

U.S. DEPARTMENT OF LABOR
Bureau of Labor Statistics

OFFICE OF PRICES AND
LIVING CONDITIONS

Bias in the Consumer Price Index: What Is the Evidence?

Brent R. Moulton

Working Paper 294
October 1996

I thank Katharine Abraham, Dennis Fixler, John Greenlees, Bill Hawkes, Alan Krueger, Brad De Long, Jim MacDonald, Marshall Reinsdorf, Timothy Taylor, and Kimberly Zieschang for advice and comments. The opinions expressed in this paper are those of the author and do not represent an official policy of the Bureau of Labor Statistics or the views of other BLS staff members.

Economists have noted for decades that the U.S. Consumer Price Index (CPI) may tend to overstate changes in the cost of living. But bias in the CPI became an important policy issue only recently, when it became part of the debate over a balanced budget. Alan Greenspan (1995) triggered this controversy in January 1995 when he told the Budget Committees of Congress, "[T]he official CPI may currently be overstating the increase in the true cost of living by perhaps 1/2 percent to 1-1/2 percent per year. ... If the annual inflation adjustments to indexed programs and taxes were reduced by 1 percentage point ... the annual level of the deficit will be lower by about \$55 billion after five years." Subsequently, the Senate Finance Committee (1995) held a series of hearings on the Consumer Price Index, and then appointed an advisory commission of experts to investigate the bias.¹ The panel's interim report estimated that the CPI has had a bias of 1.5 percent per year during recent years, and projected a bias of 1 percent per year in the future (Advisory Commission, 1995).

The public debate over bias in the CPI was preceded by a flurry of new research, much of it conducted by economists at the Bureau of Labor Statistics and other statistical agencies around the world, as well as by academic economists. This research focused on identifying and measuring the biases in the consumer price index. Table 1 summarizes a range of estimates that have been presented. The range is clearly quite wide. The diversity of beliefs is probably even greater than indicated in this tabulation, because several of the experts testifying before the Committee declined to give an estimate (for example,

¹ The commission is chaired by Michael Boskin. The members are Ellen Dulberger, Robert Gordon, Zvi Griliches, and Dale Jorgenson.

Katharine Abraham, Janet Norwood, Robert Pollak, and Joel Popkin), and most of these individuals were critical of the larger estimates. In many cases, the same evidence has been interpreted in a number of different ways. Griliches (1995) said, "the Committee assumes that we already know that the CPI is overstated. But the scientific basis for this judgment is much weaker than the [Committee's] questions seem to imply ... The various 'guesstimates' in these sources are not independent of each other." For some of the sources of bias, the evidence is based on case studies of a small number of commodities. The differences between estimates seems to be largely determined by the willingness of experts to extrapolate from these case studies to estimates for broader categories of goods. The available research results may reflect a kind of selection effect, where researchers have tended to study the goods for which there is a strong presumption of possible bias--- computers, prescription drugs, etc.

Bias in the CPI impinges on most of the measurements economists make of economic growth and well-being. The CPI is often used directly to deflate nominal measures to "real" units, such as real wages. The CPI component indexes are also used by the Bureau of Economic Analysis to deflate personal consumption expenditures in constructing the national income accounts, so biases in the CPI could lead to biased measures of real growth and productivity.² The poverty thresholds are an example of another important economic indicator that is escalated by the CPI. As an example of the potential impact of CPI bias, consider the lower endpoint of the interval estimate given by

² Components of the producer price index and import and export price indexes are also used to deflate components of the national accounts, and presumably are affected by some of the same biases as the CPI.

the CPI Advisory Commission, an upward bias of 0.7 percent per year. Over a 25 year period, a bias of that magnitude would cause the growth of quantities deflated by the CPI, such as real wages, to be understated by 19 percent. Bias of this magnitude substantially affects how we assess the growth of our economy and the well-being of its members. Larger biases would have a more dramatic impact. The upper endpoint of the estimate given by the CPI Advisory Commission is 2.0 percent per year, which over a 25 year period would imply that growth of real quantities are understated by about 64 percent.

Teachers of economics may find that the discussion of bias in the CPI provides a useful case study in economic data and policy. Most of the biases discussed below are related to simple economic theories that are usually taught in a principles course. Discussion of the CPI may provide students with an interesting application of economic theory that has important policy implications, and also teaches them some of the difficulties associated with measurement of economic variables.

In this article I do not intend to provide another set of "guesstimates." However, for economists who are consumers of the information contained in the CPI -- or of the related data derived in part from using the CPI to make adjustments, such as the GDP accounts, productivity, and real wages -- it may be useful to describe how the CPI is constructed and then to review the recent evidence on bias in the CPI. I begin with a brief description of the CPI program's sampling and estimation methods. Then, I will review the evidence on each of the sources of bias, trying to indicate the nature of the evidence and its strengths and weaknesses.

An Overview of the CPI

The Consumer Price Index is designed to measure the change in the cost of purchasing a fixed market basket of goods and services representing average consumption patterns during some base period. An index based on a fixed, historical market basket is called a Laspeyres index.³ The actual index is constructed in two stages.

At the stage of index aggregation, the CPI is built up from 44 geographical strata (for example, the Denver metropolitan area) and 207 strata of items (for example, women's shoes), which combine to create 9,108 strata indexes. These strata indexes are aggregated into the all-items CPI using weights derived from the Consumer Expenditure Survey. The weights have historically been changed at roughly 10-year intervals. The current weights were introduced into the index at the beginning of 1987 and are based on expenditures during 1982-84.⁴ Current plans for CPI revision call for the next change of weights to occur at the beginning of 1998, using weights from 1993-95. The relative importance of major CPI categories are shown in Table 2.

³ In the CPI, the index $I_{T,0}$ measures the relative change between periods 0 and T in the cost of a fixed basket of goods from a base period B :

$$I_{T,0} = \frac{\sum_i Q_{Bi} P_{Ti}}{\sum_i Q_{Bi} P_{0i}},$$

where Q_{Bi} is the quantity of good i consumed during period B , and P_{0i} and P_{Ti} are the prices during periods 0 and T .

⁴ Thus, it would be accurate to call the CPI a "modified" Laspeyres index, modified because the market basket refers to a different, earlier, period (say 1982-84) than the period over which the prices are compared (say, 1987 to the present).

At the disaggregated stage, each of the strata indexes is estimated, using a representative sample of outlets and prices. To determine which actual outlets should be visited to determine prices, a Point-of-Purchase Survey is conducted, in which consumers are asked detailed questions about the outlets at which they purchased consumer goods and services. The allocation of consumer expenditures across outlets is estimated from the Point-of-Purchase Survey, which is conducted 1-2 years before a sample is selected.⁵ From the responses to this survey, a sample of outlets is selected, with probability of selection proportional to expenditures. Within each selected outlet, a BLS field economist then selects one or more specific varieties of items — again with probability proportional to sales — and the prices of these items will then be checked for the following five years. Through this process, the samples for about one-fifth of the indexes are replaced (or "rotated") each year.⁶ Detailed checklists are employed to ensure that precisely the same item is repriced each month. Any changes in the quality or characteristics of the sample item are noted and lead to the quality adjustment procedures discussed later in this paper. The overall sample sizes are 95,000 items from 22,000 retail outlets for commodities and services other than shelter, which are repriced monthly or bimonthly, and 35,000 rental units for measuring changes in rent and owners' equivalent rent, which are repriced every

⁵ When the sample in a particular city is replaced, prices are collected for both the old and new samples and the indexes from the old sample are linked to the indexes from the new sample. Linkage procedures are discussed later in this paper.

⁶ The selection of samples according to formal rules of probability, as well as the regular replacement of samples, were processes introduced in 1978 in response to recommendations made by the Stigler Commission, a committee of eminent economists and statisticians charged with reviewing government price statistics (Stigler, 1961).

six months.⁷ The rent sample is continuously augmented with a sample of new construction.

A number of practical problems arise in this process. One especially worth noting, because it relates to the conceptual issue of how the price index is developed, is that direct information on base-period quantities is generally not available; instead, the household surveys provide information on total base-period expenditures on categories of items. Direct information on the base-period price of the sample items is also generally not available, because the outlet sample of items is selected after the household expenditure data are collected and processed. One area where quantity information is available is for rent, where expenditure and price are the same, since the consumer is paying for the use of one housing unit. However, for commodities and services other than shelter, the index is calculated using information on base-period expenditures, adjusted by relative price changes since that time.⁸ At the disaggregated stage, the base-period expenditure weights refer to the Point-of-Purchase Survey period.

⁷ A sample of about 35,000 owner units is also interviewed biannually, but the change in owners' equivalent rent is imputed from the rent change of renter units. Prices are collected at bimonthly or semi-annual frequency for strata of items that tend to have infrequent price changes, and the price changes for those items receive 1/2 or 1/6 of the weight of price changes for items that are priced every month.

⁸ When direct information on base-period quantities are not available, the index formula is made operational using information on base-period expenditures,

$$E_{Bi} = Q_{Bi} P_{Bi},$$

and relative price change:

$$I_{T,0} = \frac{\sum_i E_{Bi} (P_{Ti} / P_{Bi})}{\sum_i E_{Bi} (P_{0i} / P_{Bi})}.$$

Commodity Substitution Bias

A true economic cost-of-living index would measure the change in the cost of obtaining a fixed level of economic well-being, or utility. However, the consumer price index calculates the change in cost of obtaining a fixed basket of goods, which is not quite the same thing. Most economists have used the cost-of-living index concept as the standard against which biases of the CPI are to be measured.

The substitution bias reflects the failure of the fixed basket index to account for the fact that consumers will tend to substitute relatively less expensive items in place of items that have become relatively more expensive. Several empirical strategies have been employed to estimate the substitution bias. A strategy that was frequently employed prior to 1980 was to estimate a system of demand equations and then, using this information about what substitution would occur as prices changed, to calculate directly the exact cost-of-living index associated with the demand system (Braithwait, 1980).

The more recent literature has avoided the difficulties of estimating a complete demand system by relying on the concept of *superlative* price indexes, which was introduced by Diewert (1976). In contrast to the Laspeyres index, which requires information on expenditures from only one period, and which for all subsequent periods can be calculated using price information only, superlative indexes generally require complete information on expenditures or quantities for each period. Diewert showed that certain superlative index numbers -- for example, either the Fisher or the Törnqvist

index⁹—will closely approximate an exact cost-of-living index. Essentially, using quantity and price information from each time period allows the substitution to be taken into account.

Several recent empirical studies have compared a Laspeyres index (like the CPI) to superlative indexes. For example, Manser and McDonald (1988) relied on Personal Consumption Expenditure data covering the period 1959-85, whereas Aizcorbe and Jackman (1993) used data from the Consumer Expenditure Survey and strata price indexes covering the period 1982-91. Both studies found that the Laspeyres index tends to grow 0.2 to 0.25 percentage points per year faster than alternative measures that allow for consumer substitution, such as the Fisher or Tornqvist superlative indexes. However, these papers have not unequivocally resolved such questions as whether the substitution

⁹ The formula for the Fisher index is

$$F = (L \cdot P)^{1/2},$$

where L is the simple two-period Laspeyres index,

$$L = \frac{\sum Q_{0i} P_{Ti}}{\sum Q_{0i} P_{0i}},$$

and P is the index weighted by current quantities or *Paasche* index,

$$P = \frac{\sum Q_{Ti} P_{Ti}}{\sum Q_{Ti} P_{0i}}.$$

The Laspeyres index is usually an overestimate of the cost-of-living index, while the Paasche index is usually an underestimate. The Törnqvist formula is

$$T = \exp\left\{\sum \left[\frac{S_{0i} + S_{Ti}}{2}\right] \ln\left(\frac{P_{Ti}}{P_{0i}}\right)\right\},$$

where

$$S_{0i} = \frac{Q_{0i} P_{0i}}{\sum Q_{0i} P_{0i}}$$

is the expenditure share for good i in period 0, and similarly for S_{Ti} .

bias increases with the inflation rate, or whether the rate of substitution bias grows with the time elapsed since the last market basket update.

These superlative indexes also rely on certain assumptions that should be remembered. For example, these measurements have assumed that consumer tastes have remained constant over the measurement period. If demand for certain goods were to shift exogenously—for example, a shift in demand for eggs due to scientific information about the effects of cholesterol on health—it could confound the measurement of the substitution effect because the quantity consumed could fall at the same time that the relative price is falling.¹⁰ This "simultaneity" problem has not been adequately addressed in these studies of substitution effects. More generally, since these studies have largely been based on the expenditures of a representative consumer, thus overlooking issues of aggregation across consumers, they run the risk of confounding true substitution effects with the results of geographical shifts, demographic changes, and changes in consumption motivated by factors other than changes in relative prices. Also, these studies have usually treated prices and expenditures as known amounts rather than as estimates subject to sampling error.

It should also be recognized that Diewert's (1976) original result showed that the superlative indexes provide a close approximation to any exact cost-of-living index only if the income elasticities of consumers are equal to one for all goods (that is, if preferences are homothetic). If income elasticities of demand differ from 1, then a rising income level

¹⁰ Brown and Schrader (1990) studied the demand for eggs, a case in which consumption has fallen despite a declining relative price.

will be shifting the composition of the desired consumer basket over time.¹¹ Users of superlative indexes should recognize that income effects can matter, especially for comparisons over long periods of time, such as several decades, and avoid confusing these income effects with substitution effects from price changes.

The existing studies of substitution bias have been limited to studying substitution at the level of strata indexes and above. This may capture substitution from say, canned soup to frozen meals, but it will not capture substitution inside a given category, like from one type of frozen meal to another. The standard data sources are not available to investigate this issue because the Consumer Expenditure Survey does not provide information below the stratum level. Recently, however, Reinsdorf (1996) and Bradley (1996) have used supermarket scanner data to study low-level substitution bias for a handful of grocery items. These results, although very limited, do suggest that low-level substitution effects may also be important.

Formula Bias

A few years ago, BLS researcher Marshall Reinsdorf (1993) set out to explain a striking empirical result: the fact that CPI for many food items had grown 1.5-2.0 percent

¹¹ Diewert (1976), however, also showed that with non-homothetic preferences, under certain conditions a superlative index will approximate the cost-of-living index for an intermediate utility level, providing a justification for focusing on the superlative index in spite of non-homothetic preferences. More generally, results from Caves, Christensen and Diewert (1982) show that the Törnqvist index, in particular, is the geometric mean of the cost-of-living indexes with reference period and current utility levels under non-homothetic translog preferences. Balk (1990) proposed an econometric method for calculating approximate cost-of-living index numbers for arbitrary base-period income levels.

faster per year than average price series calculated using the same data. In this paper, he attributed the difference entirely to outlet substitution effects, which are discussed in the next section. Subsequent research has shown that much of the difference between the CPI and average prices is attributable to what has become known as "functional form" or "formula" bias (Reinsdorf, 1994).¹²

Formula bias arises in this way. Remember, about one-fifth of the sample is rotating each year. The base price for the sample item should represent its average price during the expenditure base period. Because the sample item had not yet been selected during the base period, neither the base price nor the base-period quantity is observable and a method is required for estimating the base price. From 1978 until 1996 the BLS used the following procedure: take the price of the sample item during the sample replacement or "link" month and deflate it to the base period using the overall price index for the stratum. This procedure causes items that are on sale or otherwise have an unusually low price when they are introduced to the sample to receive a disproportionately large weight, because the expenditure weight is divided by an atypically low base price for the item on sale.¹³ These items are likely to go off sale the next period, and thus show a

¹² The formula bias problem is closely related to a more general problem of upward bias for indexes that are calculated using averages of ratios (Carruthers, Sellwood, and Ward, 1980; Szulc, 1983; Dalén 1992; Diewert, 1995a).

¹³ Specifically, the true modified Laspeyres can be written as

$$\sum W_i (P_{Ti} / P_{0i}),$$

where the weight

$$W_i = E_{Bi} (P_{0i} / P_{Bi}) / \sum E_{Bi} (P_{0i} / P_{Bi})$$

price rise. Conversely, a relatively smaller weight is applied to items that are off sale when the new sample is introduced, and may go on sale the following period. The net effect is that the estimator may apply too much weight to price increases and too little weight to price decreases immediately after the introduction of a new sample or a new sample item. In other words, transitory price movements are systematically related to the weights because of the way those weights have been constructed.

Table 3 shows a numerical example of how formula bias can appear immediately following sample replacement. In this example the sample consists of prices from three outlets for a relatively homogeneous item, such as tomatoes. Each month two of the outlets sell the item at full price, \$2.00, while at one of the outlets the item is on sale for \$1.25. To keep the problem simple, assume that all outlets have expenditure weights equal to \$100. For the old sample, assume all items have base prices equal to \$1.00, so that the implicit base-period quantity purchased at each outlets is equal to dollar expenditures divided by base price, i.e., \$100/\$1.00 per pound = 100 pounds. Although the item is on

represents the share of expenditures that would be spent on i during period 0 if quantities were held fixed at period B levels. If period 0 represents the link month, then the base-price setting method formerly used by the CPI reduces to

$$\sum S_{Bi}(P_{Ti}/P_{0i}),$$

where

$$S_{Bi} = E_{Bi}/\sum E_{Bi}$$

is the base period expenditure share. If (P_{0i}/P_{Bi}) is smaller than the average for the stratum, as, for example, if item i is on sale during the link period, then

$$S_{Bi} > W_{Bi},$$

so the CPI method will apply too much weight to the outlet, relative to the modified Laspeyres target.

sale at outlet B in June and at outlet C in July, the overall price index does not show any overall price change between these months. This is because, with equal expenditures and equal base prices in each outlet, the index change is calculated as the ratio of the sums of the prices times the inferred base period quantities:

$$I_{July} = \frac{100 \times \$2.00 + 100 \times \$2.00 + 100 \times \$1.25}{100 \times \$2.00 + 100 \times \$1.25 + 100 \times \$2.00} = \frac{\$525}{\$525} = 1.$$

But beginning in August a new sample of outlets is drawn, which raises the problem of linking the two indexes. For August, prices from the old sample are used in the index calculation, and again no price change is reflected in the index. When the new outlets enter the sample, the first step is to discount them back to the base period given the overall inflation in the strata; since there hadn't been any overall inflation, the base prices in the new sample are taken to be the same as the prices when these outlets are first sampled in August. Notice that as a result of the sample replacement, the base prices have shifted. The outlet with the sale price during August, outlet E, was implicitly assigned a quantity weight of $\$100/\1.25 per pound = 80 pounds. The two outlets that did not have a sale were implicitly assigned quantity weights of $\$100/\2.00 per pound = 50 pounds. The change in the stratum index from August to September now involves calculating the ratio that multiplies the implied quantities at each outlet times the price at that outlet, and then divides the September figure by the August figure:

$$I_{Sept} = \frac{80 \times \$2.00 + 50 \times \$2.00 + 50 \times \$1.25}{80 \times \$1.25 + 50 \times \$2.00 + 50 \times \$2.00} = \frac{\$322.50}{\$300.00} = 1.075.$$

But the reason behind this conclusion is that the outlet with the sale during August (outlet E) received a low base price, and so the increase in price at that outlet was overweighted compared to the fall in price at outlet D. In the October index, again no change in the index is reported. Although the formula bias cannot be guaranteed to disappear after the first months, empirical studies and simulations have suggested that the bias is usually concentrated in the first month after the calculation of the proxy base prices.

A number of methods have been studied by BLS researchers for improving the estimation procedure. Early research focused on use of alternative estimation formulas, like using the geometric mean rather than the arithmetic mean (Moulton, 1993; Reinsdorf and Moulton, 1994; Moulton and Smedley, 1995). The geometric mean has several attractive econometric characteristics—in particular, it requires only information on base-period expenditures, and it is an exact index formula under Cobb-Douglas preferences—and several other countries have recently adopted a geometric mean estimator for strata indexes. If the objective is to estimate a modified Laspeyres index, however, then the geometric mean has been shown to produce inflation estimates that are systematically too small (McClelland, 1996).

BLS recently adopted a new method for addressing this estimation issue, which it has applied to food-at-home items beginning in January 1995, and to all other items beginning in June and July 1996. The approach is to hold out the new samples for three to four months after the base prices are estimated (Armknecht, Moulton, and Stewart, 1995; McClelland, 1996; Moulton, 1996). In the example in Table 3, note that the upward bias occurs in calculating the August to September index change, immediately following the

rotation. Substantial simulation and analysis indicates that this pattern is typical: the formula bias is concentrated in the first month or two following the introduction of the new sample. Holding out the new sample for several months thus sidesteps the problem.

The magnitude of the formula bias problem, including the problem of inappropriately weighting replacement samples, and a related formula bias in calculation of homeowners' equivalent rent, was estimated to be about 0.24 percent per year during 1993-94.¹⁴ However, the changes that BLS implemented in January 1995 and June and July 1996 have effectively eliminated this formula bias.

Outlet Substitution Bias

Another potential problem related to sample rotation, also originally raised by Reinsdorf (1993), is the treatment of new discount outlets in BLS sampling and estimation procedures. Discount outlets, like all other types of outlets, are selected for CPI samples in proportion to consumer expenditures as reported in the Point-of-Purchase Survey. The new outlets are linked into the survey as described in the example in the previous section. But the linkage procedure means that prices in the old and new outlets are not compared

¹⁴ These calculations are based on unpublished estimates made by Karin Smedley and Claire Gallagher of BLS. The changes implemented in January 1995 had an estimated effect of about 0.14 percent per year, and the changes implemented in 1996 had an estimated effect of about 0.10 percent per year. Simulations of the geometric mean resulted in somewhat larger estimates, as large as 0.5 percent per year including homeowners' equivalent rent, but some of the difference between the geometric mean and the CPI is due to the fact that it is estimating a different population target than the modified Laspeyres index (Moulton and Smedley, 1995).

directly. Thus any savings that consumers potentially receive from switching to discount outlets (after netting out quality differences) are not reflected by the CPI.

In many respects, the entry of discount outlets is a special case of the new goods problem. Market entrants succeed by either providing improved services or lower prices, and in principle these gains to consumer well-being should be reflected in a true cost-of-living index (Fixler, 1993). But measuring the bias that results from linking of new outlets is a tricky empirical problem, since the measurement depends both on the relative quality of services provided by the new and old outlets and the price response of the old outlets. If the price difference of the discount outlet largely reflects a lower level of retail services, then directly comparing the prices between traditional and discount outlets would overstate the consumer gains from entry of the discount outlets. Similarly, if the traditional outlets respond to the entry of discounters by lowering their prices, then the CPI procedures would reflect those price changes.

At this point, the empirical evidence on the effect of discount outlets is quite limited. Reinsdorf (1993) compared prices for food and motor fuel between old and new samples during an overlap period when the samples were undergoing rotation. The differences indicated that prices in the new samples were about 1.25 percent lower than in the old samples over a 2-year period. These results would be consistent with an upward bias of 0.25 percent per year (since one-fifth of the sample rotates each year), assuming that the price differences are not offset by any declines in quality. But the Reinsdorf results are barely statistically significant, and it would be very useful if this research could be replicated for additional years.

A rough calculation of the effects of discount stores can be made using information on the size of the price differential between discount stores and traditional outlets, and the rate of growth of the discount stores. MacDonald and Nelson (1991) provide evidence on the price differential, by comparing the price of food warehouse outlets to traditional outlets. They found that prices in the warehouse stores were about 13.4 percent lower. The rate of growth of the share of warehouse stores between 1983 and 1991 according to data published by the trade journal, *Progressive Grocer*, was about 0.7 percent per year. Combining these---a price difference of 13.4 percent, together with a growth in market share of 0.7 percent per year---would imply a maximum bias for grocery-store food of about $0.134 \times 0.7 = 0.1$ percent per year, assuming no quality differential. If the warehouse stores provide significantly lower retail services, the quality adjustment would further reduce the estimated bias.

The entry of discount outlets is not confined to food; in recent years, it has probably been more important in categories like consumer electronics and hardware. Evidence on outlet substitution bias for these other expenditure categories is not yet available. Some retail services have also seen growth of discount outlets, e.g., discount brokerages. But many of the services categories, which account for 57 percent of the CPI market basket, are categories like rent, electric and gas utilities, and college tuition, which are probably not much affected by outlet substitution factors. Outlet substitution is clearly an important issue whether the bias is 0.25 percent or less than 0.1 percent annually, but like the related problem of new goods, precise estimation is difficult.

Quality Adjustment

Many of the procedures used by the CPI program in processing data are specifically designed for separating price changes from quality changes. The data collection begins with detailed checklists that the data collectors use to assure that precisely the same item is repriced from period to period. If the sample item has changed in any observable way, one of three general procedures may be applied to the data. An economist with specialized knowledge of the item examines information on the two versions of the item and determines whether: a) the change has not resulted in a significant change in the quality of the item, so that the prices of the old version and the new version can be directly compared; or b) a significant change in quality occurred and information is available for estimating the dollar value of the change in quality; or c) a significant change in quality occurred and information on the value of the change in quality is not available.

Manufacturers of a product are one possible source of information of what a certain change in quality cost. An alternative method is hedonic regression analysis. The hedonic method estimates the price-quality relationship by running regressions of price on characteristics of goods. The coefficients of these regressions can then be used to infer the value of changes in characteristics of the goods in the sample. For example, the observed valuation of computers with different processor speeds could be used to estimate the quality improvement of a new computer with a faster processor. The CPI has used

hedonic methods since 1988 for calculating the effects of depreciation on rent, and since 1991 for quality changes in apparel.¹⁵

When hedonic methods aren't practical, then some other method must be found for linking or imputing the effect of the quality change on price. To understand how such a linkage can work, consider a simplified situation where a certain product is available one month, but then is replaced the next month by a product of different quality. In a linkage calculation, the first step is to calculate the rate of inflation during that month based only on a class of other, similar goods, and completely ignoring the good which was replaced. For the sake of this example, say that the inflation rate based on the other goods was 2 percent, but that the new and improved product, when it appeared, cost 5 percent more than the earlier version. Then, linkage effectively assumes that of the 5 percent, $2/5$ was due to the overall rise in the price of goods, and the other $3/5$ was due to a quality improvement.

The method of linking can produce the optimal quality adjustment under certain assumptions; for example, if price levels are continuously at a competitive equilibrium (more precisely, a competitive hedonic equilibrium in which prices reflect all quality differences), and all items are close substitutes. Most of the evidence suggests, however, that when price changes are relatively small and quality improvements are substantial,

¹⁵ For interpretation of the traditional methods used in hedonic quality adjustment, see Gordon (1990), Griliches (1990), and Triplett (1990). For description of changes in CPI quality adjustment procedures, see Randolph (1988), Liegey (1993), and Reinsdorf, Liegey, and Stewart (1996). In recent years a number of researchers have extended the theory of hedonic quality adjustment, integrating the method with the theory of the cost-of-living index (Triplett, 1983a; Fixler and Zieschang, 1992; Feenstra, 1995) and accounting for non-competitive market structure and consumer heterogeneity (Berry, Levinsohn, and Pakes, 1995; Goldberg, 1995).

linking tends to understate the value of quality improvements from one version to the next. For example, in the move from 486 to Pentium-based personal computers, if one subtracts out the relatively small overall change in price levels, the remaining difference in price doesn't seem large enough to capture the true increase in computer performance from one generation of chip to the next. On the other hand, for goods that do not have substantial quality improvements, the method of linking may attribute too much quality change and too little price change to the replacement of models. Thus, in principle, any bias due to the method of linking could be either upwards or downwards. Recognizing the problems associated with linking, the BLS has taken steps in recent years to reduce the dependency on linking, and increase the use of direct comparisons and direct quality adjustments. When linking must be done, BLS has adopted new methods to determine a more comparable class of other goods from which to calculate the inflation rate.

The direct quality adjustments and implicit quality adjustments due to linkage can be significant---the change in the price index for new cars from 1967 to 1994 would have been 80 percent greater if no adjustments had been made for quality improvements.¹⁶ In the past, the BLS has tended to rely on the linkage approach, and a number of studies have compared BLS indexes to hedonic indexes for specific items or groups of items. The most extensive of these studies is Gordon (1990), which found that because the BLS indexes failed to capture quality improvements fully, the inflation rate in consumer

¹⁶ The automobile index in the CPI rose 172.1 percent from December 1967 to December 1994, while over the same period, without the quality adjustments, the new car component would have risen 313.4 percent (Abraham, 1995).

durables was biased upward by 1.5 percent per year over the period 1947-83.¹⁷ Large upward biases were also found by Berndt, Griliches, and Rappaport (1995) for personal computers and Griliches and Cockburn (1994) for prescription drugs, which were attributed to overreliance on linking. In contrast, Reinsdorf, Liegey, and Stewart (1996) suggest that linking may have led to downward bias for the apparel indexes, at least during 1978-86.

To date, the empirical research on quality bias has been heavily concentrated in durables and apparel. Since researchers choose to focus on sectors for which a problem is perceived, the results from the sectors that have been studied may not necessarily provide a useful guide to sectors that have not been studied. There are many sectors of the CPI, particularly services (including medical services), for which little research has been done that would provide information on the magnitude or, in some cases, even the direction of quality bias.¹⁸ Research on quality changes can be difficult and tedious, and not all quality changes are amenable to hedonic or related procedures. Nevertheless, it would be very useful if researchers could try to fill in some of the gaps.

¹⁷ Because the BLS has introduced several improvements to quality adjustment procedures since the beginning of Gordon's study (Reinsdorf, Liegey, and Stewart, 1996), the relevance of his estimate of bias to the current CPI is questionable.

¹⁸ Shapiro and Wilcox (1996) provide an interesting case study of improvement in medical treatment of cataracts that is suggestive of large quality bias for medical services.

New Goods

Sometimes new goods provide a service similar to an existing good, but with higher quality or a lower price: a generic drug provides the same service as its branded predecessor; a compact disk provides higher quality recorded music than a vinyl record. In other cases, new goods offer an additional variety of choices, but without fundamentally changing the services provided, as happens when new varieties of blue jeans are introduced. Finally, some new goods provide entirely new services that were previously unavailable, like interactive video games or cellular telephones. For the consumer price index, the appearance of new goods presents at least two important problems: bringing new goods into the samples on a timely basis; and accounting for differences in price between new goods and the old goods that provided the same or similar services (Armknrecht, Lane, and Stewart, 1994).

One of the purposes of the CPI's periodic sample rotation is to bring new goods into the sample in a timely manner. As discussed earlier, one-fifth of the sample is replaced each year. BLS is planning to change the Point-of-Purchase Survey procedures during the next CPI revision so that more frequent sample replacements could be made as needed for specific categories of items.

The procedure of sample rotation by itself, however, may not appropriately account for improvements to consumer well-being that result from introduction of the new goods. The sample rotation results in a linking of the old and new samples, hence the implicit assumption is that prices in both samples fully reflect quality differences. In some

cases this assumption may be appropriate, if prices of the old goods fall as a result of direct competition from the new goods. But if the new sample includes items which provide the same services at a cheaper price, or new services that were previously unavailable, and the prices of the old goods do not fall commensurably, then benefits of those improvements will not be fully reflected in the measured price change.

Measurement of the new goods bias appear to be pretty much guesswork at present. Hicks (1940) showed that for consistency with the economic theory of the cost-of-living index, the consumer's surplus derived from introduction of the new good should be measured by reference to its reservation price. But estimation of reservation prices is a tricky econometric problem. These estimates appear to me to potentially confound several effects. The Hicksian consumer surplus from the introduction is the pure new goods effect, but the introduction of new goods is often immediately followed by significant quality improvements and price declines. Clearly separating the pure effects of new goods from quality change bias and substitution bias is a difficult empirical proposition. Most of the recent estimates of new goods bias in the CPI, including the estimate in the Advisory Commission report, are based on back-of-the-envelope calculations, and it seems possible that some of what is being counted as a new goods effect is also being included in estimates of quality change or substitution. If true, this double-counting would have the effect of overstating the overall upward bias in the CPI.

Several recent papers suggest that the CPI is missing some very large gains in consumer welfare because of the new goods problem. Hausman (1994) studied introduction of new brands of breakfast cereals, estimating the Hicksian consumer surplus

directly using econometric estimation of a demand system, and concluded that increases in consumer surplus that are missed by the CPI lead the price index for cereal to be substantially overstated. This finding is surprising for a commodity that is not undergoing major technological improvements.

However, Hausman's method for estimating the reservation prices for the new brands involves extrapolation of the demand functions well outside of the sample region. Nordhaus (1994) analyzed the cost of indoor illumination, and showed that there have been dramatic reductions in the price of light, as measured in lumens, when new technologies (such as compact fluorescent bulbs) are introduced. These product innovations would typically be linked into the CPI as the sample rotation picks up new products, rather than appearing as price declines in an existing product.

On the other hand, many economists have expressed skepticism about possibly exaggerated claims for the importance of new goods. If a new good replaces other close substitutes, it may be implausible to claim that its reservation price is much higher than its market price. If the price of a new good falls dramatically, consumers will tend to apply it to low valued uses (e.g., computers used for playing games, lights left on in unoccupied rooms.) The current state of empirical research has not done much to narrow the set of plausible beliefs about the effects of new goods.

Concluding Comments

Besides the various components of CPI bias, other CPI issues deserve our attention and may affect the interpretation and policy uses of the measure. For example, should an escalator intended for a specific demographic group, such as Social Security recipients, reflect the expenditure patterns of that group? The CPI Advisory Commission (1995) has argued not, based on several studies that found little difference between indexes calculated for specific groups. If the basis for this is empirical, however, changes in economic conditions could cause this result to be reversed. Furthermore, the biases themselves could have differing impacts across different demographic groups. In my view, further research is warranted on possible differences of inflation between the elderly and the non-elderly, the poor and the non-poor, and other groups with different expenditure patterns. Another issue is the effects of non-market goods, which affect consumer well-being, but are not measured in studies that estimate cost-of-living indexes based on prices of market goods. The purposes for which an index is to be used are clearly relevant in determining how an index is to be defined. For example, Triplett (1983*b*) argued that the usual expenditure-based cost-of-living index may not be the appropriate index for escalation of Social Security benefits and pensions. Others, including Griliches (1995), have questioned the policy of fixed escalation formulas.

The state of knowledge about the substitution and formula biases in the CPI has increased substantially in the last couple of years, and recent changes to BLS procedures have essentially eliminated the formula bias. However, the other categories of CPI bias call

out for additional information. Although I have suggested that the net effect of outlet substitution bias may not be large, entry of new firms is a pervasive phenomenon in the retail sector and one that is amenable to further research. Many researchers have tackled parts of the quality adjustment problem, but for selected categories of items. The new goods problem is the least amenable to systematic study, though a number of provocative papers have recently been written.

Perhaps the most encouraging outcome to date is the renaissance of research on price measurement issues. New data sources, such as supermarket scanner data and microdata from retail and trade associations, are providing detailed information that previously was not available. Although confidentiality restrictions have prevented general distribution by BLS of microdata on prices, BLS has policies that allow researchers access to the microdata for specific research projects (de Wolf, 1995). The tough measurement problems associated with quality change and new goods provide ample opportunities for researchers to do significant empirical and theoretical work, which ultimately may lead to improvements in the quality of the price data produced by government agencies. If the advance of a science is constrained by the quality of its measurement, then these are issues that should engage our best researchers.

References

Abraham, Katharine G., "Prepared Statement." In "Consumer Price Index: Hearings Before the Committee on Finance, United States Senate." Senate Hearing 104-69, Washington, D.C.: U.S. Government Printing Office, 1995, pp. 107-9.

Advisory Commission to Study the CPI, "Toward a More Accurate Measure of the Cost of Living: Interim Report to the Senate Finance Committee," mimeo, Senate Finance Committee, Washington, D.C., 1995.

Aizcorbe, Ana M., and Patrick C. Jackman, "The Commodity Substitution Effect in CPI Data, 1982-91." *Monthly Labor Review*, December 1993, 116:12, 25-33.

Armknrecht, Paul A., Brent R. Moulton, and Kenneth J. Stewart, "Improvements to the Food at Home, Shelter, and Prescription Drug Indexes in the U.S. Consumer Price Index." Bureau of Labor Statistics Working Paper No. 263, Washington, 1995.

Armknrecht, Paul A., Walter F. Lane, and Kenneth J. Stewart, "New Products and the U.S. Consumer Price Indexes." In Bresnahan, T. and R.J. Gordon, eds., *The Economics of New Goods*. Forthcoming, 1996.

Balk, Bert M., "On Calculating Cost-of-Living Index Numbers for Arbitrary Income Levels," *Econometrica*, January 1990, 58:1, 75-92.

Berndt, Ernst R., Zvi Griliches, and Neal J. Rappaport, "Econometric Estimates of Price Indexes for Personal Computers in the 1990s," *Journal of Econometrics*, July 1995, 68:1, 243-68.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Automobile Prices in Market Equilibrium," *Econometrica*, July 1995, 63:4, 841-90.

Boskin, Michael J., "Prepared Statement." In "Consumer Price Index: Hearings Before the Committee on Finance, United States Senate." Senate Hearing 104-69, Washington, D.C.: U.S. Government Printing Office, 1995, pp. 109-15.

Bradley, Ralph, "The Use of Scanner Data as a Means to Reduce the Mean Squared Error in the CPI," mimeo, Bureau of Labor Statistics, 1996.

Braithwait, Steven D., "Substitution Bias of the Laspeyres Price Index: An Analysis Using Estimated Cost-of-Living Indexes," *American Economic Review*, March 1980, 70:1, 64-77.

Brown, Deborah J., and Lee F. Schrader, "Cholesterol Information and Shell Egg Consumption," *American Journal of Agricultural Economics*, August 1990, 72:3, 548-55.

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

Caves, Douglas W., Laurits R. Christensen, and W. Erwin Diewert, "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity," *Econometrica*, November 1982, 50:6, 1393-414.

Congressional Budget Office, *Is the Growth of the CPI a Biased Measure of Changes in the Cost of Living?* Washington, D.C.: Congressional Budget Office, October 1994.

Dalén, Jörgen, "Computing Elementary Aggregates in the Swedish Consumer Price Index," *Journal of Official Statistics*, 1992, 8:2, 129-47.

Darby, Michael R., "Statement." In "Consumer Price Index: Hearings Before the Committee on Finance, United States Senate." Senate Hearing 104-69, Washington, D.C.: U.S. Government Printing Office, 1995, pp. 173-76.

de Wolf, Virginia A., "Researcher Access to Confidential Microdata at the Bureau of Labor Statistics," Proceedings of the Section on Government Statistics American Statistical Association, 1995.

Diewert, W. Erwin, "Exact and Superlative Index Numbers," *Journal of Econometrics*, 1976, 4, 115-145.

Diewert, W. Erwin, "Axiomatic and Economic Approaches to Elementary Price Indexes." University of British Columbia Discussion Paper No. 95-01, 1995a.

Diewert, W. Erwin, "Prepared Statement." In "Consumer Price Index: Hearings Before the Committee on Finance, United States Senate." Senate Hearing 104-69, Washington, D.C.: U.S. Government Printing Office, 1995b, pp. 115-18.

Feenstra, Robert C., "Exact Hedonic Price Indexes," *Review of Economics and Statistics*, November 1995, 77:4, 634-53.

Fixler, Dennis, "The Consumer Price Index: Underlying Concepts and Caveats," *Monthly Labor Review*, December 1993, 116:12, 3-12.

Fixler, Dennis, and Kimberly D. Zieschang, "Incorporating Ancillary Measures of Process and Quality Change into a Superlative Productivity Index," *Journal of Productivity Analysis*, 1992, 2:2, 245-67.

Goldberg, Pinelopi Koujianou, "Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry," *Econometrica*, July 1995, 63:4, 891-951.

Gordon, Robert J., *The Measurement of Durable Goods Prices*. Chicago: University of Chicago Press, 1990.

Gordon, Robert J., "Prepared Statement." In "Consumer Price Index: Hearings Before the Committee on Finance, United States Senate." Senate Hearing 104-69, Washington, D.C.: U.S. Government Printing Office, 1995, pp. 122-26.

Greenspan, Alan, "Prepared Statement." in U.S. Congress, "Review of Congressional Budget Cost Estimating: Joint Hearing before the House of Representatives Committee on the Budget and the Senate Committee on the Budget." Serial No. 104-1, Washington, D.C.: U.S. Government Printing Office, 1995, pp. 130-49.

Griliches, Zvi, "Hedonic Price Indexes and the Measurement of Capital and Productivity: Some Historical Reflections." In Berndt, E.R., and J.E. Triplett, eds., *Fifty Years of Economic Measurement: The Jubilee of the Conference on Research in Income and Wealth*. Chicago: University of Chicago Press, 1990, pp. 185-202.

Griliches, Zvi, "Prepared Statement." In "Consumer Price Index: Hearings Before the Committee on Finance, United States Senate." Senate Hearing 104-69, Washington, D.C.: U.S. Government Printing Office, 1995, pp. 129-32.

Griliches, Zvi, and Iain Cockburn, "Generics and New Goods in Pharmaceutical Price Indexes," *American Economic Review*, December 1994, 84:5, 1213-32.

Hausman, Jerry A., "Valuation of New Goods Under Perfect and Imperfect Competition." NBER working Paper No. 4970, 1994; forthcoming in Bresnahan, T. and R.J. Gordon, eds., *The Economics of New Goods*.

Hicks, John, "The Valuation of the Social Income," *Econometrica*, August 1940, 7, 108-24.

Jorgenson, Dale W., "Statement." In "Consumer Price Index: Hearings Before the Committee on Finance, United States Senate." Senate Hearing 104-69, Washington, D.C.: U.S. Government Printing Office, 1995, pp. 36-41.

Klumpner, Jim, "Fact and Fancy: CPI Biases and the Federal Budget," *Business Economics*, April 1996, 31:2, 22-29.

Lebow, David E., John M. Roberts, and David J. Stockton, "Monetary Policy and 'The Price Level,'" mimeo, Federal Reserve Board, Washington, D.C., 1994.

Liegey, Paul R., Jr., "Adjusting Apparel Indexes in the Consumer Price Index for Quality Differences." In Foss, M.F., M.E. Manser, and A.H. Young, eds., *Price Measurements and Their Uses*, Chicago: University of Chicago Press, 1993, pp. 209-26.

MacDonald, James M., and Paul E. Nelson, Jr., "Do the Poor Still Pay More? Food Price Variations in Large Metropolitan Areas," *Journal of Urban Economics*, November 1991, 30:3, 344-59.

Manser, Marilyn E., and Richard J. McDonald, "An Analysis of Substitution Bias in Measuring Inflation, 1959-85," *Econometrica*, 56:4, July 1988, 909-30.

McClelland, Robert, "Evaluating Formula Bias in Various Indexes Using Simulations," mimeo, Bureau of Labor Statistics, 1996.

Moulton, Brent R., "Basic Components of the CPI: Estimation of Price Changes," *Monthly Labor Review*, December 1993, 116:12, 13-24.

Moulton, Brent R., "Estimation of Elementary Indexes of the Consumer Price Index," mimeo, Bureau of Labor Statistics, 1996.

Moulton, Brent R., and Smedley, Karin E., "A Comparison of Estimators for Elementary Aggregates of the CPI," mimeo, Bureau of Labor Statistics, 1995.

Nordhaus, William D., "Do Real Output and Real Wage Measures Capture Reality? The History of Lighting Suggests Not." Cowles Foundation Discussion Paper No. 1078, 1994; forthcoming in Bresnahan, T. and R.J. Gordon, eds., *The Economics of New Goods*.

Pakes, Ariel, "Statement." In "Consumer Price Index: Hearings Before the Committee on Finance, United States Senate." Senate Hearing 104-69, Washington, D.C.: U.S. Government Printing Office, 1995, pp. 44-48.

Randolph, William C., "Housing Depreciation and Aging Bias in the Consumer Price Index," *Journal of Business and Economic Statistics*, July 1988, 6:3, 359-71.

Reinsdorf, Marshall, "The Effect of Outlet Price Differentials in the U.S. Consumer Price Index." In Foss, M.F., M.E. Manser, and A.H. Young, eds., *Price Measurements and Their Uses*, Chicago: University of Chicago Press, 1993, pp. 227-54.

Reinsdorf, Marshall, "Price Dispersion, Seller Substitution, and the U.S. CPI." Working Paper No. 252, Washington, Bureau of Labor Statistics, 1994.

Reinsdorf, Marshall, "Constructing Basic Component Indexes for the U.S. CPI from Scanner Data: A Test Using Data on Coffee." Working Paper No. 277, Bureau of Labor Statistics, 1996.

Reinsdorf, Marshall, Paul Liegey, and Kenneth Stewart, "New Ways of Handling Quality Change in the U.S. Consumer Price Index." Working Paper No. 276, Bureau of Labor Statistics, 1996.

Reinsdorf, Marshall, and Brent R. Moulton, "The Construction of Basic Components of Cost of Living Indexes." BLS working paper no. 261, 1995; forthcoming in Bresnahan, T. and R.J. Gordon, eds., *The Economics of New Goods*.

Shapiro, Matthew D., and David W. Wilcox, "Causes and Consequences of Imperfections in the Consumer Price Index," mimeo, University of Michigan, 1996; forthcoming in *NBER Macroeconomics Annual 11*.

Stigler, George, ed., *The Price Statistics of the Federal Government. Report to the Office of Statistical Standards, Bureau of the Budget.* New York: National Bureau of Economic Research, 1961.

Szulc, Bohdan J., "Linking Price Index Numbers." In Diewert, W.E. and C. Montmarquette, eds., *Price Level Measurement: Proceedings from a Conference Sponsored by Statistics Canada,* Ottawa: Statistics Canada, 1983, pp. 537-66.

Triplett, Jack E., "Concepts of Quality in Input and Output Price Measures: A Resolution of the User Value–Resource Cost Debate." In Foss, M.F., ed., *The U.S. National Income and Product Accounts: Selected Topics,* Chicago: University of Chicago Press, 1983a, pp. 269-311.

Triplett, Jack E., "Escalation Measures: What is the Answer? What is the Question?." In Diewert, W.E. and C. Montmarquette, eds., *Price Level Measurement: Proceedings from a Conference Sponsored by Statistics Canada,* Ottawa: Statistics Canada, 1983b, pp. 457-87.

Triplett, Jack E., "Hedonic Methods in Statistical Agency Environments: An Intellectual Biopsy." In Berndt, E.R., and J.E. Triplett, eds., *Fifty Years of Economic Measurement: The Jubilee of the Conference on Research in Income and Wealth,* Chicago: University of Chicago Press, 1990, pp. 207-33.

U.S. Senate, "Consumer Price Index: Hearings before the Committee on Finance, United States Senate." Senate Hearing 104-69, Washington, U.S. Government Printing Office, 1995.

Wynne, Mark A., and Fiona D. Sigalla, "The Consumer Price Index," *Federal Reserve Bank of Dallas Economic Review,* Second Quarter 1994, 1-22.

*Table 1.***Recent Estimates of Bias in the U.S. Consumer Price Index**

<i>Author(s)</i>	<i>Point Estimate</i>	<i>Interval Estimate</i>
Advisory Commission to Study the CPI (1995)	1.0	0.7 - 2.0
Michael Boskin (1995)	1.5	1.0 - 2.0
Congressional Budget Office (1995)	----	0.2 - 0.8
Michael R. Darby (1995)	1.5	0.5 - 2.5
W. Erwin Diewert (1995b)	----	1.3 - 1.7
Robert J. Gordon (1995)	1.7	----
Alan Greenspan (1995)	----	0.5 - 1.5
Zvi Griliches (1995)	1.0	0.4 - 1.6
Dale W. Jorgenson (1995)	1.0	0.5 - 1.5
Jim Klumpner (1996)	----	0.3 - 0.5
Lebow, Roberts, and Stockton (1994)	----	0.4 - 1.5
Ariel Pakes (1995)	0.8	----
Shapiro and Wilcox (1996)	1.1	0.7 - 1.6
Wynne and Sigalla (1994)	less than 1.0	----

*Table 2.***Relative Importance of U.S. CPI Categories, December 1995**

<i>Expenditure category</i>	<i>Relative Importance</i>
Food and Beverages	17.3
Housing	41.3
Apparel and Upkeep	5.5
Transportation	17.0
Medical Care	7.4
Entertainment	4.4
Other Goods and Services	7.1
Total	100.0
Commodities	42.9
Services	57.1
Total	100.0

Note: Relative importance is based on share of consumer out-of-pocket expenditures during 1982-84, updated by price change through December 1995.

Table 3.

Numerical Example of Formula Bias

<i>item</i>	<i>Base Price^a</i>	<i>Base Price^b</i>	<i>Price</i>				
			<i>June</i>	<i>July</i>	<i>August (link)</i>	<i>September</i>	<i>October</i>
Old sample							
A	1.00		2.00	2.00	1.25		
B	1.00		1.25	2.00	2.00		
C	1.00		2.00	1.25	2.00		
New sample							
D		2.00			2.00	1.25	2.00
E		1.25			1.25	2.00	2.00
F		2.00			2.00	2.00	1.25
Index	100	100	100	100	100	107.5	107.5
Percentage change				0	0	7.5%	0

^a Estimated base-period prices for outlets in the old sample.

^b Estimated base-period prices for outlets in the new sample.