W. Erwin Diewert is a professor at the University of British Columbia.

Commentary

W. Erwin Diewert

In the first section of this commentary, I review the main points made by Matthew D. Shapiro and David W. Wilcox in their excellent article, "Alternative Strategies for Aggregating Prices in the CPI." In the remaining sections of this commentary, I broaden my discussion and consider the following four general measurement problems related to the Consumer Price Index (CPI) and inflation measurement: alternative approaches to the determination of the "correct" functional form for the CPI, the problems raised by the seasonality of prices and quantities, the problems that occur when aggregating prices at the lowest level of aggregation (i.e., the new goods and new outlets problems), and the determination of the "correct" domain of definition for the CPI or a general measure of inflation. The final section provides conclusions.

A SUMMARY OF THE MAIN RESULTS IN SHAPIRO AND WILCOX

Shapiro and Wilcox used the database of Aizcorbe and Jackman (1993) on U.S. CPI components for the years 1984-94. This data set consists of monthly price relatives and monthly or quarterly expenditure shares on 207 classes of goods and services for 44 U.S. regions, or 9,108 commodities in all. Shapiro and Wilcox constructed a variety of monthly price indexes, using the superlative Törnqvist and Fisher functional forms, as well as the fixed-basket Laspeyres form and the geometric functional form. They obtained five main results:

- The two superlative indexes, the Törnqvist and Fisher ideal price indexes, approximated each other very closely. This confirms the previous numerical results of Fisher (1922), Diewert (1978, p. 894), Aizcorbe and Jackman (1993), many others, and the theoretical results of Diewert (1978, p. 888).[1]

- The two superlative indexes averaged 0.3 of a percentage point per year *below* the corresponding fixed-basket CPI Laspeyres indexes during the period 1984-94. Therefore, an estimate of the substitution bias in the U.S. CPI during this period is 0.3 of a percentage point per year. This result confirms the earlier estimates made by Aizcorbe and Jackman (1993).

- The geometric indexes averaged 0.41 of a percentage point per year *below* the corresponding Laspeyres indexes and hence averaged 0.11 of a percentage point per year *below* the corresponding superlative indexes. These estimates were consistent with Shapiro and Wilcox's a priori assumption that Cobb-Douglas preferences for consumers (consistent with geometric indexes) would overstate the amount of substitution that exists between the 9,000 aggregate commodities distinguished in the Aizcorbe-Jackman data set.

- The disaggregated superlative indexes, defined over 9,000 commodities, were very close to their counterparts that assumed no interarea substitution of commodities between the 44 U.S. regions. These latter indexes were computed as superlative indexes for the 207 commodities within each region, but then these 44 separate superlative indexes were aggregated to the national level using the Laspeyres formula. This is perhaps the authors' most important result. Shapiro and Wilcox (correctly, I believe) interpreted their result as evidence of strong national trends in

[1] Fisher (1922) tabled his numerical results on pages 505 (the Törnqvist price index is his number 123) and 512 (the Fisher ideal price index is his number 353). Diewert (1978, p. 888) showed that the Fisher ideal price index, $P_2(p^0, p^1, x^0, x^1)$ in his notation, numerically approximated the Törnqvist price index, $P_0(p^0, p^1, x^0, x^1)$, to the second order around an equal price (i.e., $p^0 = p^1$) and equal quantity (i.e., $x^0 = x^1$) point.

regional commodity prices. Shapiro and Wilcox also correctly noted that an implication of their result may well be that it is not necessary to have so much regional detail to construct a national price index. This implication requires careful study by statistical agencies.

- A final extremely important result obtained by Shapiro and Wilcox is that the CES index derived by Moulton (1996) using the lagged expenditure weights of two years ago exactly captures the trend rate of growth in the Törnqvist index provided that the elasticity of substitution parameter σ is chosen to be 0.7. Hence this new CES index could be used to predict a superlative index on a monthly basis using data that are presently available to statistical agencies.

Overall, Shapiro and Wilcox have made an extremely useful practical contribution to the literature, and I can find little to criticize in their article. However, I would like to indicate some of the problems involved in constructing a CPI (or any measure of inflation) that Shapiro and Wilcox did not discuss. The first problem is, How should the functional form for the CPI be chosen?

ALTERNATIVE APPROACHES TO CHOOSING AN INDEX NUMBER FORMULA

Shapiro and Wilcox take it for granted that the "correct" method for aggregating strata prices and quantities to form an overall measure of consumer price change is to use a superlative index number formula. However, this is not the only approach suggested to choose the functional form for the CPI. In this section, I shall outline four alternative theoretical approaches used to justify particular functional forms for the price index. For the purposes of this section, I shall assume that price and quantity data, p_n^t and q_n^t, are available for commodities $n = 1, 2, ..., N$ and for time periods $t = 0, 1$. I define the period t price and quantity vectors by $p^t \equiv [p_1^t, \ldots, p_N^t]$ and $q^t \equiv [q_1^t, \ldots, q_N^t]$ respectively for $t = 0, 1$. The purpose of the various approaches is to find a single summary measure of the "average" price change between periods 0 and 1. In other words, the various approaches to the index number problem attempt to determine the functional form for the price index, $P(p^0,p^1,q^0,q^1)$.

The Fixed Basket Approach

The first approach to measuring aggregate price change between periods 0 and 1 dates back several hundred years.[2] This fixed-basket approach took the ratio of the costs of buying the same basket of goods in period 1 to period 0. Two choices are natural for the reference basket—the period 0 commodity vector q^0 and the period 1 commodity vector q^1. These two choices lead to the Laspeyres price index P_L defined as

$$(1) \quad P_L\,(p^0,p^1,q^0,q^1) \equiv p^1 \cdot q^0 / p^0 \cdot q^0$$

and the Paasche price index P_P defined as

$$(2) \quad P_P\,(p^0,p^1,q^0,q^1) \equiv p^1 \cdot q^1 / p^0 \cdot q^1,$$

where $p \cdot q \equiv \Sigma_{n=1}^{N} p_n q_n$ denotes the inner product of the vectors p and q.

The problem with the index number formulas defined by Equations 1 and 2 is that they are equally plausible, but in general they will give different answers. This suggests that we take some sort of evenly weighted average of the two indexes as our final estimate of price change between periods 0 and 1. In other words, define the class of *symmetric fixed-basket price indexes* by

$$(3) \quad P_m\,(p^0,p^1,q^0,q^1) \equiv m[P_L\,(p^0,p^1,q^0,q^1),\, P_p\,(p^0,p^1,q^0,q^1)],$$

where $m(x, y)$ is a homogeneous symmetric mean of the two positive numbers x and y.[3] The simplest choices for the function of two variables, m, are: $m(x, y) \equiv \tfrac{1}{2}x + \tfrac{1}{2}y$ (the arithmetic mean) and $m(x,$

2 The early history of this approach is outlined in Diewert (1993a, pp. 34-36).

3 For the mathematical properties of homogeneous symmetric means, see Diewert (1993b, pp. 361-64).

$y) \equiv x^{1/2} y^{1/2}$ (the geometric mean). With these choices for m, Equation 3 becomes the Sidgwick-Bowley index, P_{SB}, and the Fisher (1922) ideal index, P_F, respectively:

(4) $P_{SB}(p^0,p^1,q^0,q^1)$
$$\equiv {}^1\!/_2(p^1 \cdot q^0/p^0 \cdot q^0) + 1/2(p^1 \cdot q^1/p^0 \cdot q^1);$$

(5) $P_F(p^0,p^1,q^0,q^1)$
$$\equiv [p^1 \cdot q^0 \;\; p^1 \cdot q^1/p^0 \cdot q^0 \;\; p^0 \cdot q^1]^{1/2} \;.$$

Diewert (1978, p. 897) showed that all of the symmetric mean fixed-basket price indexes of the form in Equation 3 numerically approximate each other to the second order around a point where $p^0 = p^1$ and $q^0 = q^1$, so for normal time-series data (where prices and quantities change gradually between the two periods), the price indexes defined by Equations 4 and 5 will be numerically very close and thus it will not matter very much which formula is chosen. However, a strong theoretical reason exists for preferring the Fisher formula in Equation 5 to the Sidgwick-Bowley formula in Equation 4: P_F satisfies the time-reversal test, while P_{SB} does not. The price index, P, satisfies the time-reversal test if the following property is satisfied:

(6) $P(p^1,p^0,q^1,q^0) = 1/P(p^0,p^1,q^0,q^1)$.

In other words, when we reverse the order of time, the new measure of aggregate price change is the reciprocal of the old measure. This is a very desirable property for a price index and in the Appendix, I show that the Fisher ideal index–defined Equation 5 is the *only* member of the class of symmetric fixed-basket indexes that has this property. Therefore, the first approach to the determination of the functional form for the price index leads to the Fisher ideal index defined by Equation 5.

The Test or Axiomatic Approach

If there is only one commodity, then a very reasonable measure of price change going from period 0 to 1 is just the relative price of the single commodity p_1^1/p_1^0. Note that this functional form for the price index (when $N = 1$) satisfies the time-reversal test, Equation 6 above. Note also that p_1^1/p_1^0 is increasing and homogeneous of degree one in p_1^1 and is decreasing and homogeneous of degree minus one in p_1^0. Now let the number of commodities, N, be greater than one. The test approach asks that $P(p^0, p^1, q^0, q^1)$ satisfy mathematical properties analogous to the properties of the single-commodity price index. For example, we can ask that $P(p^0, p^1, q^0, q^1)$ satisfy the time-reversal test, Equation 6, or that $P(p^0, \lambda p^1, q^0, q^1) = \lambda P(p^0, p^1, q^0, q^1)$, where λ is a positive number, or that $P(p^0, p^1, q^0, q^1)$ be increasing in the components of p^1. The test approach to the determination of $P(p^0, p^1, q^0, q^1)$ is an exercise in the theory of functional equations (i.e., assume that P satisfies enough "reasonable" tests or properties so that the functional form for P is determined).

Still at issue: Just what are the "reasonable" tests that an index number formula P should satisfy? Current consensus seems to be that the Fisher ideal price index P_F satisfies more "reasonable" axioms than its competitors.[4] The test approach, therefore, leads to the Fisher ideal index as being the "best" functional form.

The Economic Approach

In the case of a single consumer, the economic approach to the determination of the functional form for the price index works as follows: Assume that a consumer's preferences for the N commodities can be represented by the utility function $f(q_1, ..., q_N) = f(q)$. Define the consumer's expenditure or cost function C, which is dual to f by

(7) $C(u, p) \equiv min_q \{p \cdot q : f(q) \geq u\}$,

where $p \equiv [p_1, \ldots, p_N]$ is a vector of commodity prices that the consumer faces and u is a reference utility level that must be attained. The Konüs (1939) price index between periods 0 and 1 is defined as the

[4] Diewert (1992) showed that P_F satisfied 20 "reasonable" tests.

ratio of the minimum costs of achieving the reference utility level u when the consumer faces the period 0 and 1 prices vectors, p^0 and p^1:

$$(8) \quad P_K(p^0, p^1, u) \equiv C(u, p^1)/C(u, p^0).$$

Konüs (1924, pp. 17-9) showed that when we choose the reference utility level u in Equation 8 to be the base-period utility level $u^0 \equiv f(q^0)$ and the period 1 utility level $u^1 \equiv f(q^1)$, we obtain the following observable bounds on the two theoretical Konüs indexes P_K :

$$(9) \quad P_K(p^0, p^1, u^0) \le P_L(p^0, p^1, q^0, q^1);$$

$$(10) \quad P_K(p^0, p^1, u^1) \ge P_P(p^0, p^1, q^0, q^1),$$

where P_L and P_P are the observable Laspeyres and Paasche indexes defined earlier by Equations 1 and 2. The bounds of Equations 9 and 10 on the unobserved Konüs price indexes $P_K(p^0, p^1, u^0)$ and $P_K(p^0, p^1, u^1)$ are completely nonparametric. That is, they are valid no matter what the functional form for the utility function f (or its dual cost function C) is.

To further progress with the economic approach, economists—starting with Konüs and Byushgens (1926)—have assumed specific functional forms for f or C and then deduced that certain theoretical Konüs indexes are *exactly* equal to specific index number formulas. For example, Diewert (1976, p. 122) assumed that the consumer's preferences could be represented by a general translog cost function. He then showed that if the reference utility level for the Konüs index were chosen to be $u^* \equiv [u^0 u^1]^{1/2}$, the geometric mean of the period 0 and 1 utility levels, then

$$(11) \quad P_K(p^0, p^1, u^*) = P_T(p^0, p^1, q^0, q^1),$$

where P_T is the Törnqvist price index. P_T is defined by

$$(12) \quad \ln P_T(p^0, p^1, q^0, q^1) \equiv \Sigma_{n=1}^{N}(1/2)(s_n^0 + s_n^1) \ln(p_n^1 / p_n^0),$$

where the period t expenditure share for commodity n is defined by $s_n^t \equiv p_n^t q_n^t / p^t \cdot q^t$ for $n = 1, ..., N$ and $t = 0, 1.$[5] Diewert (1976) called P_T a superlative index number formula because it is exactly equal to a theoretical Konüs price index of the form $C(u^*, p^1)/C(u^*, p^0)$, where the functional form for C is *flexible*.

Diewert (1976, pp. 129-32) provided many other examples of superlative index number formulas, including the Fisher ideal price index, P_F, defined by Equation 5. It turns out that P_F is exactly equal to $P_K(p^0, p^1, u)$ for any positive reference utility level u, provided that

$$f(q) \equiv [q^T A q]^{1/2} \text{ or } C(u, p) \equiv u\,[p^T B p]^{1/2},$$

where A and B are symmetric N by N matrixes that satisfy certain properties.[6]

There are at least two ways that the economic approach to the determination of the functional form for the price index can be extended from the one-consumer or household case to the many-household case: Assume that each household has the same (homothetic) preferences or apply the theory of bounds to a generalized version of the single consumer Konüs price index defined by Equation 8. In both these approaches, it is assumed that each household faces the same vector of prices in each period.

Using the first approach, the Fisher price index (defined by Equation 5) and the Törnqvist price index P_T (defined by Equation 12) again emerge as superlative functional forms.[7]

The second approach is more interesting. To develop the bounds approach in the context of many households, we need a many-household counterpart to the single-consumer Konüs price index P_K defined earlier by Equation 8. The counterpart I choose is a generalization of Pollak's (1981, p. 328) Scitovsky-Laspeyres price index. Let there be H households in the economy and represent the preferences of household h by the expenditure function $C^h(u_h, p)$ for $h = 1, ..., H$. Diewert (1983, p. 190) defined the *Pollak Plutocratic cost of living index* for period 0 prices $p^0 \equiv [p_1^0, ..., p_N^0]$, period 1

[5] Equality 11 does not require the assumption of homothetic preferences but it does require a particular choice for the reference utility level u^*.

[6] Preferences are homothetic for both these functional forms.

[7] The vectors q^0 and q^1 are now interpreted as total market-demand vectors in periods 0 and 1, respectively. I assume that all H households are present in both periods.

prices $p^1 \equiv [p_1^1, ..., p_N^1]$, and reference household utility levels $u \equiv [u_1, ..., u_H]$ by

$$(13)\quad P_{PP}(p^0, p^1, u) \equiv \Sigma_{h=1}^{H} C^h(u_h, p^1) / \Sigma_{h=1}^{H} C^h(u_h, p^0).$$

Note that I am assuming each household faces the same price vectors (p^0 in period 0 and p^1 in period 1) and that $P_{PP}(p^0, p^1, u)$ is the ratio of the market costs of achieving the same household utility levels $u \equiv [u_1, ..., u_H]$ for the two periods. In a straightforward extension of the bounds in Equations 9 and 10, Diewert (1983, p. 191) showed that when we choose the reference utility vector u in Equation 13 to be $u^0 \equiv [u_1^0, ..., u_H^0]$(the base-period utility vector) and $u^1 \equiv [u_1^1, ..., u_H^1]$ (the period 1 utility vector), we then obtain the following bounds for the plutocratic price index:

$$(14)\quad P_{PP}(p^0, p^1, u^0) \le P_L(p^0, p^1, q^0, q^1) \equiv p^1 \cdot q^0 / p^0 \cdot q^0;$$

$$(15)\quad P_{PP}(p^0, p^1, u^1) \ge P_P(p^0, p^1, q^0, q^1) \equiv p^1 \cdot q^1 / p^0 \cdot q^1,$$

where $q^0 \equiv \Sigma_{h=1}^{H} q^{h0}$ and $q^0 \equiv \Sigma_{h=1}^{H} q^{h1}$ are the period 0 and 1 observed market demand vectors and P_L and P_P are the aggregate Laspeyres and Paasche price indexes. Diewert (1983, p. 191) also showed that a reference household utility vector $u^* \equiv [u_1^*, ..., u_H^*]$ exists such that each u_h^* lies between the household h period 0 and 1 utility levels, u_h^0 and u_h^1 respectively, for $h = 1, ..., H$ and

$$(16)\quad min\{P_L, P_P\} \le P_{PP}(p^0, p^1, u^*) \le max\{P_L, P_P\}.$$

Thus, when evaluated at the intermediate utility levels u^*, the theoretical Pollak plutocratic price index $P_{PP}(p^0, p^1, u^*)$ lies between the observable Paasche and Laspeyres aggregate price indexes, P_P and P_L.[8]

As was the case with the Fixed Basket Approach earlier, it is useful to obtain a single point estimate for the theoretical index $P_K(p^0, p^1, u^*)$ by taking a symmetric average of its bounds (P_L and P_P). If we want our final estimate to satisfy the time-reversal test, we can again apply Proposition 1 in the Appendix and conclude that our "best" functional form from the viewpoint of the bounds approach is the Fisher ideal price index P_F.

Thus from the viewpoint of economic approaches to choosing the functional form for the price index, it appears that either the Törnqvist index P_T (defined by Equation 12) or the Fisher ideal price index P_F (defined by Equation 5) are "best" choices.

The Stochastic Approach

The stochastic approach to the determination of the price index can be traced back to the work of Jevons and Edgeworth more than 100 years ago.[9]

The basic idea behind the stochastic approach is that each price relative p_n^1/p_n^0 for $n = 1, ..., N$ can be regarded as an estimate of a common inflation rate α between periods 0 and 1. In other words, it is assumed that

$$(17)\quad p_n^1/p_n^0 = \alpha + \varepsilon_n$$

where $n = 1, ..., N$, α is the common inflation rate, and the ε_n are random variables with mean 0 and variance σ^2. The least-squares or maximum-likelihood estimator for α is the Carli price index P_C defined as

$$(18)\quad P_C(p^0, p^1) \equiv \Sigma_{n=1}^{N} (1/N)(p_n^1/p_n^0)$$

Unfortunately, P_C does not satisfy the time-reversal test, $P(p^1, p^0) = 1/P(p^0, p^1)$.[10]

Let us change our stochastic specification as follows: Assume the logarithm of each price relative, $\ln(p_n^1/p_n^0)$ is an unbiased estimate for the logarithm of the inflation rate between periods 0 and 1:

$$(19)\quad \ln(p_n^1/p_n^0) = \beta + \varepsilon_n,$$

where $n = 1,...,N$, $\beta \equiv \ln\alpha$, and the ε_n are random variables with mean 0 and variance σ^2. The least-squares or maximum-likelihood estimate for β is the logarithm of the geometric mean of the price relatives. Hence the corresponding estimate for

[8] This result does not require the assumption of homothetic preferences.

[9] See Diewert (1993a, pp. 37-8; 1995a) for references to the early literature.

[10] In fact, $P_C(p^0, p^1)\,P_C(p^1, p^0) \ge 1$, with a strict inequality unless $p^1 = \lambda p^0$. Fisher (1922, p. 66) seems to have been the first to point out this upward bias of the Carli index and he urged (to no avail) statistical agencies not to use this index number formula.

the common inflation rate α is the Jevons price index $P_J(p^0, p^1)$, defined as the geometric mean of the price relatives:

$$(20)\quad P_J(p^0, p^1) \equiv \Pi_{n=1}^{N} (p_n^1/p_n^0)^{1/N}$$

The Jevons price index P_J satisfies the time-reversal test and hence is much more satisfactory than the Carli index P_C.

However, both the Jevons and Carli price indexes suffer from a fatal flaw: Each price relative p_n^1/p_n^0 is regarded as being equally important and is given an equal weight in the index number formulas in Equations 18 and 20.

Theil (1967, pp. 136-37) proposed a solution to the lack of weighting in Equation 20. He argued as follows: Suppose we draw price relatives at random in such a way that each dollar of expenditure in the base period has an equal chance of being selected. Then the probability that we will draw the nth price relative is equal to $s_n^0 \equiv p_n^0 q_n^0 / p^0 \cdot q^0$, the period 0 expenditure share for commodity n. Then the overall mean (period 0 weighted) logarithmic price change is $\Sigma_{n=1}^{N} s_n^0 \ln(p_n^1/p_n^0)$. Now repeat the above mental experiment and draw price relatives at random in such a way that each dollar of expenditure in period 1 has an equal probability of being selected. This leads to the overall mean (period 1 weighted) logarithmic price change of $\Sigma_{n=1}^{N} s_n^1 \ln(p_n^1/p_n^0)$. Each of these measures of overall logarithmic price change seems equally valid. Therefore, as usual, we could argue for taking a symmetric average of the two measures in order to obtain a final single measure of overall logarithmic price change.[11] Theil (1967, p. 137) argued that a nice symmetric index number formula can be obtained if we make the probability of selection for the nth price relative equal to the arithmetic average of the period 0 and 1 expenditure shares for commodity n, $(1/2) s_n^0 + (1/2) s_n^1$. Using these probabilities of selection, Theil's final measure of overall logarithmic price change was

$$(21)\quad \ln P_T(p^0, p^1, q^0, q^1) = \Sigma_{n=1}^{N} (1/2)(s_n^0 + s_n^1)\ln(p_n^1/p_n^0).$$

Taking antilogs of both sides of Equation 21, we obtain our old friend P_T, the Törnqvist-Theil price index defined earlier in Equation 12. This index number formula appears to be "best" from the viewpoint of Theil's stochastic approach to index number theory.

Additional material on stochastic approaches to index number theory and references to the literature can be found in Diewert (1995a) and Mark A. Wynne (pp. 157-163 in this issue).

Summarizing the results of this review of the four alternative approaches to the determination of the index number formula: All four approaches lead to the choice of either the Fisher ideal formula P_F (defined by Equation 5) or the Törnqvist-Theil formula P_T (defined by Equations 12 or 21) as being "best." From an empirical point of view, it will not matter very much whether P_F or P_T is chosen because the two indexes approximate each other to the second order around an equal price and quantity point.[12]

I will now discuss some other problems associated with the measurement of consumer price change.

THE PROBLEM OF SEASONALITY

The analysis in the previous section assumed that all N commodities were available in periods 0 and 1 and that price and quantity information on these N commodities could in principle be collected. However, many commodities are *seasonal* (i.e., they are available in certain seasons but not in other seasons). The availability of some commodities in some seasons makes short-term price comparisons for those seasonal commodities impossible: We cannot compare the incomparable!

The existence of seasonal commodities means that traditional index number theory (as outlined in the previous section) is simply not applicable to the entire universe of goods. To apply traditional index number theory when there are seasonal goods, we will have to restrict the index number comparisons to subsets of nonseasonal

[11] The arithmetic mean works best in this context.

[12] See Diewert (1978), p. 888.

goods. The difficulties posed by the existence of seasonal goods for the measurement of short-run price change have not been adequately recognized by both index number theorists and statistical agencies.[13]

PROBLEMS AT THE LOWEST LEVEL OF AGGREGATION

Statistical agencies around the world find that 2 percent to 4 percent of their price quotations on individual commodities disappear each month. Some of these disappearances are a result of the existence of seasonal goods, but many disappearances are a result of existing commodities' replacement by "new" commodities. Just as the existence of seasonal commodities makes it impossible to apply traditional index number theory, the existence of new commodities leads to difficulties. However, in theory, traditional index number theory can be adapted to deal with the appearance of new commodities. Suppose a new good made its first appearance in period 1. Then, in theory, we could find a reservation price for the new commodity that would induce the consumer to demand zero units of the new good in period 0.[14] This period 0 reservation price could be used in the index number formula as the period 0 price for the new good. Diewert (1980, pp. 501-03) suggested an econometric approach to the estimation of these reservation prices and Hausman (forthcoming) actually implemented an econometric approach. However, these econometric approaches are very data and labor intensive and hence are not a practical solution to the new goods problem for statistical agencies that have to deal with thousands of new commodities each month.

Many economists (myself included) believe the magnitude of the new goods problem is increasing.[15] The increasing proliferation of new products on the market creates severe problems for statistical agencies in their attempts to measure general price change: They cannot compare the price of a new product in this period to its price in the base period if the product did not exist in the base period!

Statistical agencies face a related problem when they attempt to produce measures of business output growth and price change: 10 percent to 15 percent of all firms disappear each year and are replaced by new firms. This *new firms problem* again leads to a lack of compar-ability in the price and quantity statistics of one period compared with an earlier period.

Another problem that occurs at the lowest level of aggregation has to do with the very definition of prices and quantities inserted into an index number formula. In the CPI context, prices are usually list prices taken from various outlets during the reference period. However, *sales* of a particular commodity at a particular outlet will typically *not* take place at this list price. Often variations will arise in transaction prices for the same commodity at the same outlet during the reference period. Bilateral index number theory requires that this within-the-period variation in prices and quantities be aggregated into single-price and quantity numbers that will be inputs into the index number formula. How should this micro aggregation be accomplished? Hill (1993, p. 399) recommended the use of the commodities' unit value and the total quantity transacted during the period as the theoretically correct price and quantity.[16] With the growth of scanner data and other electronically stored databases, it is now feasible for statistical agencies to switch from the traditional sampling of list prices methodology to a unit-value methodology.

WHAT IS THE "CORRECT" DOMAIN OF DEFINITION FOR THE PRICE INDEX?

Up to this point, I have not questioned the appropriateness of assuming that the usual domain of definition for the CPI is appropriate for measuring short-run inflation in an economy.[17] However, articles in this issue by William A. Allen (p. 173), Stephen G. Cecchetti (p. 139) and Mark A. Wynne (p. 157) assert that the appropriateness of the unadjusted CPI as "the"

[13] For discussions of these difficulties and references to the literature, see Hill (1995) and Diewert (1996b).

[14] This idea was first suggested by Hicks (1940, p. 114).

[15] See Diewert (1996c, pp. 33-4).

[16] For more extensive discussions of these issues and references to the literature, see Diewert (1995b, pp. 20-4, and 1996a, pp. 261-62).

[17] The CPI domain of definition (i.e., the set of transactions it includes) is essentially all household purchases of consumer goods and services in the economy during a specified period, with some adjustments for housing.

measure of short-run inflation is very much open to question. Therefore, in this section I will review nine classes of adjustments to the usual CPI that have been suggested in the literature. These adjustments are made to create a "better" measure of short-run inflation.

Exclusion of Seasonal Goods

It seems obvious that goods available in one season but not in another should be excluded from the domain of definition of a short-run inflation index. A case can also be made for excluding other goods that might be available throughout the year but are nevertheless subject to strong seasonal price fluctuations (e.g., fresh fruits and vegetables). For a more comprehensive discussion of index number problems in the context of seasonal goods, see Diewert (1996b).

Exclusion of Durable Goods

If a consumer purchases a durable good during the CPI reference period, then he or she receives a benefit from the use of the durable for more than one period. Hence, it seems incorrect to attribute the full cost of the durable to the period of purchase. On the other hand, spreading out the cost of the purchase over the useful life of the durable good leads us into the complexities of constructing user costs for the durable. For example, what is the correct opportunity cost of capital, what is the correct depreciation rate, and should capital gains be included?[18] The difficulties involved in constructing "objective" user costs for durable goods have led many economists to advocate the removal of at least some classes of durable good (such as housing) from the domain of definition of the CPI to make it a more accurate measure of short-run inflation. In this issue, Allen showed that in the United Kingdom, quite different measures of inflation are obtained, depending on how housing is treated in the CPI.

Inclusion of Future Goods or Savings

Because existing goods can be substituted for future goods, it seems theoretically appropriate to include discounted future (expected) prices of consumption goods in future periods in a more comprehensive index of consumer prices. Allen discusses in more detail this suggestion for broadening the domain of definition of the CPI (and the associated measurement difficulties).

Inclusion of Leisure

Again, from a theoretical point of view, it seems appropriate to include leisure as a commodity in the domain of definition of the CPI, because leisure can be substituted for increased consumption of goods. As is the case with future goods or savings, practical measurement difficulties have prevented statistical agencies from including leisure in the CPI.

Exclusion of Commodity Taxes

If the Government suddenly increases commodity taxes, then it is likely the CPI will also have a sudden upward movement. Many economists have therefore suggested that commodity taxes should be removed (somehow) from the CPI and that the resulting "tax-free" CPI should be used to monitor short-run inflation.[19] The problem with this proposal is that there is no unambiguous, completely accurate method for removing all indirect commodity taxes. In other words, any attempt to do this will be a complex exercise in applied general-equilibrium modelling rather than in economic measurement. Moreover, the fact that the government has caused consumer prices to increase rather than some other economic phenomenon seems somewhat immaterial: In either case, households are facing higher prices, and we may want to measure this fact!

[18] See Diewert (1980, pp. 475-86) for a discussion of these and other problems associated with the construction of user costs.

[19] Allen discusses this proposal.

Exclusion of Food and Energy

Because food and energy prices supposedly fluctuate more violently than the prices of other commodities, it is often advocated that these components of the CPI be excluded to obtain a more accurate measure of short-run inflation. This proposal causes a number of problems: (a) It is not certain that food and energy prices fluctuate more than other prices.[20] (b) Why remove food and energy prices but not other highly fluctuating components? (c) If food and energy price relatives have different long-run trends compared with the remaining CPI components, then their removal from the CPI will lead to a systematic bias in the resulting CPI *ex* food and energy. Cecchetti and Wynne discuss these problems further.

Exclusion of Volatile Commodities

Rather than simply eliminate all price quotations that pertain to volatile components of the CPI, Cecchetti proposes the elimination of individual volatile price quotations by suggesting the use of various limited-influence estimators such as the weighted median or the 10 percent trimmed mean. An implicit assumption in proposals of this type is that all price relatives have the same mean (but some have higher variances), and hence we can still unbiasedly estimate the mean of the price distribution by dropping the higher variance-price relatives from our average.

Many years ago, Keynes (1930, p. 78) criticized the assumption that all price relatives have the same mean by noting that the interaction between commodity suppliers and demanders will usually lead to *systematic* variations in relative prices. For example, Keynes pointed out that usually wages rise faster than consumer prices. Similarly, land prices tend to rise more quickly than wage rates over long periods, and wage rates tend to rise faster than capital input prices. It is simply not legitimate therefore to assume that *all* price relatives have the same mean and hence the legitimacy of limited influence estimators of inflation are thrown into doubt. These systematic differences in the rate of change of prices for many commodity classes mean that *the overall level of price change cannot be defined independently of the domain of definition of the price index*. Thus it seems to me all stochastic approaches to the measurement of price change that do not specify both a specific domain of definition for the price index and a Theil-type economic weighting scheme are difficult to interpret at best and at worst are biased measures of price change. Note that this criticism does not apply to Theil's stochastic approach discussed earlier. He weighted each price relative according to its economic importance and, moreover, Theil did not assume each price relative had the same mean. This neo-Keynsian criticism of stochastic approaches to the measurement of price change also applies to the neo-Edgeworthian model proposed by Diewert (1995a) and implemented by Wynne in this issue. The weights assigned to price relatives in this approach are determined by the variances of individual price relatives and not by their economic importance. For references to the literature and critiques of the new stochastic approaches to the measurement of inflation, see Diewert (1995a) and Wynne in this issue.

Consumption vs. Output vs. Transactions

Instead of taking all consumption purchases in the economy during a specified period as the appropriate domain of definition for the price index or inflation measure, we could take other subsets of transactions as the appropriate domain of definition for the price index. For example, we could take the value of gross outputs (at producer prices) less the value of intermediate input purchases for private producers as our reference set of transactions and the resulting price index would be the (private sector) *output price index*. Or we could take the value of all primary input purchases by private producers as our domain of definition and the resulting price index would be

[20] See Cecchetti, pp. 144-46 of this issue.

the (private sector) *input price index*. Either of these price indexes could be used as a measure of short-run inflation. Note, however, that the input price index will typically grow more quickly than the output price index if there is technical progress or productivity improvement in the economy.

This difference in the two indexes illustrates that systematic trends in prices occur and hence the choice of the domain of definition of the price index matters.[21] Wynne discussed another possible domain of definition for the price index: namely, Irving Fisher's (1911) set of all monetary *transactions* that take place in the economy during a reference period. The problem with this domain of definition is that sales and purchases of financial assets (currency trades in particular) would receive overwhelming weight in this domain of definition, and the resulting price index would not be representative of the economy's real transactions.

Choice of the Reference Period

Up to now, I have not questioned the choice of a month as the appropriate length of time to make price comparisons. Following Fisher (1922, p. 318) and Hicks (1946, p. 122), the reference period should be short enough so that variations in the prices of commodities within the period can be ignored. In other words, the period should be short enough so that inflation within the period can be neglected. To be certain we have minimized price variation within the period, why not choose the reference period to be a week or even a day, rather than a month? Statistical agencies do not choose a one-day reference period for the CPI for many reasons: (a) It would be too costly to collect and publish information on a daily basis. (b) Many goods would not be transacted on a daily basis, making it difficult to form day-to-day price comparisons on transactions. (c) Such a short-term index would probably be dominated by statistical noise. In other words, it would be difficult to separate the daily trend in prices from seasonal trends and random noise.[22]

To eliminate these noise and seasonal components, Cecchetti uses a 36-month moving average of monthly CPI inflation rates to define his "core" inflation rate. This means that policymakers have to wait 18 months to be able to determine the "true" monthly inflation rate![23] This 18-month wait to know the "truth" can probably be shortened to about 6 months.[24] However, in this issue, Cecchetti asserts that the existing theory of the CPI does not deal adequately with the seasonality or noise problems. In other words, existing theory does not recognize that monthly CPI estimates are so noisy that they must be intertemporally smoothed to obtain meaningful estimates of trend inflation. Thus there is a need to extend existing bilateral index number theory to a multilateral, many-period framework.

Obviously, the problems involved in measuring general price change are far from being definitively resolved.

CONCLUSION

My main conclusions are the following:

- The four leading approaches to the problem of making price comparisons between two periods all lead to the choice of either the Fisher ideal index P_F (defined by Equation 5) or the Törnqvist-Theil index P_T (defined by Equation 12) as the "best" functional forms for making bilateral price comparisons.

- The disappearance of old commodities and existing firms, the appearance of new commodities and firms and the existence of seasonal commodities mean that traditional bilateral index number theory is not applicable without further modifications and that existing statistical agency index numbers are not necessarily reliable.

- It is important to determine what should be the "correct" domain of definition for the price index that central

[21] Hill (1995) and Diewert (1996b, pp. 29-31) considered the advantages and disadvantages of different domains of definition for a short-term inflation index.

[22] In fact, in this issue, Cecchetti shows that even the U.S. monthly CPI has large seasonal and noise components.

[23] If policymakers use U.S. seasonally adjusted data to make their decisions, then they will have to wait approximately 18 months before the seasonal-adjustment procedure yields stable estimates for a particular month.

[24] Diewert (1996b) suggested an index number method for constructing trend inflation rates free of seasonal influences that requires a wait of only 6 months. His method is based on the earlier work of Mudgett (1955, p. 97) and Stone (1956, pp. 74-5).

bankers should target as an indicator of general inflation. Although index number and price measurement theorists seem to have therefore resolved some of the problems involved in measuring general inflation, a large number of problems remain unresolved.

REFERENCES

Aizcorbe, Ana M., and Patrick C. Jackman. "The Commodity Substitution Effect in CPI Data, 1982-91: Anatomy of Price Change," *Monthly Labor Review* (December 1993), pp. 25-33.

Diewert, W.E. "Exact and Superlative Index Numbers," *Journal of Econometrics* (May 1976), pp. 115-45.

_______. "Superlative Index Numbers and Consistency in Aggregation," *Econometrica* (July 1978), pp. 883-900.

_______. "Aggregation Problems in the Measurement of Capital," *The Measurement of Capital*, Dan Usher, ed., University of Chicago Press, 1980, pp. 433-528.

_______. *Price Level Measurement: Proceeding From a Conference Sponsored by Statistics Canada*, Minister of Supply and Services Canada, 1983.

_______. "Fisher Ideal Output, Input and Productivity Indexes Revisited," *The Journal of Productivity Analysis* (September 1992), pp. 211-48.

_______. "The Early History of Price Index Research," *Essays in Index Number Theory*, Vol. 1, W.E. Diewert and A.O. Nakamura, eds., Amsterdam: North-Holland Publishing Co., 1993a, pp. 33-65.

_______. "Symmetric Means and Choice Under Uncertainty," *Essays in Index Number Theory*, Vol. 1, W.E. Diewert and A.O. Nakamura, eds., Amsterdam: North-Holland Publishing Co., 1993b, pp. 355-433.

_______. "On the Stochastic Approach to Index Numbers," Discussion Paper 95-31, Department of Economics, University of British Columbia, 1995a.

_______. "Axiomatic and Economic Approaches to Elementary Price Indexes," NBER working paper no. 5104 (May 1995b).

_______. "Price and Volume Measures in the System of National Accounts," *The New System of National Economic Accounts*, John Kendrick, ed., Kluwer Academic Publishers, 1996a, pp. 237-85.

_______. "Seasonal Commodities, High Inflation and Index Number Theory," Discussion Paper 96-06, Department of Economics, University of British Columbia, 1996b.

_______. "Comment on CPI Biases," *Business Economics* (April 1996c), pp. 30-35.

Fisher, Irving. *The Purchasing Power of Money; Its Determination and Relation to Credit, Interest and Crises*, London: Macmillan, 1911.

_______. *The Making of Index Numbers; A Study of their Varieties, Tests, and Reliability*, Houghton Mifflin, 1922.

Hausman J. "Valuation of New Goods Under Perfect and Imperfect Competition," *New Goods*, T. Bresnahan and R.J. Gordon, eds., University of Chicago Press, forthcoming.

Hicks, John R. "The Valuation of the Social Income," *Economica* (May 1940), pp. 105-24.

_______. *Value and Capital; An Inquiry into Some Fundamental Principles of Economic Theory*, 2nd ed., Oxford: Clarendon Press, 1946.

Hill, P. "Price and Volume Measures," *System of National Accounts 1993*, Eurostat, IMF, OECD, U.N., and the World Bank, 1993, pp. 379-406.

_______. *Inflation Accounting: A Manual on National Accounting under Conditions of High Inflation*, OECD, Paris, 1995.

Keynes, John Maynard. A Treatise on Money, Vol. 1, Harcourt, Brace, and Company, 1930.

Konüs, A.A. "The Problem of the True Index of the Cost of Living," *Econometrica* (January 1939), pp. 10-29.

_______ and S.S. Byushgens. "K probleme pokupatelnoi cili deneg," *Voprosi Konyunkturi* (vol. 11, no. 1, 1926), pp. 151-72.

Mudgett, Bruce D. "The Measurement of Seasonal Movements in Price and Quantity Indexes," *Journal of the American Statistical Association* (March 1955), pp. 93-8.

Pollak, Robert A. "The Social Cost of Living Index," *Journal of Public Economics* (June 1981), pp. 311-36.

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

Theil, Henri. *Economics and Information Theory*, Amsterdam: North-Holland Publishing Co., 1967.

Appendix

An Optimality Property for the Fisher Ideal Index

As noted in Diewert (1993b, pp. 361-64), a homogeneous symmetric mean $m(a, b)$ defined for all positive numbers a and b is usually assumed to satisfy the following properties: $m(a, b)$ is a symmetric [*i.e.*, $m(a, b)=m(b, a)$[, strictly increasing, continuous, and (positively) homogeneous of degree one function of two variables that has the mean-value property [*i.e.*, $m(a, a)=a$]. Diewert (1993b, p. 362) shows that, with these properties, $m(a, b)$ is positive if a and b are positive. For the proof of Proposition 1 below, m must satisfy only two of the above properties:

(A1) $m(a, b) > 0$ for $a > 0$, $b > 0$ (positively);

(A2) $m(\lambda a, \lambda b) = \lambda m(a, b)$ for $\lambda > 0$, $a > 0$, $b > 0$ (linear homogeneity).

Proposition 1: Let the symmetric mean fixed-basket price index $P_m(p^0, p^1, q^0, q^1)$ be defined by Equation 3 in the text, where the mean function satisfies Equations A1 and A2. If, in addition, P_m satisfies the time-reversal test (b), then P_m must equal the Fisher ideal index P_F (defined by Equation 5).

Proof: Substituting Equations 1, 2, and 3 into the time-reversal test (Equation 6), m must satisfy the following functional equation:

$$m(p^0 \cdot q^1 / p^1 \cdot q^1, p^0 \cdot q^0 / p^1 \cdot q^0)$$
$$= 1/m(p^1 \cdot q^0/p^0 \cdot q^0, p^1 \cdot q^1/p^0 \cdot q^1)$$

or

$$m(b^{-1}, a^{-1}) = 1/m(a, b), \tag{A3}$$

where a and b are defined as the Laspeyres and Paasche price indexes:

$$a \equiv p^1 \cdot q^0/ p^0 \cdot q^0 = P_L; \quad b \equiv p^1 \cdot q^1/ p^0 \cdot q^1 = P_P. \tag{A4}$$

Hence using Equations A1 and A3, m must satisfy:

$$\begin{aligned} 1 &= m(a, b)m(b^{-1}, a^{-1}) \\ &= am(1, b/a)a^{-1}m(a/b, 1), \text{ using Equation A2} \\ &= m(1, x)m(x^{-1}, 1), \text{ defining } x \equiv b/a \\ &= m(1, x)x^{-1}m(1, x^{-1}), \text{ using Equation A2.} \end{aligned} \tag{A5}$$

Equation A5 can be rewritten as

$$[m(1, x)]^2 = x \text{ or, using Equation A1}$$
$$m(1, x) = x^{1/2}. \tag{A6}$$

Using Equation A2 again, we have

$$\begin{aligned} m(a, b) &= am(1, b/a) \\ &= a(b/a)^{1/2} \text{ using Equation A6} \\ &= a^{1/2}b^{1/2}. \end{aligned} \tag{A7}$$

Substitution of Equation A7 into Equation 3 shows that $P_m = P_F$.