\lambda Q^{12} + (1 - \lambda)(Q^{21})^{-1}$ lies between these per capita Paasche and Laspeyres quantity indexes. This suggests that we can approximate the theoretical index $\lambda Q^{12} + (1 - \lambda)(Q^{21})^{-1}$ by an average of these Paasche and Laspeyres indexes such as the Fisher index $Q_F(p^1, p^2, \overline{x}^1, \overline{x}^2) \equiv (p^2 \cdot \overline{x}^2/p^2 \cdot \overline{x}^1)^{1/2}(p^1 \cdot \overline{x}^2/p^1 \cdot \overline{x}^1)^{1/2}$ where the per capita quantity vectors $\overline{x}^i$ are defined by (57).

To summarize this section: we have shown that the translog price index P_T and the Fisher quantity index Q_F, which had very satisfactory economic interpretations in the case of one consumer, also have reasonable economic interpretations in the many consumer case.

8. Aggregation Over Producers

We assume that there are two regions or countries that are to be compared. The two countries could represent the same country at different time periods. The private production sector in country i uses a vector of primary inputs $v^i \equiv (v_1^i, \ldots, v_M^i)$, where v_m^i is the amount of input m used in country i. These inputs are different types of labor, capital, land and other natural resources. There are N net outputs that can be produced by the private production sector in each country. These goods are different types of consumer and investment goods, exports and imports. The net output vector for country i is $y^i \equiv (y_1^i, \ldots, y_N^i)$ and the corresponding price vector is the positive vector $w^i \equiv (w_1^i, \ldots, w_N^i) \gg 0_N$. We assume $w^i \cdot y^i \equiv \sum_{n=1}^N w_n^i y_n^i > 0$ for $i = 1, 2$. If $y_n^i < 0$, then good n is utilized as an input in country i; this good could be an imported good or it could be an intermediate input that is produced by the government sector in country i. Note that in contrast to the consumer case, we no longer assume that quantity vectors are nonnegative.

The technology set for the private production sector in country i is the set $S^i \equiv \{(y^i, v^i)\}$, a feasible set of net output vectors y^i and primary input vectors v^i. If knowledge is freely transferable across countries, then $S^i = S$ for $i = 1, 2$ so that there is a common technology set across countries. However, we do not require this assumption in what follows.

Define country i's *private national product function* g^i by

$$(58) \qquad g^i(w, v) \equiv \max_y\{w \cdot y : (y, v) \quad \text{belongs to} \quad S^i\}, \quad i = 1, 2.$$

The number $g^i(w, v)$ is the maximum value of outputs (less the value of imports) that the private production sector of country i can produce, given that each producer in the country faces the price vector w and the aggregate private economy has at its disposal the primary input vector v. If each producer faces the same price vector w and behaves competitively, then we do not have to concern ourselves with individual producer output vectors: all that matters is the aggregate net output vector. The national product function was introduced into the economics literature by Samuelson [1953–54]; it is sometimes called a variable or restricted profit function or a net revenue function.

The Fisher–Shell [1972a] *output price index* of country 2 relative to 1 using the country i technology set S^i and primary input vector v^i is defined as

$$P^i(w^1, w^2) \equiv g^i(w^2, v^i)/g^i(w^1, v^i), \qquad i = 1, 2. \tag{59}$$

We assume optimizing behavior on the part of producers in each country so that the observed country i price and net output vectors, w^i and y^i, satisfy $w^i \cdot y^i = g^i(w^i, v^i) > 0$ for $i = 1, 2$.

Assume that the private national product function $g^i(w, v)$ has a translog functional form for each i (recall (41) except that now $\ln g^i(w, v)$ is a quadratic form in the logarithms of w_n and v_m) and further assume that the coefficients for the quadratic terms in the logarithms of output prices are the same across the two countries. Then Diewert [1986] showed that the geometric mean of the two theoretical output price indexes defined by (59) is exactly equal to the observable translog price index $P_T(w^1, w^2, y^1, y^2)$, where P_T is defined by (11); i.e.,

$$[P^1(w^1, w^2)P^2(w^1, w^2)]^{1/2} = P_T(w^1, w^2, y^1, y^2). \tag{60}$$

Thus the translog price index P_T again turns out to have a strong microeconomic justification.

The Malmquist quantity index has been applied to the problem of constructing output indexes by Caves, Christensen and Diewert [1982b], and Diewert [1986]. It is first necessary to define the country i output deflation function d^i: for a primary input vector v and a net output vector y, define

$$d^i(y, v) \equiv \min_\delta\{\delta : (y/\delta, v) \text{ belongs to } S^i,\ \delta^i \geq 0\}, \quad i = 1, 2 \tag{61}$$

where S^i is the technology set for country i. Thus $d^i(y, v)$ denotes the amount the net output vector y must be deflated so that the deflated output vector and the reference input vector v are just on the frontier of the country i production possibilities set S^i.

The Malmquist output index of country 2 relative to country 1 using the country i technology and primary input vector v^i is:

$$Q^i(y^1, y^2) \equiv d^i(y^2, v^i)/d^i(y^1, v^i), \qquad i = 1, 2. \tag{62}$$

Assume the observed country i net output vector y^i solves the country i private product maximization problem, $\max_y\{w^i \cdot y : (y, v^i) \text{ belongs to } S^i\}$ for $i = 1, 2$. Then Diewert [1986] established the following bounds for the Malmquist indexes Q^i defined by (62):

$$Q^1(y^1, y^2) \geq w^1 \cdot y^2/w^1 \cdot y^1 \equiv Q_L(w^1, w^2, y^1, y^2); \tag{63}$$

$$Q^2(y^1, y^2) \leq w^2 \cdot y^2/w^2 \cdot y^1 \equiv Q_P(w^1, w^2, y^1, y^2) \tag{64}$$

where Q_L and Q_P are the Laspeyres and Paasche quantity indexes.

It is also possible to define a deflation function d that uses a convex combination of the input vectors, $\lambda v^1 + (1 - \lambda)v^2$, and a convex combination of the technology sets for the two countries, $\lambda S^1 + (1 - \lambda)S^2$:

$$d(y, \lambda) \equiv \min_{\delta > 0}\left\{\delta : \left[y/\delta, [\lambda v^1 + (1 - \lambda)v^2]\right] \text{ belongs to } \lambda S^1 + (1 - \lambda)S^2\right\}. \tag{65}$$

The *Malmquist λ weighted average output index* for country 2 relative to 1 may be defined as:

$$Q^\lambda(y^1, y^2) \equiv d(y^2, \lambda)/d(y^1, \lambda). \tag{66}$$

Assuming maximizing behavior, Diewert [1986] shows that there exists a λ such that $0 \leq \lambda \leq 1$ and $Q^\lambda(y^2, y^2)$ lies between the Laspeyres and Paasche quantity indexes Q_L and Q_P defined in (63) and (64). Thus the Fisher quantity index, $Q_F \equiv (Q_L Q_P)^{1/2}$, should provide an adequate approximation to the theoretical output index $Q^\lambda(y^1, y^2)$.

Our conclusion is that the translog price index P_T and the Fisher quantity index Q_F have reasonably strong justifications in the bilateral aggregate private production context.

We now turn our attention to the multilateral case.

9. Multilateral Test Approaches

We are finally in a position to study the multilateral index number problem which was set out in the introduction. To review the notation, there are I positive price vectors $p^i \equiv (p_1^i, \ldots, p_N^i)$ and I quantity vectors $x^i \equiv (x_1^i, \ldots, x_N^i)$ with $p^i \cdot x^i > 0$ for $i = 1, \ldots, I$. We wish to find $2I$ positive numbers P^i (price indexes) and X^i (quantity indexes) such that $P^i X^i = p^i \cdot x^i$ for $i = 1, \ldots, I$. The I data points (p^i, x^i) will typically be observations on production or consumption units that are separated spatially but yet are still comparable. For the sake of definiteness, we shall refer to the I data points as countries. Each commodity n is supposed to be the same across all countries. This can always be done by a suitable extension of the list of commodities.

Our first approach to the construction of a system of multilateral price and quantity indexes is based on the use of a bilateral quantity index Q. In this method, the first step is to pick the 'best' bilateral index number formula: e.g., the Fisher index Q_F defined by (14) or the implicit translog quantity index defined by $\widetilde{Q}_T(p^1, p^2, x^1, x^2) \equiv p^2 \cdot x^2/p^1 \cdot x^1 P_T(p^1, p^2, x^1, x^2)$ where P_T is defined by (11). Secondly, pick a numeraire country, say country 1, and then calculate the aggregate quantity for each country i relative to country 1 by evaluating the quantity index $Q(p^1, p^i, x^1, x^i)$. In order to put these relative quantity measures on a symmetric footing, we convert each relative to country 1 quantity measure into a share of world quantity by dividing through by $\sum_{k=1}^{I} Q(p^1, p^k, x^1, x^k)$. For a general numeraire country j, define the *share of world quantity for country i, using country j as the numeraire country*, by:

$$(67) \qquad \sigma_i^j(p, x) \equiv Q(p^j, p^i, x^j, x^i) \Big/ \sum_{k=1}^{I} Q(p^j, p^k, x^j, x^k), \quad i = 1, \ldots, I,$$

where $p \equiv (p^1, \ldots, p^I)$ is the $N \times I$ matrix of price data and $x \equiv (x^1, \ldots, x^I)$ is the $N \times I$ matrix of quantity data. Once the numeraire country j has been chosen and the country i shares σ_i^j calculated, we may set $X^i \equiv \sigma_i^j$ and $P^i \equiv p^i \cdot x^i / X^i$ for $i = 1, \ldots, I$. Thus we have provided a solution to the multilateral index number problem (1). Of course, one is free to renormalize the resulting P^i and X^i if desired; i.e., all X^i can be multiplied by a number provided all P^i are divided by this same number. Kravis [1984] calls this method the *star system*, since the numeraire country plays a starring role: all countries are compared with it and it alone.

We shall assume throughout this section that the index number formula Q satisfies the quantity counterpart to the bilateral price tests BT1 through BT6. (The quantity counterpart to BT1 is $Q(\alpha p^1, \beta p^2, x^1, x^2) = 1$ if $p^1 = p^2$ and $x^1 = x^2$, the quantity counterpart to BT2 is $Q(p^1, p^2, x^1, \alpha x^2) = \alpha Q(p^1, p^2, x^1, x^2)$ for $\alpha > 0$, and so on.) This assumption is not restrictive since our best bilateral formulae, Q_F and $\widetilde{Q}_T$, both satisfy these tests.

Of course, the problem with the star system for making multilateral comparisons is its lack of invariance to the choice of the numeraire or star country. Different choices for the base country will in general give rise to different indexes P^i and X^i. This problem can be traced to the lack of circularity of the bilateral formula Q: if Q satisfies the time reversal test BT5 and the circular test BT9 for quantity indexes, then $\sigma_i^j = \sigma_i^k$ for all i, j and k; i.e., the shares σ_i^j defined by (67) do not depend on the choice of the numeraire country j. However, given that the bilateral formula Q does not satisfy the circularity test (as is the case with Q_F and $\widetilde{Q}_T$), how can we generate multilateral indexes that treat each country symmetrically?

Fisher [1922] recognized that the simplest way of achieving symmetry was to average base specific index numbers over all possible bases. Thus define country i's share of world output $S_i(p, x)$ by

$$(68) \qquad S_i(p, x) \equiv \sum_{j=1}^{I} \sigma_i^j(p, x)/I, \quad I = 1, \ldots, I$$

where the σ_i^j are defined by (67). We can now define country i quantities and prices by

$$(69) \qquad X^i \equiv S_i(p, x), \; P^i \equiv p^i \cdot x^i / X^i, \quad i = 1, \ldots, I.$$

Fisher [1922] called this method of constructing multilateral indexes the *blend method* while Diewert [1986] called it the *democratic weights method*, since each share of world output using each country as the base is given an equal weight in the formation of the average.

Of course, there is no need to use an arithmetic average of the σ_i^j as in (68); one can use a geometric average:

$$(70) \qquad \sigma_i(p, x) \equiv \left[\prod_{j=1}^{I} \sigma_i^j(p, x) \right]^{1/I}, \quad i = 1, \ldots, I.$$

Using (70), the resulting shares no longer sum to one in general, so country i's share of world output is now defined as:

$$(71) \qquad S_i(p, x) \equiv \sigma_i(p, x) \Big/ \sum_{k=1}^{I} \sigma_k(p, x), \quad i = 1, \ldots, I.$$

If the Fisher index Q_F is used in the definition of the σ_i^j, then

$$(72) \quad S_i(p, x)/S_j(p, x) = \left[\prod_{k=1}^{I} Q_F(p^k, p^i, x^k, x^i) \Big/ \prod_{m=1}^{I} Q_F(p^m, p^j, x^m, x^j) \right]^{1/I}$$

and in this case, the multilateral method defined by (70) reduces to a method recommended by Eltetö and Köves [1964] and Szulc [1964], the *EKS method*. Instead of using the Fisher formula in (72), Caves, Christensen and Diewert [1982a] advocated the use of the translog quantity index Q_T while Diewert [1986] suggested the use of the implicit translog quantity index $\widetilde{Q}_T$, since $\widetilde{Q}_T$ is well defined even in the case where some quantities x_n^i are negative (whereas Q_T is not). We call the indexes generated by (69) and (71) for a general bilateral index Q, *generalized EKS* indexes.

When forming averages of the σ_i^j as in (68) or (70), there is no necessity to use equal weights: one can define country j's value share of world output as $\beta_j \equiv p^j \cdot x^j / \sum_{k=1}^{I} p^k \cdot x^k$ (this requires all prices to be measured in units of a common currency) and then we may define a plutocratic share weighted average of the σ_i^j:

$$(73) \qquad S_i(p, y) \equiv \sum_{j=1}^{I} \beta_j(p, x) \sigma_i^j(p, x).$$

Diewert [1986] called this method of constructing multilateral indexes the *plutocratic weights method*.

Another multilateral method that is based on a bilateral index Q may be described as follows. Define

$$(74) \qquad \alpha_i(p, x) \equiv \sum_{j=1}^{I} [Q(p^j, p^i, x^j, x^i)^{-1}]^{-1}, \quad i = 1, \ldots, I.$$

If there is only one commodity so that $N = 1$ and the bilateral index Q satisfies BT1, BT2 and BT3, then $\alpha_i = \left[\sum_{j=1}^{I}(x^i/x^j)^{-1}\right]^{-1} = \left[\sum_{j=1}^{I} x^j/x^i\right]^{-1} = x^i/\sum_{j=1}^{I} x^j$ which is country i's share of world product. In the general case where $N > 1$, the 'shares' α_i do not necessarily sum up to unity, so it is necessary to normalize them:

$$(75)\qquad S_i(p,x) \equiv \alpha_i(p,x)\Big/\sum_{k=1}^{I}\alpha_k(p,x), \quad i = 1,\ldots,I.$$

Diewert [1986] called this the *own share method* for making multilateral comparisons.

The above methods for achieving consistency and symmetry rely on averaging over various bilateral index number comparisons. Fisher [1922] realized that symmetry could be achieved by making comparisons with an average; he called this broadening the base. Thus the *basket method* (which corresponds to Fisher's [1922] formula 6053 and to method 8 described in Ruggles [1967]) may be described as follows. The price level of country i relative to country j is set equal to $p^i \cdot (\sum_{k=1}^{J} x^k/I)/p^j \cdot (\sum_{k=1}^{J} x^k/I)$. This index number formula is a Scrope index (6), where the reference quantity vector is chosen to be the average market basket, $\sum_k x^k/I$. The same result could be achieved if we chose x to be the total market basket, $\sum_k x^k$. Now define $Q^{ji} \equiv p^i \cdot x^i/p^j \cdot x^j[p^i \cdot (\sum_k x^k)/p^j \cdot (\sum_k x^k)]$ to be the implicit output of country i relative to j. Choose a j as a numeraire country and calculate country i's share of world output as:

$$(76)\qquad S_i(p,x) \equiv Q^{ji}\Big/\sum_{k=1}^{I} Q^{jk} = \left(p^i \cdot x^i/p^i \cdot \sum_k x^k\right)\Big/\sum_{m=1}^{I}\left(p^m \cdot x^m/p^m \sum_k x^k\right), \quad i = 1,\ldots,I.$$

Note that the final expression for S_i does not depend on the choice of the numeraire country j. As usual, once the share functions, S_i, have been defined, the aggregate X^i and P^i may be defined by (69).

A variation on the basket method due to Geary [1958] and Khamis [1972] is defined by (77)–(79) below:

$$(77)\qquad \pi_n \equiv \sum_{i=1}^{I} p_n^i x_n^i/P^i \sum_{k=1}^{I} x_n^k, \qquad n = 1,\ldots,N;$$

$$(78)\qquad P^i \equiv \sum_{n=1}^{N} p_n^i x_n^i\Big/\sum_{m=1}^{N}\pi_m x_m^i, \quad i = 1,\ldots,I;$$

$$(79)\qquad X^i \equiv p^i \cdot x^i/P^i, \qquad i = 1,\ldots,I.$$

π_n is interpreted as an average international price for good n. From (78), it can be seen that P^i, the price level or purchasing power parity for country i, is a Paasche-like price index for country i except that the base prices are chosen to be the international prices π_n. The π_n and $(P^i)^{-1}$ can be solved for as a system of simultaneous linear equations (up to a scalar normalization) or the $(P^i)^{-1}$ may be determined as the components of the eigenvector that corresponds to the maximal positive eigenvalue of a certain matrix. The P^i can be normalized so that the quantities X^i defined by (79) sum up to unity. This *GK method* for making multilateral comparisons has been widely used in empirical applications: see Kravis *et al.* [1975].

We have defined seven methods for making multilateral comparisons: the star method (67), the democratic (68) and plutocratic (73) weights methods, the generalized EKS method (71), the own share method (74), the basket method (76) and the GK method (79). How can we discriminate among them?

One helpful approach would be to define a system of *multilateral tests* and then evaluate how the above methods satisfy these tests.

In the bilateral situation, it was natural to phrase the tests in terms of the price index P or the quantity index Q, since if either of these functions were given (along with a single normalization such as $P^1 = 1$), then the aggregates P^1, P^2, X^1 and X^2 were all determined.

In the multilateral situation, it seems natural to phrase the tests in terms of the properties of the system of world output share functions, $S(p,x) \equiv [S_1(p,x),\ldots,S_I(p,x)]$, since given these share functions, we may set $X^i \equiv S_i(p,x)$ and $P^i = p^i \cdot x^i/X^i$ for $i = 1,\ldots,I$.

The tests MT1 to MT6 listed below are multilateral counterparts to the bilateral tests BT1 to BT6 applied to quantity indexes rather than price indexes. MT0 is a preliminary test that does not have a bilateral counterpart. Recall that $p \equiv (p^1,\ldots,p^I)$ and $x \equiv (x^1,\ldots,x^I)$.

MT0: *Share Test:* $\sum_{i=1}^{I} S_i(p,x) = 1$.

MT1: *Multilateral Identity Test:* $S_i(\alpha_1 p^1,\ldots,\alpha_1 p^I,\beta_1 x^1,\ldots,\beta_I x^I) = \beta_i$ for $i = 1,\ldots,I$ for all $\alpha_i > 0$, $\beta_i > 0$ if $p^1 = \cdots = p^I$, $x^1 = \cdots = x^I$ and $\sum_j \beta_j = 1$.

MT2: *Proportionality Test:* For $i = 1,\ldots,I$ and $\lambda_i > 0$, $S_i(p,x^1,\ldots,x^{i-1},\lambda_i x^i,x^{i+1},\ldots,x^I)/S_j(p,x^1,\ldots,x^{i-1},\lambda_i x^i,x^{i+1},\ldots,x^I) = \lambda_i S_i(p,x)/S_j(p,x)$ for $j = 1,\ldots,i-1,i+1,\ldots,I$.

MT3: *Invariance to Changes in Scale Test:* $S_i(\alpha_1 p^1,\ldots,\alpha_I p^I,\beta x^1,\ldots,\beta x^I) = S_i(p,x)$ for all $\alpha_i > 0$, $\beta > 0$, $i = 1,\ldots,I$.

MT4: *Invariance to Changes in Units Test:* $S_i(\alpha_1 p_1^1,\ldots,\alpha_N p_N^1;\ \ldots;\ \alpha_1 p_1^I,\ldots,\alpha_N p_N^I;\ \alpha_1^{-1}x_1^1,\ \ldots,\ \alpha_N^{-1}x_N^1;\ \ldots;\ \alpha_1^{-1}x_1^I,\ \ldots,\ \alpha_N^{-1}x_N^I) = S_i(p,x)$ for $i = 1,\ldots,I$ and $\alpha_1 > 0,\ldots,\alpha_N > 0$.

MT5: *Symmetric Treatment of Countries Test:* Let $\hat{p}$ denote a permutation of the I columns of the $N \times I$ matrix p, let $\hat{x}$ denote the same permutation of the I columns of x, and let $\widehat{S}(p,x)$ denote the same permutation of the I columns of the row vector $S(p,x)$. Then $\widehat{S}(p,x) = S(\hat{p},\hat{x})$.

MT6: *Symmetric Treatment of Commodities Test:* Let $\tilde{p}$ denote a permutation of the N rows of p and let $\tilde{x}$ denote the same permutation of the N rows of x. Then $S_i(\tilde{p},\tilde{x}) = S_i(p,x)$ for $i = 1,\ldots,I$.

MT7: *Country Partitioning Test:* Let $S_j^I \equiv S_j(p^1,\ldots,p^I;\, x^1,\ldots,x^I)$ for $j = 1,\ldots,I$ and let $0 < \lambda_i < 1$. Define $S_j^{I+1} \equiv S_j[p^1,\ldots,p^I,p^i;\, x^1,\ldots,x^{i-1},\, \lambda_i x^i,\, x^{i+1},\ldots,x^I,\, (1-\lambda_i)x^i]$ for $j = 1,\ldots,I,\ I+1$. Then $S_j^I = S_j^{I+1}$ for $j = 1,\ldots,i-1,i+1,\ldots,I$, $S_i^{I+1} = \lambda_i S_i^I$ and $S_{I+1}^{I+1} = (1-\lambda_i)S_i^I$. This property is to hold no matter which country i is partitioned.

The functions S_j^I are the share functions for the initial world economy that consists of I countries. The functions S_j^{I+1} are the share functions for a new world economy, where the original country i with price vector p^i and quantity vector x^i, has been partitioned into two countries with price vectors p^i and p^i and quantity vectors $\lambda_i x^i$ and $(1-\lambda_i)x^i$. MT7 says that under these conditions, the original country i share S_i^I splits into $\lambda_i S_i^I$ and $(1-\lambda_i)S_i^I$ and the remaining shares are unaffected.

MT8: *Irrelevance of Tiny Countries Test:* Let $\lambda_i > 0$ and define $S_j^I(\lambda_i) \equiv S_j(p,x^1,\,\ldots,\,x^{i-1},\,\lambda_i x^i,x^{i+1},\,\ldots\,,x^I)$ for $j = 1,\ldots,I$. Define $S_j^{I-1} \equiv S_j(p^1,\,\ldots\,,\,p^{i-1},\,p^{i+1},\ldots,p^I;\, x^1,\ldots,x^{i-1},\, x^{i+1},\ldots,x^I)$ for $j = 1,\ldots,I-1$. Then $\lim_{\lambda_i\to 0} S_j^I(\lambda_i) = S_j^{I-1}$ for $j = 1,\ldots,i-1$ and $\lim_{\lambda_i\to 0} S_j^I(\lambda_i) = S_{j-1}^{I-1}$ for $j = i+1,\, i+2,\ldots,I$. This property is to hold for all choices of the disappearing country i.

In the above test, the quantity vector for country i is deflated down to a zero vector. Consider the resulting system of limiting share functions. For all countries except i, the limiting share is equal to the share we would get if we simply deleted the data for country i and defined a system of share functions for a world economy consisting of only $I-1$ countries. MT8 is a country counterpart to the irrelevance of tiny commodities test, (12).

The above multilateral tests were proposed by Diewert [1986]. Our final multilateral test is a new one.

MT9: *Bilateral Property Test:* When $I = 2$, $S_2(p^1,p^2,\,x^1,x^2)/S_1(p^1,p^2,\,x^1,x^2)$ satisfies the bilateral tests BT1–BT6 for quantity indexes.

When $I = 2$, the multilateral system collapses down to a bilateral system. Hence it seems perfectly sensible to demand that S^2/S^1 should satisfy tests BT1–BT6 at least, since our best bilateral indexes, Q_F and $\widetilde{Q}_T$, satisfied these tests.

The seven multilateral methods mentioned above satisfy most of the multilateral tests, assuming that the bilateral index Q satisfies BT1–BT6 (recall that all of the multilateral methods utilize a bilateral index Q except the basket and GK methods).

The star system fails MT5 (it is obviously not symmetric) and it fails MT8 when the tiny country is chosen to be the numeraire country.

The plutocratic weights method fails MT2 and MT3; thus the resulting quantity indexes are not invariant to country inflation rates, a very severe defect.

The democratic weights method fails the multilateral proportionality test MT2 and the two consistency in aggregation tests, MT7 and MT8. This method is dominated by the generalized EKS method which fails only MT7 and MT8.

The basket and GK methods fail MT2 and MT9: both methods fail BT2 when $I = 2$. When $I = 2$, $S_2/S_1 = p^2\cdot x^2 p^1\cdot(x^1+x^2)/p^1\cdot x^1 p^2\cdot(x^1+x^2)$ for the basket method, and $S_2/S_1 = p^2\cdot x^2/p^1\cdot x^1 P_{GK}(p^1,p^2,x^1,x^2)$ for the GK method where the GK bilateral price index is defined as

$$(80)\quad P_{GK}(p^1,p^2,x^1,x^2) \equiv \left[\sum\nolimits_{n=1}^N p_n^2 x_n^1 x_n^2/(x_n^1+x_n^2)\right] \Big/ \left[\sum\nolimits_{k=1}^N p_k^1 x_k^1 x_k^2/(x_k^1+x_k^2)\right].$$

The five multilateral methods that use a bilateral index Q as a building block all have the property that $S_2/S_1 = Q$ when $I = 2$, which explains why these methods pass MT9.

Unfortunately, none of the above multilateral methods satisfies all nine multilateral tests. Our tentative conclusion is that if a symmetric multilateral method is desired, then the choice seems to be between the EKS (which fails MT7 and MT8, the consistency in aggregation properties) and the own share method (which fails the multilateral proportionality test, MT2). However, the systematic study of multilateral methods has only begun, so it may well be that better methods will be discovered in the future.

We turn now to another class of methods for constructing multilateral indexes.

10. Neostatistical Approaches to Multilateral Indexes

In one of the early statistical approaches to the construction of a price index, P^2/P^1, the index was found by minimizing the sum of squared residuals, $\sum_{n=1}^N (p_n^2/p_n^1 - P^2/P^1)^2$ with respect to P^2/P^1. The resulting price index turns out to equal (4). Note that this price index was defined independently of quantities.

Theil [1960] initiated a *neostatistical approach* where the price and quantity indexes, P^i and X^i, are simultaneously determined. *Theil's best linear price and quantity indexes* may be found by solving the following constrained minimization problem:

$$(81)\quad \min_{P^1,\ldots,P^I,X^1,\ldots,X^J} \sum\nolimits_{i=1}^I \sum\nolimits_{j=1}^I (p^i\cdot x^j - P^i X^j)^2$$

subject to a normalization on the P^i or X^i such as

(82) $$\sum_{i=1}^{I} X^i = 1,$$

which means that we can interpret the Theil X^i as shares. We may define $e_{ij} \equiv p^i \cdot x^j - P^i X^j$ as an error and then the interpretation of (81) becomes straightforward: we choose the P^i and X^i to minimize the sum of squared errors subject to (82).

The Theil index numbers have not been widely used, since they do not satisfy the product test equalities:

(83) $$P^i X^i = p^i \cdot x^i, \qquad i = 1, \ldots, I.$$

Kloek and de Wit [1961] suggested a number of variants of the Theil indexes including one where the constraints (83) were imposed. Thus define the Kloek and de Wit multilateral indexes as the P^i and X^i which solve (81) subject to (82) and (83). Unfortunately, the resulting indexes do not have very satisfactory properties: they fail the multilateral tests MT2, MT3, MT7, MT8 and MT9 (when $I = 2$, the resulting bilateral quantity index fails BT2 and BT3 for quantity indexes.)

Another neostatistical approach has been suggested by van Yzeren [1956] which he called the *balanced method*. The price indexes P^i are determined (up to a normalization or factor of proportionality) by solving the following minimization problem:

(84) $$\min_{P^1,\ldots,P^I} \sum_{i=1}^{I} \sum_{j=1}^{I} (P^i)^{-1} (p^i \cdot x^j / p^j \cdot x^j) P^j.$$

The X^i are then determined by (83) and the price normalization may be chosen so that (82) is satisfied. In this case, the errors e_{ij} may be defined by $(P^i X^j)^{-1} (p^i \cdot x^j / p^j \cdot x^j) P^j X^j = 1 + e_{ij}$ and the P^i may be found by minimizing

$$\sum_{i=1}^{I} \sum_{j=1}^{I} e_{ij}$$

subject to a normalization.

The van Yzeren system of share functions does rather well in the multilateral test examination: the share functions pass all tests except the consistency in aggregation tests MT7 and MT8. When $I = 2$, X^2/X^1 turns out to equal the Fisher quantity index, Q_F. Thus the balanced method satisfies the same tests as the EKS method. However, the EKS method has a more satisfactory economic interpretation and is easier to construct numerically.

Many other neostatistical approaches to the construction of multilateral indexes could be explored. However, the resulting methods seem to be rather arbitrary and, moreover, they lack economic interpretations.

11. Other Aspects

There are many aspects of index number theory that we cannot cover in this brief survey, such as: (i) *sampling problems* (see Fisher [1922] and Allen [1975]), (ii) the treatment of *seasonality* (see Turvey [1979], Balk [1980] and Diewert [1983c]), (iii) consistency in aggregation and the theory of *subindexes* (see Vartia [1976a], Pollak [1975] and Diewert [1983a]), (iv) productivity indexes (see Jorgenson and Griliches [1967], Caves, Christensen and Diewert [1982b] and Denny and Fuss [1983]), and (v) *econometric approaches* to cost of living indexes (see Jorgenson and Slesnick [1983]).

However, one area of concern that must be discussed is the *new goods problem*. Suppose that we are in the time series context and we have price and quantity data for $N-1$ commodities in periods 1 and 2, p_n^t and x_n^t for $t = 1, 2$ and $n = 1, \ldots, N-1$. Suppose in addition, that x_N^2 units of a new good are sold at the price p_N^2 during period 2. How are we to compute the bilateral price index, $P(p^1, p^2, x^1, x^2)$, when we do not know p_N^1, the price of the new good in period 1? Of course, we can assume $x_N^1 = 0$, so determining the quantity of the new good in period 1 is no problem.

From the viewpoint of the microeconomic approach to index number theory, Hicks [1940] provided a formal solution to this new good problem: if we are in the consumer context, p_N^1 should be the price which would just make the consumer's demand for good N in period 1 equal to zero. The practical problem is that this shadow price is not observable: we require a knowledge of the consumer's indifference surfaces to calculate it. Of course, econometric techniques could be used to estimate these shadow prices (see Diewert [1980] for an example of such a technique in the producer context), but most index number practitioners will find it inconvenient to resort to econometrics. In practice, most official indexes ignore the existence of new goods.

In order to illustrate the price index bias that can result from the omission of new goods, we shall present a hypothetical example. Suppose that there are three periods, one 'old' good with constant price and quantity, $p_1^t = x_1^t = 1$ for $t = 1, 2, 3$ and one 'new' good which appears in period 2, so that $x_2^1 = 0$. Typically, new goods follow a product cycle: they are introduced at a relatively high price and then the price declines over time. Thus we assume that $p_2^2 = 2$ and $p_2^3 = 1$, so that the period 2 price for the new good is twice as high as the period 3 price. We assume that the quantity purchased of the new good in period 2 is $f > 0$, where f is a fraction which represents the period 2 proportion of new goods to old goods. We assume that the quantity purchased of the new good in period 3 is $2f$ and that the shadow price p_2^1 in period 1 that would make the demand for the new good equal to zero is 4.

If the new good is ignored, we find that $P^t = 1$ for $t = 1, 2, 3$ for any reasonable index number formula. The true chain Laspeyres price indexes which do not ignore the new good are: $P^1 = 1$, $P^2 = 1$ and $P^3 = (1+f)/(1+2f)$. The reader can verify that the same indexes result no matter what shadow price p_2^1 is chosen. Evaluating P^3 for various reasonable values of f yields the following period 3 price indexes: if $f = 0.01$, then $P^3 = 0.9902$; if $f = 0.02$, then $P^3 = 0.9808$ and if $f = 0.05$, then $P^3 = 0.9545$. Thus the conventional Laspeyres price index which ignores the existence of new goods will have an *upward bias* of about 1 to 4.5 percent compared with the true Laspeyres index.

In order to evaluate the bias in the conventional chained Paasche and Fisher price indexes, we have to use our assumption that $p_2^1 = 4$. Under our assumptions, we obtain the following values for the true chained Paasche price indexes in period 3: if $f = 0.01$, then $P^3 = 0.9619$; if $f = 0.02$, then $P^3 = 0.9273$ and if $f = 0.05$, then $P^3 = 0.8403$. We also obtain the following values for the true chained Fisher price indexes in period 3: if $f = 0.01$, then $P^3 = 0.9759$; if $f = 0.02$, then $P^3 = 0.9537$ and if $f = 0.05$, then $P^3 = 0.8956$. Thus the conventional Fisher ideal price index will have an upward bias of about 2.5 to 10.5 percent in period 3, depending on the fraction f of new goods introduced in period 2.

The above analysis of bias is only illustrative but it does indicate that ignoring new goods could lead to a substantial overestimation of price inflation and a corresponding underestimation of real growth rates, especially in advanced market economies where millions of new goods are introduced each year.

References for Chapter 5

Afriat, S.N., 1972b. "The Theory of International Comparisons of Real Income and Prices." In *International Comparisons of Prices and Outputs*, D.J. Daly (ed.), National Bureau of Economic Research, New York: Columbia University Press, 13–69.

Allen, R.G.D., 1949. "The Economic Theory of Index Numbers," *Economica N.S.* 16, 197–203.

Allen, R.G.D., 1975. *Index Numbers in Theory and Practice*, London: Macmillan.

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

Bowley, A.L., 1928. "Notes on Index Numbers," *Economic Journal* 38, 216–237.

Carli, G.R., 1764. "Del valore e della proporzione de'metalli monetati con i generi in Italia prima delle scoperte dell'Indie colonfronto del valore e della proporzione de'tempi nostri." In *Opere scelte di Carli*, Vol. 1, Milan, 299–366.

Caves, D.W., L.R. Christensen, and W.E. Diewert, 1982a. "Multilateral Comparisons of Output, Input and Productivity Using Superlative Index Numbers," *Economic Journal* 92, 73–86.

Caves, D.W., L.R. Christensen, and W.E. Diewert, W.E., 1982b. "The Economic Theory of Index Numbers and the Measurement of Input, Output and Productivity," *Econometrica* 50, 1393–1414.

Christensen, L.R., D.W. Jorgenson, and L.J. Lau, 1971. "Conjugate Duality and the Transcendental Logarithmic Production Function," *Econometrica* 39, 255–256.

Denny, M. and M. Fuss, 1983. "A General Approach to Intertemporal and Interspatial Productivity Comparisons." *Journal of Econometrics* 23, 315–330.

Diewert, W.E., 1974a. "Applications of Duality Theory." In *Frontiers of Quantitative Economics*, Vol. II, M.D. Intriligator and D.A. Kendrick (eds.), Amsterdam: North–Holland, 106–171.

Diewert, W.E., 1976a. "Exact and Superlative Index Numbers." *Journal of Econometrics* 4, 115–145, and reprinted as Ch. 8 in THIS VOLUME, 223-252.

Diewert, W.E., 1978b. "Superlative Index Numbers and Consistency in Aggregation." *Econometrica* 46, 883–900, and reprinted as Ch. 9 in THIS VOLUME, 253–273.

Diewert, W.E., 1980. "Aggregation Problems in the Measurement of Capital." In *The Measurement of Capital*, D. Usher (ed.), Studies in Income and Wealth, Vol. 45, National Bureau of Economic Researach, Chicago: University of Chicago Press, 433–528, and reprinted in Diewert and Nakamura [1993].

Diewert, W.E., 1981a. "The Economic Theory of Index Numbers: A Survey." In *Essays in the Theory and Measurement of Consumer Behaviour in Honour of Sir Richard Stone*, A. Deaton (ed.), London: Cambridge University Press, 163–208, and reprinted as Ch. 7 in THIS VOLUME, 177–221.

Diewert, W.E., 1983a. "The Theory of the Cost-of-Living Index and the Measurement of Welfare Change." In Diewert and Montmarquette [1983; 163–233], reprinted in Diewert [1990; 79–147] and in Diewert and Nakamura [1993].

Diewert, W.E., 1983c. "The Treatment of Seasonality in a Cost-of-Living Index." In Diewert and Montmarquette [1983; 1019–1045], reprinted in Diewert and Nakamura [1993].

Diewert, W.E., 1986. Microeconomic Approaches to the Theory of International Comparisons. Technical Working Paper 53, National Bureau of Economic Research, Cambridge, Massachusetts.

Diewert, W.E. (ed.), 1990. *Price Level Measurement*, Amsterdam: North-Holland.

Diewert, W.E. and C. Montmarquette, 1983. *Price Level Measurement: Proceedings from a conference sponsored by Statistics Canada*, Ottawa: Statistics Canada.

Diewert, W.E. and A.O. Nakamura, 1993. *Essays in Index Number Theory*, Vol. II, Amsterdam: North-Holland, forthcoming.

Dutot, C., 1738. *Réflexions politiques sur les finances et le commerce*, The Hague.

Eichhorn, W. and J. Voeller, 1976. *Theory of the Price Index: Fisher's Text Approach and Generalizations*, Lecture Notes in Economics and Mathematical Systems, Vol. 140, Berlin: Springer–Verlag.

Eltetö, O. and P. Köves, 1964. "On a Problem of Index Number Computation Relating to International Comparison," *Statisztikai Szemle* 42, 507–518.

Fisher, F.M. and K. Shell, 1972a. "The Pure Theory of the National Output Deflator." In *The Economic Theory of Price Indices*, F.M. Fisher and K. Shell (eds.), New York: Academic Press, 49–113.

Fisher, I., 1911. *The Purchasing Power of Money*. London: Macmillan.

Fisher, I., 1922. *The Making of Index Numbers*. Boston: Houghton Mifflin.

Frisch, R., 1930. "Necessary and Sufficient Conditions Regarding the Form of an Index Number Which Shall Meet Certain of Fisher's Tests," *American Statistical Association Journal* 25, 397–406.

Frisch, R., 1936. "Annual Aurvey of General Economic Theory: The Problem of Index Numbers," *Econometrica* 4, 1–39.

Geary, R.G., 1958. "A Note on Comparisons of Exchange Rates and Purchasing Power Between Countries," *Journal of the Royal Statistical Society*, Series A 121, 97–99.

Hicks, J.R., 1940. "The Valuation of the Social Income," *Economica* 7, 105–124.

Jevons, W.S., 1865. "Variations of Prices and the Value of Currency since 1762," *Journal of the Royal Statistical Society* 28 (June), 294–325.

Jorgenson, D.W. and Z. Griliches, 1967. "The Explanation of Productivity Change," *Review of Economic Studies* 34, 249–283.

Jorgenson, D.W. and D.T. Slesnick, 1983. "Individual and Social Cost-of-Living Indexes." In Diewert and Montmarquette [1983; 241–323], and reprinted in Diewert [1990; 155–234].

Keynes, J.M., 1930. *Treatise on Money*, Vol. 1. London: Macmillan.

Khamis, S.H., 1972. "A New System of Index Numbers for National and International Purposes," *Journal of the Royal Statistical Society*, Series A 135, 96–121.

Kloek, T. and G.M. de Wit, 1961. "Best Linear Unbiased Index Numbers," *Econometrica* 29, 602–616.

Konüs, A.A., 1924. English translation, titled "The Problem of the True Index of the Cost of Living," published in 1939 in *Econometrica* 7, 10–29.

Konüs, A.A. and S.S. Byushgens, 1926. "K probleme pokupatelnoi cili deneg" (English translation of Russian title: "On the Problem of the Purchasing Power of Money"), *Voprosi Konyunkturi* II(1) (supplement to the Economic Bulletin of the Conjuncture Institute), 151–172.

Kravis, I.B., 1984. "Comparative Studies of National Incomes and Prices," *Journal of Economic Literature* 22, 1–39.

Kravis, I.B., Z. Kenessey, A. Heston, and R. Summers, 1975. *A System of International Comparisons of Cross Product and Purchasing Power*, Baltimore: Johns Hopkins University Press.

Laspeyres, E., 1871. "Die Berechnung einer mittleren Waarenpreissteigerung," *Jahrbücher für Nationalökonomie und Statistik* 16, 296–315.

Malmquist, S., 1953. "Index Numbers and Indifference Surfaces," *Trabajos de Estadistica* 4, 209–242.

Marshall, A., 1887. "Remedies for Fluctuations of General Prices." *Contemporary Review* 51, 355–375. Reprinted as Ch. 8 in *Memorials of Alfred Marshall*, A.C. Pigou (ed.), London: Macmillan, 1925.

Paasche, H., 1874. "Über die Preisentwicklung der letzten Jahre nach den Hamburger Börsennotirungen. *Jahrbücher für Nationalökonomie und Statistik* 23, 168–178.

Pigou, A.C., 1912. *Wealth and Welfare*, London: Macmillan.

Pollak, R.A., 1971a. "The Theory of the Cost of Living Index," Research Discussion Paper 11, Office of Prices and Living Conditions, Bureau of Labor Statistics, Washington, D.C. In Diewert and Montmarquette [1983; 87–161], and reprinted in Diewert [1990; 5–77] and Pollak [1989; 3–52].

Pollak, R.A., 1975. "Subindexes of the Cost of Living," *International Economic Review* 16, 135–150, and reprinted in Pollak [1989; 53–69].

Pollak, R.A., 1981. "The Social Cost of Living Index," *Journal of Public Economics* 15, 311–336, and reprinted in Pollak [1989; 128–152].

Pollak, R.A., 1989. *The Theory of the Cost-of-Living Index*, Oxford: Oxford University Press.

Prais, S., 1959. "Whose Cost of Living?" *Review of Economic Studies* 26, 126–134.

Ruggles, R., 1967. "Price Indexes and International Price Comparisons." In *Ten Economic Studies in the Tradition of Irving Fisher*, W. Fellner et al. (ed.), New York: John Wiley, 171–205.

Samuelson, P.A., 1947. *Foundations of Economic Analysis*, Cambridge, Mass.: Harvard University Press.

Samuelson, P.A., 1953–54. "Prices of Factors and Goods in General Equilibrium," *Review of Economic Studies* 21, 1–20.

Samuelson, P.A. and S. Swamy, S., 1974. "Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis," *American Economic Review* 64, 566–593.

Scrope, G.P., 1833. *Principles of Political Economy*, London: Longman, Rees, Orme, Brown, Green and Longman.

Shephard, R.W., 1953. *Cost and Production Functions*, Princeton: Princeton University Press.

Stone, R., 1956. *Quantity and Price Indexes in National Accounts*, Paris: Organization for European Cooperation.

Szulc (Schultz), B.J., 1964. "Indices for Multiregional Comparisons," *Przeglad Statystyczny* (*Statistical Review*) 3, 239–254.

Theil, H., 1960. "Best Linear Index Numbers of Prices and Quantities," *Econometrica* 28, 464–480.

Törnqvist, L., 1936. "The Bank of Finland's Consumption Price Index," *Bank of Finland Monthly Bulletin* 10, 1–8.

Turvey, R., 1979. "The Treatment of Seasonal Items in Consumer Price Indices," *Bulletin of Labour Statistics*, 4th quarter, Geneva: International Labour Office, 13–33.

van Yzeren (van Ijzerin), J., 1956. *Three Methods of Comparing the Purchasing Power of Currencies*, Statistical Studies No. 7, The Hague: The Netherlands Centraal Bureau of Statistics.

Vartia, Y.O., 1976a. "Ideal Log-Change Index Numbers," *Scandinavian Journal of Statistics* 3, 121–126.

Vartia, Y.O., 1985. "Defining Descriptive Price and Quantity Index Numbers: An Axiomatic Approach." Paper presented at the Fourth Karlsruhe Symposium on Measurement in Economics, University of Karlsruhe, July.

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

Walsh, C.M., 1921. "The Best Form of Index Number: Discussion," *Quarterly Publication of the American Statistical Association* 17 (March), 537–544.

Westergaard, H., 1890. *Die Grundzüge der Theorie der Statistik*, Jena: Fischer.