Volume Title: The Economics of New Goods

Volume Author/Editor: Timothy F. Bresnahan and Robert J. Gordon, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-07415-3

Volume URL: http://www.nber.org/books/bres96-1

Publication Date: January 1996

Chapter Title: The Construction of Basic Components of Cost-of-Living Indexes

Chapter Author: Marshall B. Reinsdorf, Brent R. Moulton

Chapter URL: http://www.nber.org/chapters/c6073

Chapter pages in book: (p. 397 - 436)

10 The Construction of Basic Components of Cost-of-Living Indexes

Marshall B. Reinsdorf and Brent R. Moulton

New products pose a difficult problem in the construction of price indexes. In 1887, in one of the earliest discussions of new products and price-change measurement, Alfred Marshall suggested the use of chaining to incorporate new products into price indexes. More than a hundred years later, this solution remains the most common method of handling new products.

The possibility that chaining could lead to bias was also understood quite early. Although Pigou argued that chaining is necessary to keep the index from becoming unrepresentative due to the exclusion of new products, he warned: "It must, indeed, be conceded that, if the successive individual comparisons embodied in the chain method, each of which admittedly suffers from a small error, are for the most part to suffer from errors *in the same direction,* the cumulative error as between distant years may be large. Were people equally likely to forget how to make things now in use as to invent new things, a large cumulative error would be unlikely. But, in fact, we know that the great march of inventive progress is not offset in this way. Hence the errors introduced by the chain method are likely to be predominantly in one direction" (1932, 71–72).

The basic component indexes investigated in this paper aggregate prices for a single good or a narrow category of goods. New manufacturers, outlets, and product designs pose a problem in these indexes parallel to that posed by new goods at a higher level of aggregation. Indeed, flux in the population of outlets and varieties in which goods are offered is frequently greater than the flux in the population of consumer goods. Furthermore, since new products are often

Marshall B. Reinsdorf and Brent R. Moulton are economists at the Bureau of Labor Statistics.

The views expressed in this paper are those of the authors and do not represent views of the Bureau of Labor Statistics. The authors are grateful to Marilyn Manser, Jack Triplett, and Andy Baldwin for helpful comments.

similar in function to products that already exist, they generally enter the U.S. Consumer Price Index (CPI) by being chained into an existing component index. (For example, compact-disc players entered the CPI by being chained into one of the audio component indexes, where they may have partly replaced analog phonographs.) Reinsdorf (1993) discusses how chaining basic component indexes may prevent the CPI from crediting progress embodied in new outlets, new brands, or new varieties of goods.

The present paper examines a bias of a different type. It arises when a statistical estimator of a Laspeyres index is used for a basic component of a chained index. If competition rules out large discrepancies between the prices offered in a market, all the sellers who continuously participate in a market will tend to have similar long-run average rates of price change. Absent turnover in the population of sellers, convergence of all rates of price change to a common trend value (in effect, fulfilling the conditions for a Hicksian composite commodity) would tend to make the bias negligible in any long-run component index that passes the proportionality test.[1] In contrast, we show below that when sample estimates of short-run Laspeyres component indexes are chained to measure long-run price change, trend reversion of sellers' prices exacerbates, rather than eliminates, the bias. Prices that start out low tend to have excessive weights in these indexes, but trend reversion means that these initially low prices tend subsequently to have high growth rates. This makes prices' weights positively correlated with their rates of change, resulting in upward bias.

In exploring this effect theoretically, it is convenient to adopt the abstraction that the economy contains just two levels of aggregation, goods and sellers of goods. Goods make up a consumer's market basket, while sellers make up the supply side of a market in which a good is sold. In practice, of course, a single good may be offered by many manufacturers in many versions at many outlets. Hence, depending on the context, our "sellers" paradigm may refer to either the competing outlets that stock a good or the competing varieties of a good in a market. A tendency to follow a common price trend may be expected across outlets and varieties. This tendency arises because consumers may be apt to substitute between outlets or between varieties and, also, because the sellers in a market may use similar inputs and technologies, which would lead to common cost trends.

An overview of the body of this paper is as follows. In section 10.1 we introduce the Laspeyres approach to the construction of the basic component indexes for a good. In section 10.2 we describe the way in which this approach is implemented in practice in the U.S. CPI. In section 10.3 we discuss how consumers' tendency to substitute products whose relative price has fallen for

1. The proportionality test requires that the index comparing price vector $\boldsymbol{p}_t$ to price vector $\boldsymbol{p}_0$ equal λ whenever $\boldsymbol{p}_t = \lambda \boldsymbol{p}_0$.

products whose relative price has risen can cause upward bias in true Laspeyres component indexes. In section 10.4 we explain why chaining CPI or true Laspeyres component indexes may exacerbate rather than ameliorate their upward bias. In section 10.5 we derive simple formulas for the bias of several types of component indexes using a model of price and consumer behavior. In section 10.6 we discuss a possible solution to the bias problem, the use of geometric-mean component indexes. In section 10.7 we provide empirical evidence on how geometric mean indexes would affect the CPI. Finally, in section 10.8 we discuss additional empirical evidence on the performance of CPI component indexes.

10.1 Laspeyres Component Indexes

Laspeyres price indexes measure the comparative cost of purchasing, in a later time period, the basket of items that consumers originally purchased in a "base" period. They are widely used (at least at higher levels of aggregation) because they can be calculated with greater timeliness and at a lower cost than price indexes that require both current- and base-period quantity data.

Another advantage of Laspeyres indexes is that they have the property of consistency in aggregation. This means that repeated application of the Laspeyres formula at any number of stages of aggregation yields the same result as combining the multitude of prices into a comprehensive index number in a single step. In particular, a Laspeyres price index may be constructed in one step as a weighted average of the ratio of each seller's comparison- and base-period prices, where the weights are the sellers' shares of aggregate expenditure in the base period. It may also be constructed in two steps by calculating Laspeyres component indexes combining sellers' prices for each good, and then finding the weighted average of goods-level indexes, with each good's weight proportional to the amount that consumers spent on it in the base period. The Laspeyres component indexes may themselves be constructed as weighted averages of the ratios of sellers' comparison- and base-period prices for the particular good in question, where the weights are proportional to consumers' base-period expenditures on that good at each of its sellers. Sample estimates of such indexes of this type are the basic building blocks of the U.S. CPI.

To see explicitly how an aggregate-level Laspeyres index may be constructed from Laspeyres indexes for individual goods, let G denote the market basket of goods, let S_g denote the universe of sellers of good $g \in G$, and let K denote the population of consumers. Also, let p_{gst} be the price for good g from seller $s \in S_g$ at time t, and let q_{gskt} denote the quantity that consumer $k \in K$ buys of good g from seller s at time t. Finally, in order to guarantee that at any time t any particular consumer pays a single price for a given good, assume that q_{gskt} is positive for exactly one seller for any particular g,k,t combination. Combining all consumers' purchases of a good from all the sellers selling that good gives the quantity of that good in the aggregate market basket. The aggregate Laspeyres index is thus defined as

$$(1) \qquad I_t = \frac{\sum_{g \in G} \sum_{s \in S_g} \sum_{k \in K} p_{gst} q_{gsk0}}{\sum_{g \in G} \sum_{s \in S_g} \sum_{k \in K} p_{gs0} q_{gsk0}}.$$

Let w_{g0} be the proportion of aggregate expenditures devoted to good g in period 0:

$$(2) \qquad w_{g0} = \frac{\sum_{s \in S_g} \sum_{k \in K} p_{gst} q_{gsk0}}{\sum_{\gamma \in G} \sum_{s \in S_g} \sum_{k \in K} p_{\gamma s0} q_{\gamma sk0}}$$

Also, let L_{gt} be the Laspeyres component index for good g:

$$(3) \qquad L_{gt} = \frac{\sum_{s \in S_g} \sum_k p_{gst} q_{gsk0}}{\sum_{s \in S_g} \sum_k p_{gs0} q_{gsk0}}.$$

Then equation (1) can be written in terms of w_{g0} and L_{gt} as

$$(4) \qquad I_t = \sum_{g \in G} w_{g0} L_{gt}.$$

As a Laspeyres index, L_{gt} can itself be constructed as a weighted average of sellers' price ratios. Let σ_{gs0} be seller s's base-period share of consumers' expenditures on good g. Then,

$$(5) \qquad L_{gt} = \sum_{s \in S_g} \sigma_{gs0}(p_{gst}/p_{gs0}).$$

10.2 CPI Component Indexes

Directly applying equation (2) or equation (4) to calculate the components of a Laspeyres price index is impossible because the large number of sellers precludes taking a census of their prices. In addition, collecting base-period quantity (or expenditure) data at the seller level can be difficult and expensive. Consequently, prior to 1978 no country attempted to use the Laspeyres index concept consistently at all levels of aggregation. At that time, the Bureau of Labor Statistics (BLS) introduced a sample-based estimator of a Laspeyres component index for the lowest-level aggregates in the CPI. This approach offers three important advantages: (1) it makes possible a unified approach at all levels of aggregation; (2) it incorporates scientific sampling of sellers based on consumers' recent buying patterns; and (3) it allows calculation of reliable index standard errors that include the effects of seller sampling.

Giving small and large sellers equal sample-selection probabilities is inefficient.[2] Consequently, in most cases other than housing, BLS uses probability-proportional-to-size (PPS) sampling, where a seller's "size" is measured by the expenditures it receives in the base period. Thus, the sampling probabilities are—in theory, at least—equal to the weights that equation (4) uses to average price ratios to obtain L_t.

A simple illustration of the PPS estimator of the population value of equation (4) is as follows. Suppose that the population contains three sellers that charge prices of \$3, \$4, and \$5 in time period 0 and that consumers buy the same quantity from each seller. The three sellers' base-period expenditure shares are 3/12, 4/12, and 5/12. Hence, if at time t the sellers charge \$5, \$4, and \$3, equation (4) is

$$L_t = \frac{3}{12} \times \frac{5}{3} + \frac{4}{12} \times \frac{4}{4} + \frac{5}{12} \times \frac{3}{5} = 1.$$

Assuming that a PPS sample of size two is used to estimate L_t, the first seller should have probability of selection 1/2, the second seller should have probability of selection 2/3, and the third seller should have probability of selection 5/6. Given sampling without replacement, a sample consisting of sellers one and two must then have a 1/6 probability, selection of sellers one and three must have a 1/3 probability, and selection of sellers two and three must have a 1/2 probability. The expected value of the PPS estimator of L_t is therefore

$$\left(\text{sample selection probability of } \frac{1}{6}\right)$$
$$\times \left(\text{sample index estimate of } \frac{1}{2} \times \frac{5}{3} + \frac{1}{2} \times \frac{4}{4}\right)$$
$$+ \left(\frac{1}{3}\right) \times \left(\frac{1}{2} \times \frac{5}{3} + \frac{1}{2} \times \frac{3}{5}\right) + \left(\frac{1}{2}\right) \times \left(\frac{1}{2} \times \frac{4}{4} + \frac{1}{2} \times \frac{3}{5}\right) = 1.$$

Note that equal sampling probabilities would have implied a higher expected value of 49/45 for the index estimator. Such equal sampling probabilities could occur if consumer substitution of lower-priced items made quantities inversely proportional to prices, or if the sellers' expenditure shares had identi-

2. It is noteworthy that estimating a Laspeyres component index by drawing a small simple random sample of sellers and then treating that sample as though it were the whole population in equation (2) or equation (4) leads to upwardly biased estimates of the population Laspeyres index. This occurs because the total base-period expenditures in the sample, which is the denominator when equation (2) is applied to a sample of sellers, would be a random variable. Randomness of the denominator raises the expected value of the sample estimator of L_{gt} because Jensen's inequality implies that $E(y/x) > E(y)/E(x)$ if x has a positive variance, where x and y are arbitrary random variables that are not perfectly correlated; see Mood, Graybill, and Boes (1974, 72). In the example immediately below, the population Laspeyres index equals 1 but the simple random sample Laspeyres index estimator has an expected value of 193/189.

cal long-run averages and those averages were used as measures of size in lieu of the actual shares occurring at the base prices.

In practice, BLS estimates the σ_{gs0} of equation (4) in two stages. First, BLS uses results from a household Point-of-Purchase Survey (POPS) to estimate outlets' shares of the expenditures on the good in question by the consumers covered by the index. These expenditure shares furnish a measure of size for PPS sampling of the outlets selling the good in question. Second, each outlet selected to be in the sample is asked to furnish a revenue breakdown for the varieties of the good that it sells. BLS then uses these revenue shares to select a PPS sample of one or more detailed varieties whose price at the outlet can be tracked over time. (For example, if the good in question is white pan bread, a selected variety might be a twenty-four-ounce loaf of white Wonder Bread.) Each year this process occurs in one-fifth of the cities furnishing CPI data, so any given city gets an updated sample of sellers every five years.

Let M_1 denote such a sample. Then if the elements of M_1 are selected with probabilities proportional to consumers' expenditures at time 0, the unbiased and efficient Horvitz-Thompson estimator of L_{gt} for the city and good in question is[3]

$$\hat{L}_{gt} = (1/n) \sum_{s \in M_1} (p_{gst}/p_{gs0}). \tag{6}$$

Processing seller-level expenditure-share data and drawing a sample reflecting those data takes time. Consequently, the sampling probabilities must necessarily reflect an earlier period of time than the initial price data from the sampled sellers. Unless expenditure shares are constant, this timing difference precludes sample-based estimation of a true Laspeyres index of sellers' prices. Assuming constancy of expenditure shares is not a solution to this dilemma because it implies that elasticities of substitution equal 1 rather than 0, as would be required for the Laspeyres index to equal a cost-of-living (COL) index.

In the case of the CPI, the POPS has long recall periods for purchases of some types of goods, and six months to two years may pass before POPS responses are used to draw an outlet sample. Furthermore, once BLS selects an outlet to be in the sample and conducts PPS sampling of its varieties, several months pass before it begins collecting prices for the new index, and outlets' estimates of their revenue breakdowns are also based on lagged (and, often, approximate) averages. In addition, both POPS expenditures and outlet revenues are generally aggregated over periods of time long enough for prices to vary. Consequently, the sampling probabilities for M_1 reflect approximate aver-

3. Horvitz-Thompson estimators and their efficiency property are discussed in Cochran (1977, 258–61). Equation (6) is not precisely the estimator used for the CPI because PPS sampling does not completely obviate the need for seller weighting in practice. Also, equation (6) presumes that the entire outlet sample has a single base period, but some CPI outlet samples consist of two or more segments that have different base periods. The details of seller weighting in the CPI are quite complex; for a more thorough description, see U.S. Department of Labor (1992).

age expenditure shares over some interval of time that precedes the beginning of price data collection by months or even years. The correlation between CPI sampling probabilities and sellers' initial prices is thus likely to be closer to zero than the correlation between contemporaneous expenditures and prices implied by Leontief behavior. Consequently, sellers with low initial prices tend to receive too much weight in the CPI, but these sellers also tend to have unusually high rates of price growth as their prices revert to more normal levels. The result is an upward bias due to positive correlation between sellers' price growth rates and errors in their weights. This bias differs from the substitution bias that textbook discussions of Laspeyres price indexes cover because it is not caused by consumer behavior. Instead, it resembles the functional form bias that Irving Fisher emphasized in his discussion of the simple arithmetic averages of price ratios that Sauerbeck used for his indexes.[4]

CPI component indexes can be expected to behave similarly to true Laspeyres component indexes if all seller-level demand elasticities equal 1. Under this circumstance, expenditure shares would be uncorrelated with base-period prices regardless of when each was measured. Consequently, both L_{gt} and $E(L_{gt})$ would equal $\sum_{s \in S_g} (p_{gst}/p_{gs0})$, where $E(\cdot)$ denotes an expected value. On the other hand, if seller-specific demand elasticities exceed 1, the use of lagged expenditure shares to calculate sampling probabilities for the CPI component indexes may actually reduce their bias compared with a COL index.

10.3 Substitution Bias in Laspeyres Component Indexes

Although the estimation of a true Laspeyres component index is infeasible, the use of this concept as an estimation goal makes it important to understand its properties. Laspeyres price indexes have the disadvantage that as relative prices depart from their initial values, the relative quantities in consumers' initial market baskets may become suboptimal. At the level of goods, it has long been known that the resulting product substitution causes a Laspeyres price index to exceed the corresponding Konüs ("true") COL index. Nevertheless, U.S. consumption data indicate that Laspeyres indexes suffer less bias from substitution between goods that many economists had suspected: using 53 categories of goods Braithwait (1980) finds an average commodity-substitution bias of just 0.1 percent per year; using 101 goods categories Manser and McDonald (1988) find an average bias of under 0.2 percent per year; and, finally, using over 200 categories of goods in forty-four localities, Aizcorbe and Jackman (1993) find an average bias of slightly over 0.2 percent per year.

4. Fisher ([1927] 1967, 29–30, 86–91, and 527) discusses Sauerbeck indexes and how, in practice, sample-based estimation of a Laspeyres index is likely to entail an upward bias similar to that of the Sauerbeck index. Pigou (1932, 79) and Törnqvist (1936, 28) also discuss the upward bias of the Sauerbeck or "arithmetic average" index.

Laspeyres indexes may perform less well below the goods level of aggregation, however. For many goods, outlets and varieties often have large price changes due to sale pricing. Moreover, consumers probably substitute more readily between different sellers of the same good than between different goods. In fact, since any monopolistic competitor choosing a price on the inelastic portion of its demand curve is not maximizing profits, seller behavior alone may often guarantee that seller-level price elasticities of demand remain above 1. In the short term, therefore, bias may accumulate rapidly in a Laspeyres component index.

A formal analysis of seller-substitution bias requires a comparison between a Laspeyres component index and a COL index benchmark. In this benchmark, an explicit treatment of aggregation across consumers is necessary because price dispersion causes consumers to have different COL indexes even if they have the same homothetic utility function.

Aggregate Laspeyres price indexes have been called "plutocratic" because they can be constructed as averages that weight individual consumers' Laspeyres price indexes in proportion to their base-period expenditures (see Pollak 1980). Similarly, a Laspeyres component index can be expressed as a plutocratic average of the ratios of the time t and time 0 prices at the sellers that consumer chose in the base period. In this case, however, consumers' weights depend on their expenditures on the good in question rather than on all goods.

Since a population Laspeyres index uses plutocratic aggregation across consumers, its COL index counterpart is a plutocratic average of consumers' Konüs COL indexes:

$$(7) \qquad K_t = \sum_k s_{k0} \frac{e(\mathbf{P}_t^{kt}, u_{k0})}{e(\mathbf{P}_0^{k0}, u_{k0})} = \frac{\sum_k e(\mathbf{P}_t^{kt}, u_{k0})}{\sum_k e(\mathbf{P}_0^{k0}, u_{k0})}.$$

Here, s_{k0} is consumer k's share of aggregate expenditures in period 0, $e(\cdot)$ is the expenditure function giving the minimum cost of achieving utility u_0, and $\mathbf{P}_t^{kt}$ is the vector of prices paid for the goods in G by consumer k at time t. Dividing the numerator and the denominator by the number of consumers in K shows that the dispersed-price COL index K_t can be interpreted as the ratio of consumers' expected expenditure functions in periods 0 and t.[5]

To avoid the added complexity of separating out commodity-substitution effects from problems arising in aggregation over sellers, we assume that consumers have Leontief preferences over goods. We further assume that consumers regard sellers as perfect substitutes. This requires the presence of search costs to explain how higher-priced sellers can make positive sales. Nevertheless, we do not include search costs in the price index because a model that

5. This dispersed-price COL index is discussed in Reinsdorf (1994a). It resembles Pollak's (1981, 328) Laspeyres-Scitovsky social COL index.

incorporated them would be very complicated and might give only limited practical guidance on how to construct a component index.[6]

If preferences over goods are Leontief, but sellers are perfectly substitutable, then $e(P_t^{kt}, u_0) = \sum_g P_{gt}^{kt} q_g^{k0}$, where P_{gt}^{kt} is the price that consumer k pays at time t, and q_g^{k0} is k's quantity of good g in the market basket yielding utility u_0. Let q_g^0 denote the combined quantity of all consumers' purchases of good g at time 0, and let $\tilde{p}_{g\tau} = \sum_s \sum_k p_{gs\tau} q_{gsk\tau} / q_g^0$, the average price paid for good g at time τ. As a ratio of total expenditures on the good to the total quantity of the good that is sold, $\tilde{p}_{g\tau}$ is, in fact, a unit value. Aggregating over consumers shows that the ratio of expected expenditure functions equals $\sum_g \tilde{p}_{gt} q_g^0 / \sum_g \tilde{p}_{g0} q_g^0$. Using the fact that $w_{g0} = \tilde{p}_{g0} q_g^0 / \sum_\gamma \tilde{p}_{\gamma 0} q_\gamma^0$, this index can also be expressed as

$$K_t = \sum_{g \in G} w_{g0} (\tilde{p}_{gt} / \tilde{p}_{g0}) \tag{8}$$

Comparing equations (3) and (8) reveals that the COL index differs from the Laspeyres price index only in the way it measures price change for each good. The COL index's components are ratios of the average prices paid in the comparison and base periods. In contrast, the Laspeyres component indexes compare the average price that consumers would have paid if they repeated their base-period seller choices to the average price that they paid in the base period. Under the assumption of no quality differences between sellers, the difference between a Laspeyres component index and the ratio of the average prices paid by consumers is a measure of seller-substitution bias.

10.4 Trend Reversion of Sellers' Prices and Bias in Chained Component Indexes

Forsyth and Fowler (1981, 241) report that oscillating prices are commonly observed in constructing basic component indexes. One reason for this is that sellers' prices tend to exhibit trend reversion.[7] Although competition rarely acts quickly enough to make retail market obey Jevons's law of one price, it does prevent sellers from deviating indefinitely from their market's overall price trend. Consumers' propensity to buy from low-priced sellers rather than high-priced sellers puts downward pressure on prices that are comparatively high, and it may also put upward pressure on prices that are unusually low by making them attract high sales. Moreover, consumers' willingness to substitute

6. Reinsdorf (1994a) discusses implications of commodity substitution for dispersed-price COL indexes. Anglin and Baye (1987) develop a COL index that does include search costs.

7. Assuming independence of sellers' prices and their subsequent rates of change implies that these prices follow a random walk and hence have a nonstationary distribution whose variance grows over time without bound. Thus, at least in a linear time-series framework, the weak assumption that competing sellers' price discrepancies are bounded is sufficient for prices' changes to be negatively correlated with their starting levels. Friedman (1992) offers an interesting perspective on the ubiquity of trend reversion in time series.

an outlet or variety offering a low price sometimes makes it profitable for sellers to run off-price specials to build customer traffic, to introduce new consumers to a product, or simply to sell large quantities to price-sensitive consumers who would otherwise not buy. Consequently, prices that are low compared to the average price in a market tend to rise, while comparatively high prices tend to fall or to remain stable.

In the absence of new sellers, such reversion-to-trend behavior would cause the average yearly bias of any type of index of sellers' prices to tend toward zero in the long run as every seller's average rate of price change asymptotically approached a common value. In the short run, however, trend reversion can exacerbate the bias of a Laspeyres, Sauerbeck, or CPI component index. For example, suppose that sellers sometimes offer highly discounted "sale" prices. A Sauerbeck or CPI component index will implicitly give large quantity weights to those sellers offering sale prices in the base period, and a true Laspeyres component index will also assign them large quantities if consumers readily substitute between sellers. Such a weighting pattern makes the index rise rapidly as the sale prices revert to their regular values.

The compounding of such high short-run biases could severely affect a long-run component index calculated by linking together a succession of short-run indexes. Yet unfortunately, flux in populations of sellers makes it necessary to update the market basket periodically to reflect changes in consumers' purchasing patterns. Otherwise, the fixed market basket represented by the index might become quite unrepresentative as entry and exit of sellers caused large permanent changes in consumers' purchasing patterns.

Although it may seem paradoxical that chaining exacerbates a Laspeyres component index's substitution bias, Christensen and Manser (1976, 442) report empirical evidence of this in their index for meat. Szulc (1983) demonstrates theoretically that in a Laspeyres or a Sauerbeck index, the effect of chaining depends on the pattern of price changes. Different price-change patterns may tend to emerge at different levels of aggregation. Broadly defined goods may often be subject to price trends that persist for many months or years, causing positive autocorrelation in their changes. In contrast, at the seller level, prices' reversion to trend can be expected to cause negative autocorrelation of price changes. As Frisch (1936, 9) observes, under this circumstance, chaining a Laspeyres or a Sauerbeck index causes it to drift upward compared to its unchained counterpart.

The mathematical explanation of how negatively autocorrelated price changes lead to spurious increases when indexes that arithmetically average price ratios are chained relies on the fact that the expected value of a product of two variables equals the product of their expected values plus their covariance—see Mood, Graybill, and Boes (1974, 180). When component indexes are chained, products of variables with zero covariances tend to replace products of variables with negative covariances. In particular, a price's change from period 0 to period t equals the product of two subinterval price changes:

(9) $$p_{gst}/p_{gs0} = (p_{gst-1}/p_{gs0})(p_{gst}/p_{gst-1}).$$

Suppose that price changes are negatively autocorrelated and that the second term in the above product is replaced by a new term, p_{gzt}/p_{gzt-1}, that is independent of p_{gst-1}/p_{gs0}. Then even if the expected value of the new term equals the *unconditional* expected value of the term it replaces, the new product will have a higher expected value than the original product.

10.5 A Statistical Approach to Evaluating Bias in Basic Component Indexes

The presumption that the prices offered by competing sellers in a market tend to have stationary (inflation-adjusted) distributions suggests the use of a statistical model for evaluating alternative component index formulas. Another advantage of considering statistical properties along with COL indexation properties is that tracking changes in consumers' cost of living—though the goal of the CPI—is only one of its uses. For example, the CPI is sometimes used as an indicator of the effectiveness of monetary policy or of general price trends in the economy.

Suppose that prices are generated by the following model:

(10) $$\log p_{gst} = \pi_{gt} + u_{gs} + e_{gst},$$

where π_{gt} is the log of the true price trend for good g; u_{gs} is a stochastic permanent component of a seller s's price, e_{gst} is a stochastic transitory component; u_{gs} and e_{gst} are independent; and e_{gst} is normal with mean zero, variance σ_g^2, and $\mathrm{Corr}(e_{gst}, e_{gs\tau}) = \rho_{t-\tau}(t > \tau)$. The e_{gst} represent sellers' temporary deviations from the market trend such as could be caused by special sale pricing or by differences among sellers in the timing of price increases. They play a critical role in the results below, whereas the other disturbances end up having no effect on the value to which the index converges as the seller sample grows large.

For sellers' quantities $q_{gst} = \Sigma_k q_{gskt}$, assume the following constant demand elasticity model:

(11) $$\log q_{gst} = -\eta_g \log p_{gst} + \delta_{gt} + \nu_{gs} + \omega_{gst}.$$

Here η_g is the elasticity of demand for the item, δ_{gt} is a time-period effect, ν_{gs} is a stochastic permanent component of an item's quantity demanded, ω_{gst} is a stochastic transitory component, and ν_{gs} and ω_{gst} are independent of e_{gst}. The constant elasticity specification is chosen for tractability, and other specifications that also imply an inverse relation between quantities and prices can be expected to imply qualitatively similar results.

Henceforth we omit the g subscripts for convenience. Note that although normality is assumed for e_{st}, no distributional assumption is required for u_s,

v_s, or ω_{s0}. We do assume, however, that $\exp[(1 - \eta)(u_s + e_{s0}) + v_s + \omega_{s0}]$ and $\exp[(1 - \eta)u_s + e_{st} - \eta e_{s0} + v_s + \omega_{s0}]$ have constant, finite variances.

Consider first the Laspeyres component index of equation (3). The appendix uses the properties of the log-normal distribution to show that, as the number of sellers becomes large, the true Laspeyres component index converges in probability to

$$\text{plim } L_t = \exp(\pi_t - \pi_0) \times \exp[\eta(1 - \rho_t)\sigma^2]. \tag{12}$$

After t periods, the upward bias of the logarithm of the Laspeyres component index is $\eta(1 - \rho_t)\sigma^2$. Large demand elasticities raise the bias by causing transitory disturbances in time 0 prices to have a large effect on time 0 quantities. On the other hand, high correlations between time 0 and time t prices reduce the bias by weakening the inverse relation between sellers' price levels at time 0 and their rates of price change from time 0 to time t.

If BLS could obtain sellers' price histories back to the period covered by the POPS, it could consistently estimate the Laspeyres index with a market basket from that time period. Since this index would be linked into the CPI at a much later time, the Laspeyres index would, in effect, be aged before it is used. The assumption in the present model that sellers' prices are stationary around a common trend implies that such an aged Laspeyres index would have almost no bias. In particular, suppose that a Laspeyres index with base period 0 were linked into a chained index at time l and used to measure price change from time l to time t. The probability limit of the index from time l to time t would be

$$\text{plim } L_t/L_l = \exp(\pi_t - \pi_0) \times \exp[\eta(\rho_l - \rho_t)\sigma^2]. \tag{13}$$

Unless time l is very close to time 0, $\rho_l - \rho_t$ is likely to be close to zero. This will make the bias factor in equation (13) approach 1 even if η and σ^2 are large.

Next, consider an average-of-ratios or Sauerbeck index. Since this index does not use weights that depend on consumers' behavior, it should not depend on η. In fact, its probability limit is simply $E(p_{st}/p_{s0})$, which is the Laspeyres index's probability limit when $\eta = 1$. The Sauerbeck index is biased by a factor whose logarithm equals half the variance of sellers' rates of price change from time 0 to time t. This follows from the properties of the log-normal distribution and from equation (10), which implies that

$$\text{Var}(\log p_{st} - \log p_{s0}) = 2(1 - \rho_t)\sigma^2. \tag{14}$$

The CPI component index probably lies somewhere between a true Laspeyres index and a Sauerbeck index. Its seller sampling probabilities approximately reflect the average shares that outlets, and varieties within outlets, had of consumers' expenditures during some historical time period, which we denote as period h. Also, denote by l (for "link month") the time when BLS begins to collect prices for use in the new component index that is linked into the CPI. The CPI measure of price change from link month l to period t for an index area whose entire seller sample rotates simultaneously is

$$(15) \qquad R_{t,l} = \frac{\sum_s W_{s0} \times p_{st}/\hat{p}_{s0}}{\sum_s W_{s0} \times p_{sl}/\hat{p}_{s0}},$$

where $\hat{p}_{s0} = p_{sl}/(\hat{L}_l/\hat{L}_0)$ and $W_{s0} \propto p_{sh}q_{sh}$ by assumption. Equation (15) can be rewritten as

$$(16) \qquad R_{t,l} = \frac{\sum_s p_{sh}q_{sh} \times p_{st}/p_{sl}}{\sum_s p_{sh}q_{sh}}.$$

Assuming that no changes in sample composition occur between period l and period t, and that $\exp[(1 - \eta)(u_s + e_{sh}) + v_s + \omega_{sh}]$ and $\exp[(1 - \eta)(u_s + e_{sh}) + e_{st} - e_{sl} + v_s + \omega_{sh}]$ have constant, finite variances, the appendix shows that equation (16) converges in probability to

$$(17) \qquad \text{plim } R_{t,l} = \exp(\pi_t - \pi_l) \times \exp\{[1 - \rho_{t-l} + (\eta - 1)(\rho_{l-h} - \rho_{t-h})]\sigma^2\}.$$

Note that if time l = time h = time 0, then equation (17) equals the probability limit for the true Laspeyres index. Also, if $\rho_{l-h} = \rho_{t-h}$ or if $\eta = 1$, then equation (17) equals the probability limit of a Sauerbeck index with base month l. The remoteness of time l from time h makes near equality between ρ_{l-h} and ρ_{t-h} quite plausible.

As a numerical example, suppose that $\eta = 1.5$, $\sigma^2 = 0.02$, and $\rho_t = 0.6^t$. Then for the first period ($t = 1$), the Laspeyres index overstates inflation by $e^{0.012} - 1$, or 1.21 percent. In this example, the index continues to overestimate inflation during subsequent periods, but the magnitude of the overestimate declines. For the first six periods, the cumulative asymptotic seller-substitution bias in the example amounts to 2.90 percent. After the sixth period, very little additional overstatement of inflation occurs, and the index eventually converges to a value that is too large by 3.05 percent. By contrast, the asymptotic bias of the CPI component index in equation (17), assuming that $l - h = 5$ months, is 0.83 percent the first month. The cumulative expected overestimate for the first six periods is 2 percent, and the CPI would eventually converge to a value that is too large by 2.10 percent. We find that the cumulative overstatement of price change in the CPI formula is relatively insensitive to η and usually is close to σ^2.

10.6 The Use of Geometric Means in Basic Component Indexes

Empirical results in Reinsdorf (1993) and (1994b) suggest that in food at home and gasoline portions of the U.S. CPI, the bias of the chained Laspeyres component index estimator could be substantial. Forsyth and Fowler (1981), Szulc (1989), and Turvey (1989) suggest that using geometric means of sellers' prices or price relatives may prevent upward bias from price oscillations in component price indexes. In 1987 the International Labour Organisation adopted a motion calling on its member countries to consider the use of geometric means for constructing basic component indexes, though no European or North American country had done so by 1994.

We now consider the limiting properties of the geometric mean estimator under the statistical model of the last section. Assume again that sellers' weights (or probabilities of selection) are proportional to expenditures during initiation period h, that no changes in sample composition occur between link period l and period t, and that $\exp[(1 - \eta)(u_s + e_{sh}) + v_s + \omega_{sh}]$ and $\exp[(1 - \eta)(u_s + e_{sh}) + e_{st} - e_{sl} + v_s + \omega_{sh}]$ have constant, finite variances. The geometric mean measure of price change from link month l to period t for an index area whose entire seller sample rotates at once can then be written as

$$(18) \qquad G_{t,l} = \exp\left[\frac{\sum_s p_{sh}q_{sh} \times \log(p_{st}/p_{sl})}{\sum_s p_{sh}q_{sh}}\right].$$

The appendix shows that this expresion converges in probability to

$$(19) \qquad \operatorname{plim} G_{t,l} = \exp(\pi_t - \pi_l) \times \exp[(\eta - 1)(\rho_{l-h} - \rho_{t-h})\sigma^2].$$

If expenditure shares are unaffected by sellers' prices because seller-level price elasticities of demand equal 1, the geometric mean index is unbiased. At the seller level, however, η may often exceed 1. In this case, the gap between time h and time l (which, when $\eta = 0$, leads to bias in the CPI component index) becomes an advantage as ρ_{t-h} and ρ_{l-h} largely offset each other. For the numerical example in the preceding section, the weighted geometric mean formula converges to a value that is too large by 0.08 percent. Furthermore, the expenditure-share measures used for CPI sample selection may often reflect average behavior over an interval of time. Such average expenditure shares, which are analogous to estimates of a fixed effect in a panel data regression, may have very little correlation with prices at times l and t. If so, ρ_{t-h} and ρ_{l-h} in equation (19) should be replaced with correlation coefficients that are close to zero. In this case, equation (19) suggests that using geometric means would virtually eliminate the bias in the CPI components.

An analysis of how geometric mean component indexes perform in the context of the economic theory of the COL index is also important. We are currently pursuing research on this topic.

10.7 The Effects of Geometric Mean Component Indexes on the CPI

This section empirically compares the Laspeyres and geometric mean formulas for basic CPI components. The basic component indexes that have been recomputed cover those strata of goods that use the POPS/outlet-rotation sampling method (which carry approximately 70 percent of the weight, or relative importance, of the CPI), over the period from June 1992 to June 1993. The

CPI database was reconstituted from archived data, and two sets of basic component indexes were calculated using exactly the same price quotes.[8] One set simulates the CPI's Laspeyres-type component indexes, though it occasionally differs slightly from the published indexes because the computer program used to simulate the CPI is not identical to the program used in actual CPI production. The other set uses the alternative geometric mean formula. In aggregating both sets of basic components to higher levels, we have used the usual CPI Laspeyres formula and Consumer Expenditure Survey–based aggregation weights. There is a difference in formulas only at the lowest level of aggregation.

Table 10.1 compares the annual percentage changes for the two sets of indexes for various items over the period. Note that the geometric mean component indexes almost always exhibit lower rates of price growth than the Laspeyres component indexes do, a result that is not surprising in view of the known properties of the two types of averages.[9] More importantly, the size of the difference between the two indexes varies substantially between classes of items. For fresh fruits and vegetables and apparel, the Laspeyres indexes showed rates of change 2 to 3 percentage points higher than the geometric mean indexes. These differences are comparable in magnitude to the large differences in rates of change between CPI and average price series for food that have been noted by Reinsdorf (1993, 1994b). The large differences in annual rates of change for these expenditure classes are consistent with the model sketched out above, as fresh fruits and vegetables and apparel tend to have highly variable prices, due either to the perishability of food items or to the use of frequent sales with substantial discounting.

For other expenditure categories, however, the differences tend to be smaller, in most cases less than 1 percent a year. For some expenditure categories where sale pricing is rare, such as automobile parts and equipment and apparel services, there is little difference between the Laspeyres and geometric mean indexes.

Another implication of the model sketched out above is that the largest differences in measured rates of change between the Laspeyres and geometric mean indexes should occur immediately following sample rotation. Table 10.2 compares the rates of change of local area indexes based on the simulated Laspeyres and geometric mean estimators for the basic components.

8. Ken Stewart, Claire Gallagher, and Karin Smedley of the Division of Consumer Prices and Price Indexes of the BLS developed the computer programs and estimates for the two sets of indexes. Most of these empirical results have appeared in Moulton (1993), though in this paper we correct and extend the sample used in table 10.3.

9. A well-known mathematical result is that the geometric mean of a set of positive numbers having a positive variance must be less than the corresponding arithmetic mean. This result applies to the geometric mean index number formula in equation (18) and the Laspeyres-like index number formula in equation (15) only if the period whose inflation rate is measured begins with the link month. During subsequent periods it is possible for the geometric mean index to imply higher inflation than the Laspeyres index, although this seldom seems to occur in practice.

Table 10.1 **Rates of Change of Simulated Consumer Price Index for All Urban Consumers, U.S. City Average, June 1992–June 1993 (%)**

Expenditure Category	Laspeyres Index	Geometric Mean Index	Difference
All available items (70.3% of all items)	*2.95*	*2.48*	*0.47*
Food and beverages	*2.11*	*1.56*	*0.55*
Food	2.18	1.59	0.59
Food at home	2.37	1.52	0.85
Cereals and bakery products	3.36	2.78	0.58
Meat, poultry, fish, and eggs	3.90	3.28	0.62
Dairy products	1.61	1.29	0.33
Fruits and vegetables	1.58	−0.70	2.28
Fresh fruits and vegetables	4.09	1.09	3.00
Processed fruits and vegetables	−3.03	−3.98	0.96
Other food at home	0.86	0.39	0.48
Sugar and sweets	−0.09	−0.59	0.50
Fats and oils	−0.02	−0.46	0.44
Nonalcoholic beverages	−0.27	−0.64	0.38
Other prepared foods	2.22	1.67	0.55
Food away from home	1.85	1.70	0.14
Alcoholic beverages	1.48	1.32	0.16
Housing	—	—	—
Shelter	—	—	—
Renters' costs	—	—	—
Homeowners' costs	—	—	—
Maintenance and repairs	1.90	1.84	0.06
Fuel and other utilities	—	—	—
Fuels	3.55	3.54	0.01
Fuel oil and other household fuel commodities	0.38	0.31	0.07
Gas (piped) and electricity (energy services)	3.88	3.87	0.01
Other utilities and public services	—	—	—
Household furnishings and operation	—	—	—
House furnishings	0.14	−0.53	0.67
Housekeeping supplies	1.22	0.59	0.63
Housekeeping services	—	—	—
Apparel and upkeep	*0.59*	*−1.21*	*1.80*
Apparel commodities	0.47	−1.52	1.98
Men's and boys' apparel	0.19	−1.31	1.50
Women's and girls' apparel	0.44	−2.43	2.87
Infants' and toddlers' apparel	−0.13	−0.01	−0.13
Footwear	0.21	0.03	0.17
Other apparel commodities	1.83	−0.84	2.67
Apparel services	1.79	1.71	0.08
Transportation	—	—	—
Private transportation	—	—	—
New vehicles	2.46	2.30	0.16
New cars	2.23	2.09	0.15

Table 10.1 (continued)

Expenditure Category	Laspeyres Index	Geometric Mean Index	Difference
Used cars	—	—	—
Motor fuel	−3.03	−3.04	0.01
Maintenance and repairs	1.90	1.84	0.06
Other private transportation	—	—	—
Other private transportation commodities	−1.75	−1.85	0.10
Other private transportation services	—	—	—
Automobile insurance	5.39	5.33	0.06
Automobile finance charges	—	—	—
Automobile fees	5.21	5.19	0.02
Public transportation	13.19	12.86	0.33
Medical care	—	—	—
Medical care commodities	3.58	3.19	0.38
Medical care services	—	—	—
Professional medical services	5.37	4.99	0.38
Hospital and related services	8.74	8.24	0.49
Entertainment	—	—	—
Entertainment commodities	—	—	—
Reading materials	3.60	3.22	0.38
Sporting goods and equipment	—	—	—
Toys, hobbies, and other entertainment	1.07	0.37	0.70
Entertainment services	3.32	2.57	0.74
Other goods and services	*6.41*	*6.05*	*0.35*
Tobacco and smoking products	7.80	7.20	0.60
Personal care	2.43	2.05	0.38
Toilet goods and personal care appliances	2.40	1.89	0.51
Personal care services	2.48	2.24	0.24
Personal and educational expenses	7.06	6.83	0.23
School books and supplies	3.76	3.79	−0.03
Personal and educational services	7.27	7.03	0.24

Notes: A dash indicates that data are not available (usually because the index includes some strata that are not part of the POPS survey and sample rotation). Rates of change of the simulated Laspeyres indexes are not identical to the published rates of change of the CPI, because of differences between the index simulation and the actual index calculation and because the simulated indexes were not rounded prior to computing rates of change. For both indexes, aggregation above the level of the basic components (i.e., indexes of strata of items and areas) was based on the usual Laspeyres formula and weights that were in turn based on the Consumer Expenditure Survey.

Table 10.2 **Rates of Change of Simulated Consumer Price Index for All Urban Consumers, Selected Local Areas, June 1992–June 1993 (%)**

Local Area	Laspeyres Index	Geometric Mean Index	Difference
All available items (70.3% of all items)			
Chicago/Gary/Lake County, IL/IN/WI	3.57	2.81	0.76
Los Angeles/Anaheim/Riverside, CA	2.83	2.34	0.49
New York/Northern New Jersey/Long Island, NY/NJ/CT	3.11	2.54	0.57
Philadelphia/Wilmington/Trenton, PA/NJ/DE/MD	2.22	1.78	0.44
San Francisco/Oakland/San Jose, CA	2.66	1.77	0.89
Baltimore, MD[a]	2.02	1.63	0.40
Cleveland/Akron/Lorain, OH[a]	2.17	1.66	0.51
Miami/Fort Lauderdale, FL[a]	4.22	3.95	0.27
St. Louis/East St. Louis, MO/IL[a]	0.56	0.45	0.11
Washington, DC/MD/VA[a]	3.52	3.16	0.36
Dallas/Fort Worth, TX	1.96	1.32	0.64
Detroit/Ann Arbor, MI	2.95	2.37	0.58
Houston/Galveston/Brazoria, TX	2.18	2.35	−0.17
Pittsburgh/Beaver Valley, PA	3.21	2.66	0.55
Food at home			
Chicago/Gary/Lake County, IL/IN/WI	2.62	2.00	0.62
Los Angeles/Anaheim/Riverside, CA	4.21	3.61	0.60
New York/Northern New Jersey/Long Island, NY/NJ/CT	1.65	0.35	1.30
Philadelphia/Wilmington/Trenton, PA/NJ/DE/MD	1.69	2.13	−0.44
San Francisco/Oakland/San Jose, CA	2.62	0.06	2.56
Baltimore, MD	2.12	2.05	0.06
Boston/Lawrence/Salem, MA/NH	3.55	3.46	0.09
Cleveland/Akron/Lorain, OH	2.81	2.47	0.34
Miami/Fort Lauderdale, FL	5.81	5.34	0.47
St. Louis/East St. Louis, MO/IL	−2.55	−2.69	0.14
Washington, DC/MD/VA	1.86	2.03	−0.17
Dallas/Fort Worth, TX	2.32	1.81	0.51
Detroit/Ann Arbor, MI	1.32	1.04	0.28
Houston/Galveston/Brazoria, TX	−0.59	−1.68	1.09
Pittsburgh/Beaver Valley, PA	3.38	2.72	0.66

Notes: Rates of change of these simulated Laspeyres indexes are not identical to the published rates of change of the CPI, because of differences between the index simulation and the actual index calculation and because the simulated indexes were not rounded prior to computing rates of change. For both indexes, aggregation above the level of the basic components (i.e., indexes of strata of items and areas) was based on the usual Laspeyres formula and weights that were in turn based on the Consumer Expenditure Survey.

[a]Because of the bimonthly sampling for nonfood items, the period for these indexes is July 1992–May 1993.

From June 1992 to June 1993, only one of the local areas listed in table 10.2 had its sample replaced: San Francisco. San Francisco had the largest difference in rates of change for the Laspeyres and geometric mean component indexes for both all available items and food at home. Another local area, New York, had part of its sample replaced during this period. New York has the second largest difference for food at home. In the other areas that did not introduce a new sample during the June 1992–June 1993 period, the Laspeyres component indexes also showed a larger rate of change than the geometric mean component indexes did, but the differences were smaller than those for the cities that rotated their samples. The effect of rotation is particularly noticeable when one examines the month-to-month differences. For San Francisco, the Laspeyres component index for food at home produced a rate of change 1.11 percentage points larger than the geometric mean component index during the month after the new sample was introduced. For New York, the difference during the month following the introduction of the partial new sample was 1.49 percentage points.

10.8 Other Empirical Evidence

If using a Laspeyres type of formula causes an index to overstate significantly the inflation rate immediately following sample rotation, evidence of the effect should appear in the historical behavior of the indexes. Because the samples in the smaller urban areas do not all rotate at the same time, we examined the price changes for large urban areas (A-size primary sampling units) immediately following rotation. Rotation schedules designating the link month for the two samples were obtained for the years 1980–85 and 1988–93. The link months are listed in the note to table 10.3.

Table 10.3 presents the mean difference between the measured inflation rate for the rotated area a and the U.S. average inflation rate during two separate periods: the two-month period and the six-month period after the rotated samples are introduced. If introducing the new sample induces a positive shock to the inflation rate, it should result in positive values for the mean difference.

The results shown in the table are generally consistent with this prediction of the model. There are significant positive differences between the area inflation rates and the U.S. average inflation rates for food, especially fruits and vegetables and meat. The numerical magnitude of these differences, however, appears to be too small to explain the entire difference between the geometric mean indexes and the Laspeyres indexes. For example, if the Laspeyres index overstates the inflation rate for fruits and vegetables by about 2 percent a year, as suggested by the comparison with the geometric-mean index, and if most of the overstatement occurs shortly after each five-year rotation, then we might expect a 10 percentage point differential in the inflation rate immediately following each rotation. The observed differentials for fruits and vegetables in table 10.3 are 2.6 percent for the two-month period and 3.1 percent for the six-

Table 10.3 **Mean Differences in Measured Inflation between Consumer Price Indexes for A-Size Primary Sampling Units and U.S. Average Consumer Price Indexes**

Expenditure Category	Two Months after Sample Rotation	Six Months after Sample Rotation
All items	*0.07 (0.10)*	*−0.00 (−0.16)*
All items less shelter	0.04 (0.10)	−0.04 (0.15)
Food and beverages	0.58[a] (0.13)	0.59[a] (0.23)
Food	0.62[a] (0.14)	0.61[a] (0.24)
Food at home	0.96[a] (0.20)	1.02[a] (0.34)
Cereals and bakery products	0.60[a] (0.29)	−0.06 (0.32)
Meats, poultry, fish, and eggs	1.25[a] (0.25)	1.40[a] (0.45)
Dairy products	−0.32 (0.29)	−0.14 (0.48)
Fruits and vegetables	2.57[a] (0.58)	3.11[a] (0.96)
Other food at home	0.45 (0.29)	0.53 (0.34)
Food away from home	−0.02 (0.11)	−0.12 (0.19)
Alcoholic beverages	−0.04 (0.21)	−0.05 (0.32)
Transportation	0.05 (0.14)	−0.27 (0.23)
Motor fuels	−0.28 (0.40)	−0.66 (0.48)
Medical care	0.01 (0.20)	0.14 (0.32)
Entertainment	−0.40 (0.27)	−0.27 (0.50)
Other goods and services	0.19 (0.19)	0.10 (0.27)

Notes: Numbers in parentheses are standard errors of the means. The sample sizes are $N = 37$ for the two-month comparisons and $N = 36$ for the six-month comparisons. The indexes come from the BLS LABSTAT program. The rotation link months for A-size primary sampling units used in the analysis are as follows: Philadelphia—January 1980, January 1985, February 1989; Boston—July 1983, January 1989; Pittsburgh—October 1982, October 1991; Buffalo—August 1980, February 1985; Detroit—February 1981, August 1991; St. Louis—July 1980, September 1990; Cleveland—October 1982, July 1993; Minneapolis—October 1983; Milwaukee—January 1984; Cincinnati—September 1984; Kansas City—August 1984; Washington—January 1981, July 1991; Dallas—June 1981, October 1990; Baltimore—September 1983, July 1989, November 1993; Houston—June 1982, June 1993; Atlanta—August 1982; Miami—July 1983, July 1993; San Francisco—August 1982, November 1992; Seattle—July 1981; San Diego—September 1983; Honolulu—August 1984; Anchorage—January 1980.

[a]Significance at the 5 percent level in a one-sided test of H_0: Diff = 0 versus H_a: Diff > 0, where Diff is defined as $\text{Diff}_2 = 100 \times (R^a_{l+2,l} - R^{US}_{l+2,l})$ and $\text{Diff}_6 = 100 \times (R^a_{l+6,l} - R^{US}_{l+6,l})$, in which $R^a_{l+2,l}$ is the CPI change in prices for area a in the first two months after link month l, and $R^{US}_{l+2,l}$ is the same period's change in the U.S. average CPI.

month period after new samples are introduced. One possible explanation is that the autocorrelation of the individual transitory component of prices may diminish slowly, rather than rapidly as has been assumed.[10] Another possible explanation is that sample-initiation effects also occur between sample rotations as items drop out of the sample and are replaced with substitutes.

Another prediction of the theoretical model is that the Laspeyres index will have the greatest tendency to overstate inflation when price oscillation is

10. For example, the autocorrelations would die down slowly if the transitory component followed a fractionally integrated time-series process. See, for instance, Beran (1992).

Table 10.4 **Rates of Change of Simulated Consumer Price Index, Selected Items, June 1992–June 1993**

Item	Laspeyres Index (%)	Geometric Mean Index (%)	Difference	$\mathrm{Var}\left[\log\left(\frac{p_{\mathrm{Jun}\,93}}{p_{\mathrm{Jun}\,92}}\right)\right]$
White bread	2.70	1.86	0.84	.0290
Round roast	4.40	4.48	−0.08	.0624
Round steak	4.49	4.12	0.37	.0523
Bacon	7.36	7.43	−0.06	.0403
Pork chops	3.79	3.45	0.34	.0397
Fresh whole chicken	5.82	5.00	0.82	.0497
Bananas	−3.22	−3.89	0.67	.0976
Oranges	−4.82	−7.82	3.00	.1108
Lettuce	3.84	2.12	1.72	.1509
Tomatoes	60.00	55.69	4.31	.1603

Notes: Rates of change of the simulated Laspeyres indexes are not identical to the published rates of change of the CPI because of differences between the index simulation and the actual index calculation and because the simulated indexes were not rounded prior to computing rates of change. For both indexes, aggregation above the level of the basic components (i.e., indexes of strata of items and areas) was based on the usual Laspeyres formula and weights that were in turn based on the Consumer Expenditure Survey.

largest. Equation (14) implies that we can measure the degree of price oscillation by $\mathrm{Var}(\log p_{st} - \log p_{s0})$, the variance of the logarithms of the price relatives.

To test this model prediction, we selected ten strata of food-at-home items that, in our opinion, are likely to be nearly homogeneous. Table 10.4 presents a comparison of the Laspeyres and geometric mean indexes, as well as Var $(\log p_{\mathrm{Jun}93} - \log p_{\mathrm{Jun}92})$, for these items. The three strata with the largest variances—oranges, lettuce, and tomatoes—also have the largest differences between the Laspeyres and geometric mean indexes, more than 1 percentage point in each case. The Laspeyres index does indeed have the greatest tendency to overstate inflation when the variance in the logarithm of the price differences is largest.

10.9 Conclusion

Since 1978, the basic components of the CPI have been sample estimators of Laspeyres indexes which weight outlets and varieties by their base-period quantities. To estimate these base-period quantities, BLS divides base-period expenditure estimates by link-month prices that have been deflated back to the base period.[11] Unfortunately, this way of imputing base-period prices results in

11. Between the time that this paper was originally written and when it went to press, BLS changed to a new method of imputing base-period prices that avoids the use of link-month prices. BLS chose not to adopt the geometric mean index solution because it would have entailed a fundamental change in the index concept. A discussion of the new method of imputing base prices and its likely effect on the CPI may be found in Moulton (1996).

a positive correlation between errors in weights and price changes subsequent to the link month. In other words, the CPI component index estimator tends to give too much weight to prices that increase, and too little weight to prices that decline.

The bias from this positive correlation between weighting errors and price changes may explain much of the growth rate discrepancies between CPI component indexes and matched Average Price series reported by Reinsdorf (1993). In particular, this bias probably raised the growth rates of the CPI components by more than consumers' substitution of lower priced outlets and varieties lowered the growth rates of BLS's Average Price series. A substantial effect of this type would be consistent with Dalén's (1992, 144) finding that calculating the basic component indexes of the Swedish CPI as averages of price ratios raised their growth rates considerably.

A possible solution to the functional form bias problem may be to use geometric mean indexes. Geometric means are especially suitable for use with lagged and averaged expenditure-share data, which are generally the only data available for weighting purposes. Empirical tests of the effect of using geometric mean indexes suggest that their adoption for items other than shelter might reduce the inflation rate of the "all items" CPI by about 0.4 percent per year.

Appendix
Probability Limits under the Statistical Model of Section 10.5

Derivation of Equation (12)

In large samples, $(1/n)\sum p_{st}q_{s0}$ converges in probability to $E(p_{st}q_{s0})$ and $(1/n)\sum p_{st}q_{s0}$ converges in probability to $E(p_{s0}q_{s0})$, so L_t converges in probability to $E(p_{st}q_{s0})/E(p_{s0}q_{s0})$ (see White 1984, 22–24).

Under the model set out in equations (10) and (11), the two expectations are

$$\text{(A1)}\quad \frac{E(p_{st}q_{s0})}{E(p_{s0}q_{s0})} = \frac{E\left\{\exp\begin{bmatrix} \pi_t - \eta\pi_0 + (1-\eta)u_s \\ + e_{st} - \eta e_{s0} + \delta_0 + \nu_s + \omega_{s0}\end{bmatrix}\right\}}{E\{\exp[(1-\eta)(\pi_0 + u_s + e_{s0}) + \delta_0 + \nu_s + \omega_{s0}]\}}$$
$$= \exp(\pi_t - \pi_0)\frac{E[\exp(e_{st} - \eta e_{s0})]}{E\{\exp[(1-\eta)e_{s0}]\}}.$$

This result is derived by taking the expectation of the antilog of the sum of the right-hand sides of equations (10) and (11) for the appropriate time periods, substituting the expression in equation (10) for $\log p_{st}$ in equation (11), and noting that if two random variables x and y are independent, then $E[f(x)g(y)] = E[f(x)]E[g(y)]$. This allows $E\{\exp[(1 - \eta)u_s + \delta_0 + v_s + \omega_{s0}]\}$ to be factored out of both the numerator and denominator and cancelled. (Factorization of expectations of independent variables is discussed in Mood, Graybill, and Boes 1974, 160.)

We now use properties of the log-normal distribution to solve for the expectations. If z is normally distributed with mean μ and variance σ_z^2, then $x = e^z$ has a log-normal distribution and $E(x) = e^{\mu+\sigma_z^2/2}$ (see Mood, Graybill, and Boes 1974, 117). Thus,

$$
\begin{aligned}
E(e_{st} - \eta e_{s0}) &= E[(1 - \eta)e_{s0}] = 0, \\
\mathrm{Var}(e_{st} - \eta e_{s0}) &= (1 - 2\eta\rho_t + \eta^2)\sigma^2, \\
\mathrm{Var}[(1 - \eta)e_{s0}] &= (1 - \eta)^2\sigma^2, \\
\frac{E[\exp(e_{st} - \eta e_{s0})]}{E\{\exp[(1 - \eta)e_{s0}]\}} &= \frac{\exp[(1 - 2\eta\rho_t + \eta^2)\sigma^2/2]}{\exp[(1 - \eta)^2\sigma^2/2]} \\
&= \exp[\eta(1 - \rho_t)\sigma^2].
\end{aligned}
\tag{A2}
$$

Derivation of Equation (17)

By the same argument given above, equation (16) converges in probability to $E(p_{sh}q_{sh}p_{st}/p_{sl})/E(p_{sh}q_{sh})$. Under the model of equations (10) and (11), this ratio is

$$
\begin{aligned}
\frac{E(p_{sh}q_{sh}p_{st}/p_{sl})}{E(p_{sh}q_{sh})} &= \frac{E\left\{\exp\left[\begin{array}{c}(1 - \eta)(\pi_h + u_s + e_{sh}) \\ + \pi_t + e_{st} - \pi_l - e_{sl} + \delta_h + v_s + \omega_{sh}\end{array}\right]\right\}}{E\{\exp[(1 - \eta)(\pi_h + u_s + e_{sh}) + \delta_h + v_s + \omega_{sh}]\}} \\
&= \exp(\pi_t - \pi_l)\frac{E\{\exp[(1 - \eta)e_{sh} + e_{st} - e_{sl}]\}}{E\{\exp[(1 - \eta)e_{sh}]\}},
\end{aligned}
\tag{A3}
$$

where independence again permits $E\{\exp[(1 - \eta)u_s + \delta_h + v_s + \omega_{sh}]\}$ to be factored out of the numerator and denominator and cancelled.

Using the properties of the log-normal distribution to solve the expectations, we obtain

$$
\begin{aligned}
E[(1 - \eta)e_{sh} + e_{st} - e_{sl}] &= E[(1 - \eta)e_{sh}] = 0, \\
\mathrm{Var}[(1 - \eta)e_{sh} + e_{st} - e_{sl}] &= [2 + (1 - \eta)^2 + 2(1 - \eta)(\rho_{t-h} - \rho_{l-h}) - 2\rho_{t-l}]\sigma^2, \\
\mathrm{Var}[(1 - \eta)e_{sh}] &= (1 - \eta)^2\sigma^2, \\
\frac{E\left\{\exp\left[\begin{array}{c}(1 - \eta)e_{sh} \\ + e_{st} - e_{sl}\end{array}\right]\right\}}{E\{\exp[1 - \eta)e_{sh}]\}} &= \frac{\exp\left\{\left[\begin{array}{c}2 + (1 - \eta)^2 \\ + 2(1 - \eta)(\rho_{t-h} - \rho_{l-h}) - 2\rho_{t-l}\end{array}\right]\sigma^2/2\right\}}{\exp[(1 - \eta)^2\sigma^2/2]} \\
&= \exp\{[1 - \rho_{t-l} + (1 - \eta)(\rho_{t-h} - \rho_{l-h})]\sigma^2\}.
\end{aligned}
\tag{A4}
$$

Derivation of Equation (19)

To derive equation (19), we begin with the following result about the expected value of the product of a normal and a log-normal random variable.

LEMMA. *Suppose x and y are bivariate normal with means μ_x and μ_y, variances σ_x^2 and σ_y^2, and correlation* ρ. *Let $z = xe^y$. Then $E(z) = (\sigma_x\sigma_y\,\rho + \mu_x)$ $e^{\mu_y+\sigma_y^2/2}$ where $\sigma_x\sigma_y\rho$ is the covariance of x and y.*

PROOF. Write out the expectation:

$$E(xe^y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{xe^y}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \times \exp\left\{-\frac{1}{2(1-\rho^2)}\right.$$

$$\left.\times\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\} dx\,dy.$$

Make the simplifying transformations $u = (x - \mu_x)/\sigma_x$ and $v = (y - \mu_y)/\sigma_y$, with Jacobian of transformation $|J| = \sigma_x\sigma_y$:

$$= \frac{e^{\mu_y}}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\sigma_x u + \mu_x)$$

$$\times \exp\left[\sigma_y v - \frac{1}{2(1-\rho^2)}(u^2 - 2\rho uv + v^2)\right] dv\,du.$$

Next, complete the square on u in the exponent:

$$= \frac{e^{\mu_y}}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\sigma_x u + \mu_x)$$

$$\times \exp\left\{\sigma_y v - \frac{1}{2(1-\rho^2)}[(u-\rho v)^2 + (1-\rho^2)v^2]\right\} dv\,du,$$

then substitute $w = (u - \rho v)/\sqrt{1-\rho^2}$, with Jacobian $\sqrt{1-\rho^2}$.

$$= \frac{e^{\mu_y}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-v^2/2+\sigma_y v}\left(\int_{-\infty}^{\infty} \frac{\sigma_x\sqrt{1-\rho^2}\,w + \sigma_x\rho v + \mu_x}{\sqrt{2\pi}}\, e^{-w^2/2}\,dw\right) dv.$$

Since $\int_{-\infty}^{\infty} e^{-x^2/2}/\sqrt{2\pi}\,dx = 1$, and $\int_{-\infty}^{\infty} xe^{-x^2/2}/\sqrt{2\pi}\,dx = 0$, integration of the expression in parentheses produces

$$= \frac{e^{\mu_y}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma_x\rho v + \mu_x)e^{-v^2/2+\sigma_y v}\,dv.$$

Complete the square on v in the exponential, substitute $t = v - \sigma_y$, and solve:

$$= e^{\mu_y+\sigma_y^2/2} \int_{-\infty}^{\infty} \frac{[\sigma_x\rho(t + \sigma_y) + \mu_x]}{\sqrt{2\pi}} e^{-t^2/2}\, dt$$
$$= (\sigma_x\sigma_y\rho + \mu_x)e^{\mu_y+\sigma_y^2/2}.$$

To continue the derivation of equation (19), by the same argument given above, equation (18) converges in probability to $\exp\{E[p_{sh}q_{sh} \log(p_{st}/p_{sl})]/E(p_{sh}q_{sh})\}$. Under the model of equations (10) and (11), the ratio of expectations is

$$\text{(A5)} \quad \frac{E[p_{sh}q_{sh}\log(p_{st}/p_{sl})]}{E(p_{sh}q_{sh})} = \frac{E\left\{\begin{matrix}(\pi_t - \pi_l + e_{st} - e_{sl}) \times \\ \exp[(1-\eta)(\pi_h + u_s + e_{sh}) + \delta_h + \nu_s + \omega_{sh}]\end{matrix}\right\}}{E\{\exp[(1-\eta)(\pi_h + u_s + e_{sh}) + \delta_h + \nu_s + \omega_{sh}]\}}$$
$$= (\pi_t - \pi_l)\frac{E\{(e_{st} - e_{sl})\exp[(1-\eta)e_{sh}]\}}{E\{\exp[(1-\eta)e_{sh}]\}},$$

where independence again permits $E\{\exp[(1 - \eta)u_s + \delta_h + \nu_s + \omega_{sh}]\}$ to be factored out and canceled.

To apply the above lemma, we need to know the means of $e_{sh} - e_{sh}$ and $(1 - \eta)e_{sh}$, the variance of $(1 - \eta)e_{sh}$, and the covariance of $e_{sh} - e_{sh}$ and $(1 - \eta)e_{sh}$. These are

$$\text{(A6)} \quad \begin{aligned} E(e_{st} - e_{sl}) &= E[(1-\eta)e_{sh}] = 0, \\ \mathrm{Var}[(1-\eta)e_{sh}] &= (1-\eta)^2\sigma^2, \\ \mathrm{Cov}[(e_{st} - e_{sl}), (1-\eta)e_{sh}] &= (1-\eta)(\rho_{t-h} - \rho_{l-h})\sigma^2. \end{aligned}$$

(The expression for the covariance follows from a formula for the covariance of a linear function of random variables in Mood, Graybill, and Boes 1974, 179.) Using the above lemma about the expectation of the product of normal and log-normal random variables,

$$\frac{E\{(e_{st} - e_{sl})\exp[(1-\eta)e_{sh}]\}}{E\{\exp[(1-\eta)e_{sh}]\}} = \frac{(1-\eta)(\rho_{t-h} - \rho_{l-h})\sigma^2\exp[(1-\eta)^2\sigma^2/2]}{\exp[(1-\eta)^2\sigma^2/2]}$$
$$= (1-\eta)(\rho_{t-h} - \rho_{l-h})\sigma^2,$$

from which equation (19) immediately follows.

References

Aizcorbe, Ana, and Patrick Jackman. 1993. Commodity substitution bias in Laspeyres indexes: Analysis using CPI source data for 1982–91. *Monthly Labor Review* 116 (December): 25–33.

Anglin, Paul M., and Michael R. Baye. 1987. Information, multiprice search, and cost-of-living index theory. *Journal of Political Economy* 95 (December): 1179–95.

Beran, Jan. 1992. Statistical methods for data with long-range dependence. *Statistical Science* 7 (4): 404–27.
Braithwait, Steve D. 1980. The substitution bias of the Laspeyres price index. *American Economic Review* 70 (March): 64–77.
Christensen, Laurits R., and Marilyn E. Manser. 1976. Cost of living indexes and price indexes for U.S. meat and produce, 1947–71. In *Household production and consumption,* ed. Nestor E. Terleckyj, 399–446. NBER Studies in Income and Wealth, vol. 40. New York: Columbia University Press.
Cochran, William G. 1977. *Sampling techniques.* 3d ed. New York: John Wiley and Sons.
Dalén, Jörgen. 1992. Computing elementary aggregates in the Swedish consumer price index. *Journal of Official Statistics* 8 (2): 129–47.
Fisher, Irving. [1927] 1967. *The making of index numbers: A study of their varieties, tests and reliability.* 3d ed. Reprint, New York: Augustus M. Kelly.
Forsyth, F. G., and R. F. Fowler. 1981. The theory and practice of chain price index numbers. *Journal of the Royal Statistical Society* ser. A, 144 (2): 224–46.
Friedman, Milton. 1992. Do old fallacies ever die? *Journal of Economic Literature* 30 (December): 2129–32.
Frisch, Ragnar. 1936. Annual survey of general economic theory: The problem of index numbers. *Econometrica* 4 (January): 1–38.
Manser, Marilyn, and Richard McDonald. 1988. An analysis of the substitution bias in measuring inflation, 1959–1985. *Econometrica* 56 (July): 909–30.
Marshall, Alfred. [1887] 1925. Remedies for fluctuations of general prices. *Contemporary Review* 51:355–75. Reprinted in *Memorials of Alfred Marshall,* ed. A. C. Pigou, chap. 8. London: Macmillan.
Mood, Alexander M., Franklin A. Graybill, and Duane C. Boes. 1974. *Introduction to the theory of statistics.* 3d ed. New York: McGraw-Hill.
Moulton, Brent R. 1993. Basic components of the CPI: Estimation of price changes. *Monthly Labor Review* 116, no. 12 (December): 13–24.
———. 1996. Estimation of elementary indexes of the Consumer Price Index. Bureau of Labor Statistics, Washington, D.C. Unpublished, May.
Pigou, Arthur C. 1932. *The economics of welfare.* 4th ed. London: Macmillan.
Pollak, Robert A. 1980. Group cost-of-living indexes. *American Economic Review* 70 (May): 273–8.
———. 1981. The social cost-of-living index. *Journal of Public Economics* 15 (June): 311–36.
Reinsdorf, Marshall B. 1993. The effect of outlet price differentials on the U.S. Consumer Price Index. In *Price measurements and their uses,* ed. Murray F. Foss, Marilyn E. Manser, and Allan H. Young, 227–60. NBER Studies in Income and Wealth, vol. 57. Chicago: University of Chicago Press.
———. 1994a. The effect of price dispersion on cost of living indexes. *International Economic Review* 35, no. 1 (February): 137–49.
———. 1994b. Price dispersion, seller substitution and the U.S. CPI. Working Paper no. 252, Bureau of Labor Statistics, Washington, D.C. April.
Szulc, Bohdan J. 1983. Linking price index numbers. In *Price level measurement: Proceedings from a conference sponsored by Statistics Canada,* ed. W. Erwin Diewert and Claude Montmarquette, 537–66. Ottawa: Ministry of Supply and Services.
———. 1989. Price indices below the basic aggregation level. In *Consumer price indices: An ILO manual,* ed. Ralph Turvey, 167–77. Geneva: International Labour Office.
Törnqvist, Leo. 1936. The Bank of Finland's consumption price index. *The Bank of Finland Monthly Bulletin* October, 27–34.

Turvey, Ralph, et al. 1989. *Consumer price indices: An ILO manual.* Geneva: International Labour Office.
U.S. Department of Labor. 1992. *BLS handbook of methods.* Bureau of Labor Statistics Bulletin 2414. Washington, D.C.: Government Printing Office.
White, Halbert. 1984. *Asymptotic theory for econometricians.* San Diego, Calif.: Academic Press.

Comment W. E. Diewert

Introduction

Though the problem might appear very simple, this is far from the case. Surprisingly little work appears to have been done on it.

A. G. Carruthers, D. J. Sellwood, and P. W. Ward (1980, 16)

There is an abundant literature, both theoretical and descriptive, on the computation of consumer price indexes above the basic aggregation level, but little is written about their derivation below that level. In this respect, the index makers resemble those chefs who only allow their dishes to be presented to patrons at a certain stage of preparation, without sharing how they have been mixed and simmered in the kitchen.

B. J. Szulc (1987, 11)

In an important paper, Marshall Reinsdorf (1993) used BLS data to compare the growth of average prices in the United States with corresponding official CPI growth rates. He found that the official index for food showed average annual increases during the 1980s of 4.2 percent while the weighted mean of average prices grew at only 2.1 percent. For gasoline, Reinsdorf found that average prices fell during the 1980s about 1 percent per year more than the official CPI components for gasoline. Thus it appeared that the CPI components for food and gas were biased upward by about 2 percent and 1 percent per year, respectively, during the 1980s.

Reinsdorf (1993, 246) attributed the above results to outlet-substitution bias; that is, consumers switched from traditional high-cost retailers to new discount stores in the case of food and to self-serve gas stations from full-service stations in the case of gasoline. The existing methodology used by statistical

W. E. Diewert is professor of economics at the University of British Columbia and a research associate of the National Bureau of Economic Research.

This research was supported by a Strategic Grant from the Social Sciences and Humanities Research Council of Canada. A longer version of this comment which reviews and extends the literature on elementary price indexes is available as Diewert (1995). The author thanks Paul Armknecht, John Astin, Bert Balk, John Bossons, Jörgen Dalén, Melvyn Fuss, Robert Gordon, Marta Haworth, Peter Hill, Brent Moulton, Alice Nakamura, Marshall Reinsdorf, Jacob Ryten, Alain Saglio, Bohdan Szulc, Don Sellwood, Jack Triplett, Ralph Turvey, and Keith Woolford for valuable discussions and Louise Hebert for careful typing of a difficult manuscript.

agencies in compiling price indexes does not pick up this shift of purchasers from high- to low-cost suppliers.[1]

The more recent paper under discussion by Reinsdorf and Moulton presents an alternative explanation for Reinsdorf's earlier results:[2] when the BLS moved to probability sampling of prices in 1978, the micro price quotations were aggregated using an index number formula that generates an upward bias. In the second section of this comment, I discuss index number formulas that are used to aggregate prices at the finest level of disaggregation, and I provide Irving Fisher's (1922, 383) intuitive explanation for the Reinsdorf-Moulton empirical results. In the third section, I briefly review the recent literature on sources of bias in consumer price indexes. The fourth section concludes with a number of recommenations to statistical agencies.

The Problem of Aggregating Price Quotes at the Lowest Level

Who ever heard, for instance, of Carli and of Dutot as authorities on the subject?
F. Y. Edgeworth (1901, 404) commenting on C. M. Walsh (1901)

In order to provide an intuitive explanation for the empirical results of Reinsdorf and Moulton, it is necessary to introduce a bit of notation and define a few index number formulas. I assume that the statistical agency is collecting price quotations on a commodity at the lowest level of aggregation where information on quantities purchased is not available.[3] Assume that the physical and economic characteristics of the good are homogeneous and that N price quotes on it are collected in periods 0 and 1 respectively. Denote the period t vector of price quotes as $\boldsymbol{p}^t \equiv (p_1^t, p_2^t, \ldots, p_N^t)$ for $t = 0, 1$. Define an *elementary price index* as a function of the $2N$ prices $(p_1^0, \ldots, p_N^0; p_1^1, \ldots, p_N^1) = (p^0; p^1)$. Examples of specific functional forms for elementary price indexes are

(1) $$P_{\mathrm{CA}}(p^0, p^1) \equiv \Sigma_{n=1}^{N}(1/N)(p_n^1/p_n^0);$$

(2) $$P_{\mathrm{JE}}(p^0, p^1) \equiv \Pi_{n=1}^{N}(\mathrm{p}_n^1/p_n^0)^{1/N};$$

(3) $$P_{\mathrm{DU}}(p^0, p^1) \equiv \Sigma_{n=1}^{N}(1/N)p_n^1/\Sigma_{i=1}^{N}(1/N)p_i^0.$$

P_{CA} is the arithmetic mean of the price ratios p_n^1/p_n^0 (first suggested by Carli

1. When an outlet supplying a price quote disappears and is replaced by a new outlet, the new outlet's price quote does not immediately replace the missing price quote. Usually, price quotes are obtained from the new outlet for at least two periods, and then a price ratio using only new-outlet prices is linked into the index at the end of the second period. Thus any absolute change in prices going from the old outlet to the new outlet is ignored.

2. See also Moulton (1993); Reinsdorf (1994a); and Armknecht, Moulton, and Stewart (1994).

3. Turvey (1989, chap. 3) and Dalén (1992) refer to this situation as computing elementary aggregates while Szulc (1987) refers to it as constructing a price index below the basic aggregation level. Additional references which deal with this situation are Forsyth (1978, 352–55); Carruthers, Sellwood, and Ward (1980); Forsyth and Fowler (1981, 241); and Balk (1994).

[1804] in 1764); P_{JE} is the geometric mean of the price ratios (first suggested by Jevons [1884] in 1863), and P_{DU} is the arithmetic mean of period 1 prices divided by the arithmetic mean of period 0 prices (first suggested by Dutot [1738]).[4]

Reinsdorf and Moulton point out that the starting point for the BLS method of aggregating elementary price quotes resembles the Carli price index P_{CA} defined by equation (1).[5] In actual BLS practice, a more complicated formula than equation (1) is used (which Reinsdorf and Moulton describe; see also Armknecht, Moulton, and Stewart 1994), but as a very rough approximation, we can say that the elementary components of the U.S. CPI are computed using equation (1).

Reinsdorf and Moulton used official U.S. BLS aggregation techniques to construct consumer price index components for June 1992–June 1993, and they compared these simulated components to corresponding indexes that aggregated the elementary-level price quotes using the geometric mean formula in equation (2). Omitting housing, they found that their simulated "official" index exceeded the corresponding geometric mean index by about 0.5 percent for the year.[6]

Of course, if precisely equations (1) and (2) were being compared, we would always have

$$P_{CA}(p^0,p^1) \geq P_{JE}(p^0,p^1), \tag{4}$$

since an arithmetic mean is always equal to or greater than the corresponding geometric mean.[7] Moreover, the less proportional that prices are in the two periods (i.e., the more variable are prices), the greater the inequality in equation (4) will be.

It is likely that the inequality in equation (4) explains a large portion of the empirical results in Reinsdorf and Moulton's paper. However, at this stage, it is not clear why we should prefer the geometric average of the price relatives to the corresponding arithmetic average.

4. Unweighted price indexes of the forms in equations (1)–(3) were among the first to appear in the index number literature; see Walsh (1901, 553–58), Fisher (1922, 458–520), and Diewert (1993) for references to the early history of price indexes. Pigou (1924, 59), Frisch (1936), Szulc (1987, 13), and Dalén (1992, 139) refer to equation (1) as the Sauerbeck (1895) index.

5. Reinsdorf and Moulton note that equation (1) is called the unbiased and efficient Horvitz-Thompson estimator in the statistical literature, provided that the outlets in the statistical agency's sample were selected with probabilities proportional to their sales to consumers in the base period (period 0).

6. Armknecht, Moulton, and Stewart (1994) found that the U.S. "owners' equivalent rent" component of the U.S. CPI exceeded the corresponding geometric mean index by about 0.5 percent per year over the period March 1992–June 1994. They attributed this difference to the use of equation (1) as the elementary price index formula rather than equation (2). This upward "bias" in the owners' equivalent rent component of the CPI is likely to be present since the current implicit rent formula was introduced in January 1987.

7. Price index theorists who have used or derived the inequality in equation (4) include Walsh (1901, 517), Fisher (1922, 375–76), Szulc (1987, 12), and Dalén (1992, 142).

An explanation for our preference can be found in the work of Dalén (1992) who adapted the traditional bilateral test approach to index number theory (see Fisher 1911, 1922 and Eichhorn and Voeller 1976) to the present situation where information on quantities is missing. Dalén (138) suggested that a reasonable functional form P for an elementary price index should satisfy the following *time reversal test:*

$$(5) \qquad P(p^0, p^1)P(p^1, p^0) = 1;$$

that is, if prices in period 2 are identical to prices in period 0, then the price change going from period 0 to 1 should be exactly offset by the price change going from period 1 to 2.[8] It can be verified that the geometric mean price index P_{JE} defined by equation (2) satisifies equation (5) but the arithmetic mean price index P_{CA} defined by equation (1) will be biased upward, that is,

$$(6) \qquad P_{CA}(p^0, p^1)\, P_{CA}(p^1, p^0) \geq 1,$$

with a strict inequality if p^0 is not proportional to p^1.[9] Fisher (1922, 66, 383) seems to have been the first to establish the upward bias of the Carli price index P_{CA}[10] and he made the following observations on its use: "In fields other than index numbers it is often the best form of average to use. But we shall see that the simple arithmetic average produces one of the very worst of index numbers. And if this book has no other effect than to lead to the total abandonment of the simple arithmetic type of index number, it will have served a useful purpose" (29–30).

Unfortunately, Fisher's warning about the use of the arithmetic mean of price ratios as a functional form for an elementary price index was forgotten, not only by the compilers of the U.S. CPI, as the work of Reinsdorf and Moulton shows, but also by the compilers of the Swedish CPI for a short period in 1990, as was noted by Dalén (1992; 139).[11] Thus in view of its upward bias, the use of the Carli price index P_{CA} for aggregating elementary price quotes is definitely not recommended; the use of the geometric index P_{JE} defined by equa-

8. Fisher (1922, 82) credited the time reversal test to the Dutch economist Pierson (1896, 128). Letting P denote the index number formula, Pierson's test was $P(1_N, p_1, p_2, \ldots, p_N) = P(p_1^{-1}, p_2^{-1}, \ldots, p_N^{-1}, 1_N)$ where 1_N is a vector of ones. This can be interpreted as an invariance to changes in the units-of-measurement test. However, Pierson (130) later gave a simple example which showed that the Carli price index did not satisfy the time reversal property. Walsh (1901, 389) and Fisher (1911, 401) gave the first formal statements of the time reversal test.

9. Note that $1/P_{CA}(p^1, p^0)$ is the harmonic mean of the price ratios $p_1^1/p_1^0, \ldots, p_N^1/p_N^0$. The inequality in equation (6) now follows from the fact that the arithmetic mean of N positive numbers is always equal to or greater than the corresponding harmonic mean; see Walsh (1901, 517) and Fisher (1922, 383–84).

10. See also Pierson (1896, 130), Pigou (1924, 59 and 70), Szulc (1987, 12), and Dalén (1992, 139).

11. This bias problem is probably much more widespread; e.g., Allen (1975, 92) and Carruthers, Sellwood, and Ward (1980, 15) mentioned that the U.K. retail price index used the Carli formula at the elementary level, as well as the Dutot formula. Woolford (1994) reported that the Australian Bureau of Statistics also uses the Carli formula. Flux (1907, 619) reported that the early U.S. Bureau of Labor price index was a Sauerbeck (or Carli) index.

tion (2) or the average price index P_{DU} defined by equation (3) is definitely preferable since they both satisfy the time reversal test in equation (5).

Sources of Bias in Consumer Price Indexes

Retail markets furnish many examples of the Schumpeterian process of "creative destruction" in which more efficient producers enter and displace less efficient incumbents. The displacement of various classes of small, independent retailers by large mail order supply houses, department stores and chain grocery stores furnish historical examples of this. Recent times have seen phenomenal growth of a variety of large discount chains such as Wal-mart, Home Depot, Staples and Food Lion, as well as various "warehouse" style food stores and wholesale clubs.

M. Reinsdorf (1994b, 18)

Numerical computation of alternative methods based on detailed firm data on individual prices and quantities where new goods are carefully distinguished would cast light on the size of the new good bias.

W. E. Diewert (1993, 63)

Before we can discuss sources of bias in the computation of consumer price indexes, it is necessary to note that "bias" is a relative concept. Thus when we speak of bias, we have in mind some specific conceptual framework or purpose for the price index and if we had complete information, this underlying "truth" could be measured and "bias" would be relative to this "true index."

Economists and statisticians have been debating the question of the appropriate conceptual basis for a price index for over a hundred years.[12] The conceptual framework that I shall adopt in order to discuss bias is the COL framework due originally to Konüs (1939). More specifically, I adopt Pollak's (1981, 328) social COL index as the underlying "correct" concept.[13] This concept assumes utility maximizing (or expenditure minimizing) behavior on the part of consumers and thus is open to the criticism that it is unrealistic. However, as Pierson (1895, 332) observed one hundred years ago, consumers do purchase less in response to higher prices; that is, substitution effects do exist. The existing economic theory of COL indexes can be viewed as a way of incorporating these substitution effects into the measurement of price change (as opposed to the traditional statistical agency fixed-basket approach[14] which holds quantities fixed as prices change).

12. The debate started with Edgeworth (1888, 347), as the following quotation indicates: "The answer to the question what is the *Mean* of a given set of magnitudes cannot in general be found, unless there is given also the object for the sake of which a mean value is required." Other papers discussing different purposes and alternative conceptual frameworks include Edgeworth (1901, 409; 1923, 343–45; 1925), Flux (1907, 620), Bowley (1919, 345–53), March (1921), Mudgett (1929, 249), Ferger (1936), Mills et al. (1943, 398), Triplett (1983), Turvey (1989, 9–27), and Sellwood (1994).

13. This concept excludes the newer economic approaches to COL indexes that incorporate consumer search; see Anglin and Baye (1987) and Reinsdorf (1993, 1994a).

14. This traditional Laspeyres approach to measuring price change is comprehensively discussed in Turvey (1989). For earlier discussions, see Flux (1907, 621), Bowley (1919, 347), and Mills et al. (1943).

Instead of using the economic theory of the consumer as the theoretical basis for the construction of price indexes, it is possible to use instead a producer-theory approach to the measurement of price change; see Court and Lewis (1942–43); Fisher and Shell (1972); Samuelson and Swamy (1974); Archibald (1977); and Diewert (1983, 1054–77).[15] I will not pursue this approach here.

Once a theoretically ideal price index has been chosen, bias can be defined as a systematic difference between an actual statistical agency index and the theoretically ideal index. Instead of the term "bias," Fixler (1993, 7) and other BLS economists use the term "effect." Since most academic economists use the term "bias," I will follow in this tradition.[16]

In addition to the *elementary index functional form bias* considered in the previous section, I shall follow the examples of Gordon (1993) and Fixler (1993) and consider commodity-substitution bias, outlet-substitution bias, linking bias and new-goods bias.

The Laspeyres fixed-basket price index suffers from *commodity-substitution bias;* that is, it is biased upward compared to a COL index because it ignores changes in quantities demanded that are induced by changes in relative prices. Estimates of the size of this bias (at levels of aggregation above the elementary level) can be obtained by comparing statistical agencies' Laspeyres-type indexes with superlative index numbers such as the Fisher-Ideal index P_F^* defined as the geometric mean of the Paasche and Laspeyres indexes. Superlative indexes provide good approximations to the unobservable COL indexes.[17] Using this methodology, Manser and McDonald (1988), using 101 categories of goods and services, and Aizcorbe and Jackman (1993), using 207 categories in forty-four U.S. locations, found an average substitution bias in the U.S. CPI of about 0.2 percent per year. Using the same methodology, Généreux (1983) found the same substitution bias in the Canadian CPI over the years 1957–78. Using a different methodology, Balk (1990, 82) obtained estimates for the substitution bias in the Dutch CPI in the 0.2–0.3 percent per year range using 106 commodity groups over the years 1952–81.[18]

In the first section, I defined *outlet-substitution bias* in the context of disappearing high-cost outlets. I now want to broaden the above preliminary defini-

15. Diewert (1983, 1051–52) also compared the consumer- and producer-theory approaches.

16. Fisher (1922, 86) called an index number formula "erratic" if it did not satisfy the time reversal test and "biased" if it were "subject to a foreseeable tendency to err in *one particular direction*." Thus, using Fisher's terminology, the arithmetic and harmonic elementary price indexes, P_{CA} and P_H, are *biased,* while the Laspeyres price index, $P_L^*(p^0,p^1,q^0,q^1) \equiv p^1 \cdot q^0/p^0 \cdot q^0$, is merely *erratic*. Note that Lovitt (1928, 11) seems to have been the first to show that P_L^* was "erratic" and not "biased" in the sense of Fisher.

17. See Diewert (1976, 1978). Hill (1988, 134) assumed that superlative price indexes are essentially weighted averages of price relatives which have quantity or expenditure weights that treat the two periods under consideration in a symmetric manner.

18. A topic closely related to substitution bias is the sensitivity of the Laspeyres index to the choice of the base year or to the choice of expenditure weights for the price relatives; see Hogg (1931, 56), Mudgett (1933, 30), Saulnier (1990), Schmidt (1993), and Dalén (1994).

tion to encompass the possibility that consumers may shift their purchases from high-cost to low-cost outlets over time. Thus instead of calculating outlet-specific unit values for a commodity, a unit value could be calculated over all outlets in the market area. The difference between this market-area unit-value price relative and the corresponding Laspeyres component for the commodity in the official CPI can be defined as outlet-substitution bias.[19] This definition of outlet-substitution bias assumes that commodities should not be distinguished by their point of purchase; that is, a particular make of a video camera yields the same utility to a consumer whether it is bought in Dan's Discount Den or Regal Imports Boutique. This assumption may not be appropriate in other situations.[20] Turning to empirical evidence on the size of the outlet-substitution bias, in his direct statistical method, Reinsdorf (1993, 239–40) found that the outlet-substitution bias in the "food at home" and "motor fuel" components of the U.S. CPI was about 0.25 percent per year during the 1980s (although he regarded this as an upper bound due to possible quality differences). Saglio (1994), using Nielsen data for 915 French outlets over the years 1988–90, found that the outlet-substitution bias for milk chocolate bars averaged 0.8 percent per years; that is, the market unit value for chocolate bars of the same size and brand averaged 0.8 percent per year lower than the corresponding Laspeyres index which treated chocolate bars of the same size and brand in each outlet as separate commodities. Saglio (1994), using INSEE (Institut National de la Statistique et des Études Économiques) data on twenty-nine food groups over twelve years, also found an outlet-substitution bias of approximately 0.4 percent per year below the corresponding Laspeyres price index.

The outlet-substitution bias is formallly identical to what might be termed the *linking bias,* that is, a new good appears which is more efficient in some dimension than an existing good. After two or more periods, the statistical agency places a price relative for the new good into the relevant elementary price index, but the absolute decline in price going from the old to the new variety is never reflected in the relevant elementary price index. This source of bias was recognized by Griliches and by Gordon (1981, 130–33; 1990) as the following quotations indicate: "By and large they [statistical agencies] do not make such quality adjustments. Instead, the new product is 'linked in' at its introductory (or subsequent) price with the price indices left unchanged" (Griliches 1979, 97); "An even more dramatic case largely involving a producer durable involved the supplanting of the old rotary electric calculator by the electronic calculator; all of us can purchase for $10 or so a calculator that can perform all the functions (in a fraction of the elapsed time) of an old 1970-

19. This definition of outlet-substitution bias coincides with Reinsdorf's (1993, 228) original definition and includes both of Fixler's (1993, 7) seller- and outlet-substitution biases. It also corresponds to Saglio's (1994) point-of-purchase effect.

20. This ambiguity creates difficulties for statistical agencies; i.e., the decision whether to aggregate over outlets in a market area or not is clearly a matter of judgment.

vintage $1000 rotary electric calculator. Yet in the U.S. the electronic calculator was treated as a new product, and the decline in price from the obsolete rotary electric model to the early models of the electronic calculator was 'linked out' in the official indexes" (Gordon 1993, 239).

A more appropriate treatment of the above situation would be to calculate an average price or unit value per the relevant characteristic over the old and repackaged goods. A similar bias was recognized by Griliches and Cockburn (1994) in the context of generic drugs which are chemically identical to brand-name drugs (it should be noted that the BLS changed its procedures in January 1995 to fix this problem). An analogous bias in the statistical agency treatment of illumination was pointed out by Nordhaus (chap. 1 in this volume). These last two papers obtain very large linking biases.[21]

The *new-goods bias* results from the inability of bilateral price indexes to take into account the fact that the number of commodities from which consumers can choose is growing rapidly over time.[22] Hill makes the following comment on this situation: "In general, it may be concluded that in the real world, price indices which are inevitably restricted to commodities found in both situations will fail to capture the improvement of welfare associated with an enlargement of the set of consumption possibilities. The benefits brought by the introduction of new goods are not generally taken into account in price indices in the period in which the goods first make their appearance" (1988, 138).

Diewert (1980, 498–505; 1987, 779; 1993, 59–63), following Marshall (1887, 373) and Hicks (1940, 114), discussed the new-goods bias and suggested along with Griliches (1979, 97) and Gordon (1981, 130) that this bias could be substantially reduced by simply introducing new goods into the pricing basket in a timely fashion (this would not eliminate the bias in the period when the good makes its first appearance). Triplett (1993, 200) termed the subset of the new-goods bias caused by delays in introducing new products into an index as the *new-introductions bias*.[23] Turning now to empirical estimates of the new-goods bias, Gordon (1990) estimated that the U.S. consumer durables price index had a new-goods or quality-change bias of 1.5 percent per year over the period 1947–83. Berndt, Griliches, and Rosett (1993) provided evi-

21. Again, this source of bias creates problems for statistical agencies; i.e., when should a new product be treated as a genuinely new good rather than as a superficially repackaged old product? It should also be noted that linking bias could go in the opposite direction if firms simply repackage their products to disguise price increases.

22. Actually, what is relevant is the number of commodities that are available in the consumer's market area. Thus the growth of cities and urbanization leads to more specialized goods and services being offered by producers and hence will lead to a growth in the number of commodities that are effectively available to the consumer. Transportation and communication improvements also lead to larger choice sets, a point already noticed by Marshall (1887, 373–74).

23. Mudgett (1933, 32) noted that in 1930, the BLS had not yet added such important items of expenditure to its basket as automobile expenditures, meals outside the home, and life insurance. Gordon (1993) noted that automobiles entered the U.S. CPI in 1940, penicillin in 1951 after it had experienced a 99 percent decline from its initial price, and the pocket calculator in 1978 after it had declined in price about 90 percent since 1970. Mudgett (1929, 250) also noted that only forty commodities were comparable between 1870 and 1920 out of five hundred commodities whose prices were collected by the BLS in 1920.

dence that the BLS did not sample the prices of new drug products in a sufficiently timely fashion. They found that from January 1984 through December 1989, the BLS producer price index for prescription pharmaceutical preparations (drugs) grew at a rate of 3 percent per year higher than a superlative price index that used the monthly price and quantity sales data for 2,090 drug products sold by four major pharmaceutical manufacturers in the United States, accounting for about 29 percent of total domestic industry sales in 1989. Thus they found a combined drug-substitution and new-introductions bias of about 3 percent per year. Hausman (chap. 5 in this volume) used Nielsen scanner data from January 1990 to August 1992 on cereal consumption for seven major metropolitan areas in the United States. He used econometric techniques to estimate consumer preferences over cereals and thus he was able to estimate the Hicksian (1940, 114) reservation prices that would cause consumers to demand zero units of a new cereal. His conclusion was that an overall price index for cereals, which excluded the effects of new brands, would overstate the true COL subindex for cereals by about 25 percent over a ten-year period.[24] Finally, Trajtenberg (1990) attempted to measure reservation prices for computerized tomography (CT) scanners over the decade 1973–82. His nominal price index went from 100 to 259 but his quality-adjusted price index went from 100 to 7, implying a 55 percent drop in prices per year on average.

Summarizing the empirical evidence reviewed in this section and the previous one, we see that it is likely that in recent years, a typical official consumer price index has a 0.2 percent per year commodity-substitution bias, a 0.25 percent per year outlet-substitution bias, a linking bias of perhaps 0.1 percent per year, and a new-goods bias of at least 0.25 percent per year; that is, an upward bias of at least 0.8 percent per year. If the statistical agency is also making use of a biased elementary price index formula, this will add an additional upward bias to the official index. The reader will note that all five sources of bias were regarded as being additive, an assumption which is probably approximately correct.[25]

I conclude this section with a detailed discussion of the possible biases in the U.S. CPI. Marshall Reinsdorf and Brent Moulton have provided important empirical evidence of upward bias in the U.S. CPI due to an inappropriate choice of the functional form used to aggregate price quotations at the lowest level of aggregation. Reinsdorf and Moulton found that their geometric mean index (which used the elementary price index P_{JE} defined by equation (2) at the lowest level of aggregation) grew by 2.48 percent from June 1992 to June 1993, compared to a simulated U.S. consumer price index growth rate of 2.95 percent. Their simulations excluded housing and hence covered 70.3 percent of

24. This bias is the "pure" new-goods bias (the bias that occurs in the period when the new good is introduced) as opposed to the new-introductions bias (the bias that occurs in the second and subsequent periods after the good is introduced). Hausman found that his estimated reservation prices were approximately double the first-appearance prices of the new cereals.

25. Sellwood (1994) discussed the question of additivity. He also noted that estimates of bias have standard errors attached to them.

the U.S. CPI universe. Thus their simulated U.S. consumer price index (which largely uses the Carli-Sauerbeck price index P_{CA} defined by equation [1] at the elementary level) appears to have an upward bias of about 0.5 percent per year. Furthermore, Armknecht, Moulton, and Stewart (1994) noted that since 1987, the owner's implicit rent component of the CPI has used a Carli elementary price index, which has led to a 0.5 percent per year upward bias in that component. Thus the choice of index number formula at the elementary level is not a trivial matter.

Reinsdorf (1993, 242–47) compared the behavior of official U.S. rates of inflation for food and gasoline with corresponding rates obtained using average prices; that is, he compared CPI rates of inflation for food and gas with those obtained by using the elementary price index P_{DU} defined by equation (3). Over the 1980s, he found that means of the U.S. CPI food indexes weighted according to their importance in the CPI showed an average annual increase of 4.2 percent, while the corresponding weighted mean of the average prices grew at a rate of 2.1 percent per year. For gasoline, he found that average prices fell faster than the corresponding CPI prices at about 1 percent per year during the 1980s. Reinsdorf (242) attributed these results to outlet-substitution bias but it now seems clear that some of this upward bias in food and gas was due to the inappropriate method used by the BLS to aggregate price quotes at the elementary level. However, it is also clear that not all of Reinsdorf's results can be explained away as being elementary-level functional form bias: a substantial portion of the bias that he found must be outlet-substitution bias.

The results of Reinsdorf and Reinsdorf and Moulton suggest that outlet-substitution bias in the U.S. CPI as a whole was somewhere between 0.1 and 0.5 percent per year in the 1980s and the elementary functional form bias was somewhere between 0.35 and 0.5 percent per year in the 1990s. In addition to the above two sources of bias, we have commodity-substitution bias at levels above the elementary level, linking bias, and new-goods bias. These three sources of bias probably add an additional 0.3 to 0.7 percent per year of upward bias to traditional fixed-basket-type indexes. Adding up all of these sources of bias for the U.S. CPI leads to a total upward bias in the region of 0.75 to 1.7 percent per year in the 1990s. This is a substantial bias.[26]

Recommendations and Conclusions

> [E]very person in the room would have realized after hearing his Paper that the measurement of the cost of living was by no means a simple conception. Nobody would expect that a difficult question of engineering or a nice point of art could be put in the Press and explained in words of one syllable and in a single sentence.
>
> A. L. Bowley (1919, 371) commentary on his own paper

26. Similar sources of bias apply to the producer price index; see Gordon (1990, 1993) and Triplett (1993). Recent surveys of sources of bias in the CPI are Gordon (1993), Crawford (1993), and Wynne and Sigalla (1994).

> Would it not be well if statisticians and economists should again come together and decide authoritatively on the proper method of constructing index-numbers?
>
> C. M. Walsh (1921, 138)

A number of recommendations seem to follow from the empirical work of Reinsdorf and Moulton:

1. Statistical agencies should follow the emphatic advice of Irving Fisher (1922, 29–30) and avoid the use of the Carli arithmetic mean of price relatives formula in equation (1) to form elementary price aggregates.
2. If information on quantities is not available at the elementary or basic level, either the geometric price index in equation (2) advocated by Jevons (1884) or the average price index in equation (3) suggested by Dutot (1738) should be used.
3. At the level of the individual outlet, the best elementary average price for a homogeneous commodity would seem to be its unit value: the value of units sold during the sample period divided by the total quantity sold. If outlet unit values are available, then in aggregating over outlets there is no need to restrict ourselves to using the Jevons or Dutot formulas to construct elementary prices. From the viewpoint of economic theory, it seems preferable to use the Fisher-Ideal price index in this second stage of elementary aggregation.
4. Values and quantities should be sampled rather than just prices. Sampling values and quantities will greatly reduce the new-introductions bias.
5. Statistical agencies should consider either purchasing electronic point-of-sale data from firms currently processing these data, or the agencies should set up divisions which would compete in this area.
6. Recent economic history will have to be rewritten in view of the substantial outlet-substitution and elementary price index biases that Reinsdorf and Moulton have uncovered in U.S. price indexes. Since the United States is so large in the world economy, world inflation was lower in the 1980s than was officially recorded and world output growth (and hence productivity growth) was higher. It is very likely that many of the sources of bias in price indexes documented for the U.S. economy are also applicable to other economies.

References

Aizcorbe, A. M., and P. C. Jackman. 1993. The commodity substitution effect in CPI data, 1982–91. *Monthly Labor Review* 116 (December): 25–33.

Allen, R. G. D. 1975. *Index numbers in theory and practice.* London: Macmillan.

Anglin, P. M., and M. R. Baye. 1987. Information, multiprice search, and cost-of-living index theory. *Journal of Political Economy* 95:1179–95.

Archibald, R. B. 1977. On the theory of industrial price measurement: Output price indexes. *Annals of Economic and Social Measurement* 6:57–72.

Armknecht, P. A., B. R. Moulton, and K. J. Stewart. 1994. Improvements to the food at home, shelter and prescription drug indexes in the U.S. Consumer Price Index. Paper presented at the International Conference on Price Indices, 31 October–2 November, at Statistics Canada, Ottawa, Ontario.

Balk, B. M. 1990. On calculating cost-of-living index numbers for arbitrary income levels. *Econometrica* 58:75–92.
———. 1994. On the first step in the calculation of a consumer price index. Paper presented at the International Conference on Price Indices, 31 October–2 November, at Statistics Canada, Ottawa, Ontario.
Berndt, E. R., Z. Griliches, and J. G. Rosett. 1993. Auditing the producer price index: Micro evidence from prescription pharmaceutical preparations. *Journal of Business and Economic Statistics* 11:251–64.
Bowley, A. L. 1919. The measurement of changes in the cost of living. *Journal of the Royal Statistical Society* 82:343–72.
Carli, G. R. [1764] 1804. Del valore e della proporzione de' metalli monetati. Reprinted in *Scrittori classici italiani di economia politica.* Volume 13, 297–366. Milan: G. G. Destefanis.
Carruthers, A. G., D. J. Sellwood, and P. W. Ward. 1980. Recent developments in the retail prices index. *The Statistician* 29:1–32.
Court, L. M., and H. E. Lewis. 1942–43. Production cost indices. *Review of Economic Studies* 10:28–42.
Crawford, A. 1993. Measurement biases in the Canadian CPI: A technical note. *Bank of Canada Review,* summer, 21–36.
Dalén, J. 1992. Computing elementary aggregates in the Swedish Consumer Price Index. *Journal of Official Statistics* 8:129–47.
———. 1994. Sensitivity analysis for harmonising European consumer price indices. Paper presented at the International Conference on Price Indices, 31 October–November 2, at Statistics Canada, Ottawa, Ontario.
Diewert, W. E. 1976. Exact and superlative index numbers. *Journal of Econometrics* 4:115–45.
———. 1978. Superlative index numbers and consistency in aggregation. *Econometrica* 46:883–900.
———. 1980. Aggregation problems in the measurement of capital. In *The measurement of capital,* ed. Dan Usher, 433–528. Chicago: University of Chicago Press.
———. 1983. The theory of the output price index and the measurement of real output change. In *Price level measurement,* ed. W. E. Diewert and C. Montmarquette, 1049–1113. Ottawa: Statistics Canada.
———. 1987. Index numbers. In *The new Palgrave: A dictionary of economics.* Vol. 2, ed. J. Eatwell, M. Milgate, and P. Newman, 767–80. London: Macmillan.
———. 1993. The early history of price index research. In *Essays in index number theory.* Vol. 1, ed. W. E. Diewert and A. O. Nakamura, 33–65. Amsterdam: North-Holland.
———. 1995. Axiomatic and economic approaches to elementary price indexes. Discussion Paper no. 95-01, University of British Columbia, Vancouver.
Dutot, C. 1738. *Réflexions politiques sur les finances et le commerce.* Vol. 1. La Haye: Les frères Vaillant et N. Prevost.
Edgeworth, F. Y. 1888. Some new methods of measuring variation in general prices. *Journal of the Royal Statistical Society* 51:346–368.
———. 1901. Mr. Walsh on the measurement of general exchange value. *Economic Journal* 11:404–16.
———. 1923. The doctrine of index-numbers according to Mr. Correa Walsh. *Economic Journal* 33:343–51.
———. 1925. The plurality of index-numbers. *Economic Journal* 35:379–88.
Eichhorn, W., and J. Voeller. 1976. *Theory of the price index.* Lecture Notes in Economics and Mathematical Systems, vol. 140. Berlin: Springer-Verlag.
Ferger, W. F. 1936. Distinctive concepts of price and purchasing-power index numbers. *Journal of the American Statistical Association* 31:258–72.

Fisher, F. M., and K. Shell. 1972. The pure theory of the national output deflator. In *The economic theory of price indices*, 49–113. New York: Academic Press.

Fisher, I. 1911. *The purchasing power of money.* London: Macmillan.

———. 1922. *The making of index numbers.* 1st ed. Boston: Houghton Mifflin.

Fixler, D. 1993. The Consumer Price Index: Underlying concepts and caveats. *Monthly Labor Review* 116 (December): 3–12.

Flux, A. W. 1907. Modes of constructing index-numbers. *Quarterly Journal of Economics* 21:613–31.

Forsyth, F. G. 1978. The practical construction of a chain price index number. *Journal of the Royal Statistical Society* ser. A, 141:348–358.

Forsyth, F. G., and R. F. Fowler. 1981. The theory and practice of chain price index numbers. *Journal of the Royal Statistical Society* ser. A, 144:224–46.

Frisch, R. 1936. Annual survey of general economic theory: The problem of index numbers. *Econometrica* 4:1–39

Généreux, P. 1983. Impact of the choice of formulae on the Canadian Consumer Price Index. In *Price level measurement*, ed. W. E. Diewert and C. Montmarquette, 489–511. Ottawa: Statistics Canada.

Gordon, R. J. 1981. The Consumer Price Index: Measuring inflation and causing it. *Public Interest* 63 (spring): 112–34.

———. 1990. *The measurement of durable goods prices.* Chicago: University of Chicago Press.

———. 1993. Measuring the aggregate price level: Implications for economic performance and policy. In *Price stabilization in the 1990s*, ed. K. Shigehara, 233–76. London: Macmillan.

Griliches, Z. 1979. Issues in assessing the contribution of research and development to productivity growth. *Bell Journal of Economics* 10 (spring): 92–116.

Griliches, Z., and I. Cockburn. 1994. Generics and new goods in pharmaceutical price indexes. *American Economic Review* 84:1213–32.

Hicks, J. R. 1940. The valuation of the social income. *Economica* 7:105–40.

Hill, P. 1988. Recent developments in index number theory and practice. *OECD Economic Studies* 10 (spring): 123–48.

Hogg, M. H. 1931. A distortion in the cost of living index. *Journal of the American Statistical Association* 26:52–57.

Jevons, W. S. 1884. A serious fall in the value of gold ascertained and its social effects set forth. In *Investigations in currency and finance*, 13–118. London: Macmillan.

Konüs, A. A. 1939. The problem of the true cost of living. *Econometrica* 7:10–29. First published (in Russian) in *Economic Bulletin of the Institute of Economic Conjuncture* (Moscow, 1924) 3, no. 9–10 (September-October): 64–71.

Lovitt, W. V. 1928. Index number bias. *Journal of the American Statistical Association* 23:10–17.

Manser, M. E., and R. J. McDonald. 1988. An analysis of substitution bias in measuring inflation, 1959–85. *Econometrica* 56:909–30.

March, L. 1921. Les modes de mesure du mouvement général des prix. *Metron* 1: 57–91.

Marshall, A. 1887. Remedies for fluctuations of general prices. *Contemporary Review* 51:355–75.

Mills, F. C., E. W. Bakke, R. Cox, M. G. Reid, T. W. Schultz, and S. Stratton. 1943. An appraisal of the U.S. Bureau of Labor Statistics cost of living index. *Journal of the American Statistical Association* 38:387–405.

Moulton, B. R. 1993. Basic components of the CPI: Estimation of price changes. *Monthly Labor Review* 116 (December): 13–24.

Mudgett, B. D. 1929. Some unsettled questions in the theory of index numbers. *Journal of the American Statistical Association* 24:249–51.

———. 1933. The problem of the representative budget in a cost of living index. *Journal of the American Statistical Association* 28:26–32.
Pierson, N. G. 1895. Index numbers and appreciation of gold. *Economic Journal* 5:329–35.
———. 1896. Further considerations on index-numbers. *Economic Journal* 6:127–31.
Pigou, A. C. 1924. *The economics of welfare.* 2d ed. London: Macmillan.
Pollak, R. A. 1981. The social cost-of-living index. *Journal of Public Economics* 15:311–36.
Reinsdorf, M. 1993. The effect of outlet price differentials in the U.S. Consumer Price Index. In *Price measurements and their uses,* ed. M. F. Foss, M. E. Manser, and A. H. Young, 227–54. NBER Studies in Income and Wealth, vol. 57. Chicago: University of Chicago Press.
———. 1994a. The effect of price dispersion on cost of living indexes. *International Economic Review* 35:137–49.
———. 1994b. Price dispersion, seller substitution and the U.S. CPI. Working Paper no. 252, Bureau of Labor Statistics, Washington, D.C. March.
Saglio, A. 1994. Comparative changes in average price and a price index: Two case studies. Paper presented at the International Conference on Price Indices. 31 October–2 November, at Statistics Canada, Ottawa, Ontario.
Samuelson, P. A., and J. Swamy. 1974. Invariant economic index numbers and canonical duality: Survey and synthesis. *American Economic Review* 64:566–93.
Sauerbeck, A. 1895. Index numbers of prices. *Economic Journal* 5:161–74.
Saulnier, M. 1990. Real gross domestic product: Sensitivity to the choice of base year. *Canadian Economic Observer* 3, no. 5 (May): 1–19.
Schmidt, M. L. 1993. Effects of updating the CPI market basket. *Monthly Labor Review* 116 (December): 59–62.
Sellwood, D. 1994. Constrained macro effects of the use of different methods at the basic level in the project to harmonize consumer price indices in the European Union. Paper presented at the International Conference on Price Indices, 31 October–2 November, at Statistics Canada, Ottawa, Ontario.
Szulc, B. J. 1987. Price indices below the basic aggregation level. *Bulletin of Labour Statistics* 2:9–16. Reprinted in *Consumer price indices: An ILO manual,* ed. R. Turvey, 167–78, Geneva: International Labour Office, 1989.
Trajtenberg, M. 1990. Product innovations, price indices and the (mis)measurement of economic performance. NBER Working Paper no. 3261. Cambridge, Mass.: National Bureau of Economic Research.
Triplett, J. E. 1983. Escalation measures: What is the answer? What is the question? In *Price level measurement,* ed. W. E. Diewert and C. Montmarquette, 457–82. Ottawa: Statistics Canada.
———. 1993. Comment. In *Price measurements and their uses,* ed. M. F. Foss, M. E. Manser, and A. H. Young, 197–206. NBER Studies in Income and Wealth, vol. 57. Chicago: University of Chicago Press.
Turvey, R. 1989. *Consumer price indices: An ILO manual.* Geneva: International Labour Office.
Walsh, C. M. 1901. *The measurement of general exchange value.* New York: Macmillan.
———. 1921. *The problem of estimation.* London: P. S. King and Son.
Woolford, K. 1994. A pragmatic approach to the selection of appropriate index formula. Paper presented at the International Conference on Price Indices, 31 October–2 November, at Statistics Canada, Ottawa, Ontario.
Wynne, M. A., and F. D. Sigalla. 1994. The Consumer Price Index. *Economic Review Federal Reserve Bank of Dallas,* second quarter, 1–22.