Charles R. Hulten is a professor of economics at the University of Maryland and a research associate of the National Bureau of Economic Research. The author would like to thank Brent Moulton, Marshall Reinsdorf, John Greenlees, Walter Lane and Ralph Monaco for their comments on earlier drafts.

Quality Change in the CPI

Charles R. Hulten

When considering the effects of technical improvement on the "real price" of manufactured goods, Adam Smith observed, "Quality . . . is so very disputable a matter, that I look upon all information of this kind as somewhat uncertain." This uncertainty has scarcely diminished during the past two centuries. During congressional testimony in January 1995, Federal Reserve Board Chairman Alan Greenspan expressed the opinion that the consumer price index (CPI) was biased upward "by perhaps 1/2 to 1 1/2 percent per year." If he were with us today, Adam Smith would not be surprised to learn that unmeasured improvements in product quality are a major source of this bias.

Congress established the Advisory Commission to Study the Consumer Price Index (Boskin Commission) to pass judgment on the magnitude of the CPI bias. It finished its work late in 1996, estimating a 1.1 percent per year bias, with a range of uncertainty between 0.8 percent to 1.6 percent. Much of the reaction to the Commission's report has focused on these estimates' budgetary implications. The cumulative effects of a 1.1 percent per year bias contributes $202 billion to the federal budget deficit by the year 2008 and adds $1 trillion to the national debt, according to Commission estimates. Viewed as an implicit and unintended expenditure program, the CPI bias would be the fourth largest federal program, behind social security, health expenditures, and defense.

These estimates of the CPI bias also open a Pandora's box of questions about the true nature of U.S. economic growth. If the growth rate of consumer prices is lower than official statistics suggest, there is an opposite understatement in the corresponding component of real gross domestic product (GDP). Biases of the magnitude described by the Boskin Commission can, in fact, make a very large difference in real output. If, for example, the bias in *all* measured prices were 1.1 percent per year (not just in the CPI), 1994 real GDP per capita would have been $35,594 in 1992 prices (not the officially reported $25,335). Because real GDP per capita was $12,512 in 1960, the rate of growth of real product would have been significantly larger.

This shift in metric does not mean that people are actually better off than they thought they were as a result of the commission's report. Utility does not emanate from statistics, but from the actual goods and services consumed during the 1960-94 period. In addition, because the quantity of goods consumed is the same under any dollar metric, past utility is invariant to the choice of metric. But the use of accurate quality-corrected estimates of price and quantity is important for issues such as cost-of-living adjustments and inflation management, the calibration of poverty, and more generally, an accurate impression of the growth dynamics of the U.S. economy.

The Boskin Commission's report helps fix the potential size of the problem, but its range of possible biases—from 0.8 to 1.6—is still large. And, given the "uncertain" state of quality-change research, it is possible that the range is even larger. Indeed, an examination of the procedures used by the Bureau of Labor Statistics (BLS) to construct the CPI suggests that two biases may have this effect. One bias arises on the cost side of the quality-change problem and might extend the top of the commission's range to over 2.0 percent. A second possible bias arises from the "link" method used by BLS in estimating the quality component of price changes. This might lower the Commission's range by enough that the lower bound almost covers zero. In other words, it is possible (but not probable) that "fixing" the CPI will not lead to a major change in its growth rate. An additional implication is that the large revenue gains associated with the fix will not materialize.

THE SCOPE AND NATURE OF THE MEASUREMENT PROBLEM

The Boskin Commission identifies two general sources of measurement biases in prices (and thus quantities). One part of the bias arises from technical errors in constructing the index. The Commission estimates that these errors account for about half the total bias. The other half is because of a failure to measure adequately improvements in the quality of goods and services. I concentrate on the quality problem. Since Adam Smith's time, quality is the more "uncertain" source of bias in measured prices.

Indeed, Adam Smith saw quality change only as "somewhat uncertain." Others have been more emphatic. Shapiro and Wilcox (1996), for example, observe, "Quality change is the house-to-house combat of price measurement." It is therefore somewhat surprising that, from a strictly theoretical standpoint, no natural economic concept of "quality" exists. Differences among varieties of a product could, in principle, be accommodated by treating each variety as a distinct good in its own right, with its own price and quantity. The same index number procedures used to aggregate any collection of distinct goods could then be applied to similar varieties of a given type of good.

The concept of "quality" is useful because there are simply too many varieties of a good to handle as distinct goods. Residential houses, for example, differ as to square footage, number of bedrooms and bathrooms, fireplaces, layout, heating and cooling systems, landscaping, access to schools, access to transportation, and many other amenities. These dimensions give rise to a fearsome number of possibilities. The separate-good approach would virtually require each house to be treated as a different good if carried to its logical conclusion. Autos, personal computers, machines, tools, and appliances are among the many products that come with many different options.

The budgets and computational facilities of the U.S. statistical agencies are too small to treat each variety of a product as a separate good. Products must therefore be grouped into classes of related goods, with product variations treated as different qualities. Quality change occurs when improvements in product design cause consumers to shift their spending to superior varieties, or when the product mix changes in response to price or income changes.

MODELING QUALITY CHANGE

The Grouping of Goods on the Demand Side

The goods approach treats the products bought and sold in the marketplace as the units of analysis. Product heterogeneity and quality differentials are dealt with (in principle) by grouping "similar" items, defined with respect to the degree of substitutability.

When goods X and Y are perfect substitutes, they can be treated as the same commodity, added together to form a single good, and treated as the same good. If, on the other hand, a consumer would be indifferent between one unit of Y and $(1+\theta)$ units of X, the two goods are perfect substitutes up to some constant θ and can be added together by treating each unit of Y as though it were $(1+\theta)$ units of X. In this case, the good Y can be redefined in efficiency units as $Y^e = (1+\theta)Y$ and added together with units of X. This approach amounts to converting differences in the quality of goods into an equivalent difference in quantity, a familiar approach in aggregation theory.

The efficiency-units approach to aggregation requires an estimate of θ, usually obtained from the price side of the problem by constructing the efficiency price of Y: $P^e_Y = P_Y/(1+\theta)$. This is the price of Y expressed in equivalent units of X. It is proportionately smaller (or larger) than the ordinary price of Y by exactly the opposite extent to which the quantity Y^e is larger (smaller) than Y. However, total expenditure on the good Y is the same under both measurement conventions, since $P^e_Y Y^e$ equals $P_Y Y$. This fact is useful, because an estimate of Y^e can be obtained by deflating the total expenditure on Y by the efficiency price.

A simple example illustrates the efficiency-units approach. Suppose one unit of a new and improved toothpaste (Y) sells for $1.10 and a 10 percent quality premium exists relative to the old toothpaste (X). Here, θ is 0.10 and each unit of Y embodies 1.1 efficiency units. The efficiency price of Y is thus $1.00. If the contemporaneous price of X is also $1.00, a consumer buying one unit of both toothpaste varieties obtains the equivalent of 2.1 units of the old variety.

This example also illustrates how quality may affect the measurement of price inflation. Suppose that, from one month to the next, consumers stop buying the old toothpaste and switch to the new variety. A statistician measuring price increases will note that the price per unit of toothpaste has risen by 10 percent. This increase looks like price inflation. However, an alert statistician would observe that the quality-adjusted price of Y has not risen, so the real rate of inflation is zero. On the other hand, a statistician who fails to notice the quality differential between X and Y will overstate price inflation and understate quantity change.

The treatment of product heterogeneity in this manner works well when the marginal rate of substitution between X and Y is constant at $(1+\theta)$. The application of this aggregation procedure has its limits, however. New goods may appear for which no comparable antecedent exists, and therefore we have no opportunity to express Y in equivalent units of X (e.g., videocassette recorders). I turn to this issue later and focus next on a problem that has received less attention than new goods: the supply side of the quality problem.

Quality Measurement with Non-Proportional Cost Increases

The quality differentials that consumers face on the economy's demand side $(1+\theta)$ originate in decisions about product quality made on the supply side. Producers may compete for consumer attention by improving product quality. They may reduce quality to appeal to different segments of the market or to survive in the face of a cost squeeze. Such quality and design changes usually cost money, often in the form of research and development (R&D) programs, although this is not always the case. But, costly or not, demand-side quality decisions are by themselves insufficient to establish a relation between a product's price and quality.

The emerging literature on the supply side of the quality-measurement problem has distinguished two cases—one in which the cost of a new product variety rises in exact proportion to the relative increase in quality, and one in which the cost increase is less than proportional.[1] This distinction is important because strict proportionality leads to an "interior solution" at which the marginal rate of substitution between old and new varieties is equal to the corresponding ratio of marginal costs (i.e., the marginal rate of transformation). Nonproportionality, on the other hand, is associated with "corner solutions" in which this equality does not hold.

The model developed in Hulten (1996) may be adapted to show how the degree of proportionality affects the decomposition of prices and quality. This extension involves a new parameter, μ, defined as the elasticity of marginal cost of producing good Y with respect to θ. This parameter has a value of one if quality change drives up the marginal cost of Y in exact proportion to the increase in quality (θ). However, when improvements in quality are essentially costless, the marginal cost of the new model is the same as the old and the parameter μ is zero. In the general case in which μ is neither one nor zero, the ratio of the marginal costs is $(1+\mu\theta)$. This is the trade-off that determines relative prices of X and Y on the supply side of the market.[2]

[1] Gordon (1990) introduced this distinction. However, its origins lie in the "user value" versus "resources cost" debate of the 1950s and 1960s and Triplett's (1983) resolution. My own recent work on capital goods builds on this distinction and is the basis for the parameterization of the supply side proposed in this article.

[2] Here, I treat the cost elasticity as a constant parameter. In reality, it is more likely to be a variable whose magnitude changes through time. If, for example, an improvement in the quality of a product is the result of a costly R&D program, the variable μ can be interpreted as the markup over the marginal resource cost needed to amortize the R&D investment. This markup is likely to be larger in the early years of the product cycle.

The trade-off on the demand side continues to be $(1+\theta)$, regardless of the magnitude of μ, since the relative benefits of quality (the ratio of the marginal utility of the old and new varieties) are independent of costs. When μ equals one, MRS equals the marginal rate of transformation (MRT). An "interior solution" may be attained at which positive quantities of both goods are consumed. However, when μ is less than one, the MRT = $(1+\mu\theta) < (1+\theta)$ = MRS and a disjunction exists between the two sides of the market. This is the "corner solution" case noted above, in which the marginal cost of acquiring the superior model Y is less than the corresponding benefits. The result is that the old good X is cost-ineffective and therefore disappears from the market.

When prices are proportional to marginal costs, the parameter μ can be shown to influence the separation of a total price change into quality and pure price inflation components. When a new model Y appears in the market, the total change in price from one period to the next can be expressed as $P_Y(t)/P_X(t-1)$. This ratio can be decomposed into three components:

$$(1) \qquad \frac{P_Y(t)}{P_X(t-1)} = \frac{P_Y(t)}{P_Y^e(t)} \; \frac{P_Y^e(t)}{P_X(t)} \; \frac{P_X(t)}{P_X(t-1)} .$$

This expression is an identity and is therefore true by construction. Its relevance arises because the first term is the ratio of the actual price of Y to its quality-adjusted price. This ratio is equal to $(1+\theta)$ and is equivalent to the MRS.[3] It is therefore the pure quality component of the price change from the consumer's standpoint. The product of the first two terms is the price ratio P_Y/P_X, the marginal rate of transformation. The second term must be equal to the ratio of the MRT to the MRS, $(1+\mu\theta)/(1+\theta)$. The third term is the change in the price of good X from one period to the next, $(1+\rho)$. This is the X component of pure price inflation.

The second term is the unconventional aspect of the price decomposition given by Equation 1. It can be interpreted as the cost-effectiveness of obtaining one unit of X in the form of good Y. When μ is one, the second terms disappears (equals one), because one unit X is equivalent in cost to one quality unit of Y. But when μ equals zero, the second term takes the value $1/(1+\theta)$. It is therefore more cost-effective to obtain one quality unit of X packaged in the form of Y than it is to buy one unit of X.

The intuition behind the price decomposition is illustrated by two examples. Suppose we have two varieties of laundry detergent, but one is three times as desirable. In other words, a consumer would be willing to trade three boxes of the inferior detergent for just one of the superior variety. Assume, first, that μ equals one. This implies that the cost of achieving the quality of the superior detergent is three times greater than the cost of its inferior alternative. If the latter costs \$1 per box, the cost of the former will be \$3 per box. Note, however, that the quality-adjusted cost/price of the superior detergent is \$1, since one-third of a box of the new detergent does the job of a whole box of the inferior detergent. In this case, the left side of Equation 1 is 3.0, the first term on the right side is 3.0, the second term is 1.0, and the third term is 1.0. The \$2 price differential in the two detergents is entirely because of the additional quality of the superior good.

Now consider a second case in which μ equals zero. The superior detergent yields the same benefits as three boxes of the inferior detergent, but both varieties cost the same (\$1) because quality improvement does not drive up cost. In this situation, the quality-adjusted price of the superior item is actually one-third that of the inferior item because \$0.33 spent on the superior detergent buys as much cleaning power as \$1.00 spent on the inferior variety. Our decomposition now records a 1.0 on the left side. The first term on the right side is still 3.0, since the quality difference is the same as before. Now, however, the second term is 0.33 and the third term continues to be 1.0. In this second situation, the true rate of price

[3] Recall that Y^e is the equivalent quantity of X units embodied in one unit of the good Y. When the MRS between X and Y is $(1+\theta)$, $Y^e = (1+\theta)Y$ and $P^e_Y = P_Y/(1+\theta)$. Thus, it is P_Y/P^e_Y and not P_Y/P_X that equals the MRS. Indeed, when μ equals zero, P_Y/P_X equals one even though the MRS is $(1+\theta)$.

inflation is negative because, after the introduction of the new detergent, less money is needed to purchase the previous level of cleaning power.

These examples illustrate the role of the parameter μ when it assumes the extreme values of zero and one. More generally, the quality effect is $(1+\theta)$, and the pure price effect is the product of the last two terms, $[(1+\mu\theta)/(1+\theta)](1+\rho)$.

The practical relevance of this theoretical discussion arises because θ and μ are not directly observable and, as we will see below, all methods used to construct CPI implicitly assume that μ is equal to one. This assumption may be correct, but if it is wrong, a bias will result. The extent of the bias depends on which procedure is actually used to estimate the quality component of price. I will consider this issue after I describe the outlines of the current CPI program. For now, it is sufficient to observe that the problem of sorting out the separate contributions to the change in product price is quite difficult since two of the key elements on the right side of Equation 1, $P^e_Y(t)$ and $P_X(t)$, are not observed in the marketplace in the long run when μ has a value less than one.

New Goods

The price decomposition described above is derived assuming the MRS between the new model of a good, Y, and its predecessor, X, is constant at $1+\theta$. This presumes that the new model can be fitted into the existing categories of goods and services. Although this is often the case, at times an entirely new good appears that cannot be considered a new model of some existing product. Automated teller machines fit this definition, as do personal computers and television. When this happens, the implicit marginal rate of substitution cannot be assumed to be a constant $1+\theta$ but must be allowed to vary as it does between any two independent goods.

From a practical standpoint, no natural reference prices and quantities [i.e., no clear choice of X for the denominator of the ratio $P_Y(t)/P_X(t)$] exist. The discussion in the preceding sections therefore does not apply. In this case, the correct treatment of new goods in a cost-of-living index involves the use of reservation prices. This is the price of Y at which zero quantity of the new good would have been demanded had it been available. It is the reference price for the inclusion of the new good. If, for example, the new good becomes available at a price of $1 per unit, but demand is such that the reservation price is $2 (the point at which no one wants to buy any of the good), the relevant price for the quality calculus is $2, not the observed price of $1.

The new-goods story is more complicated, but most of the considerations in the preceding sections apply. Cost elasticity enters the decomposition of quality and pure price change. Now this decomposition is more complicated, however, by virtue of the consumers' surplus generated by the new good. This surplus must be included when measuring the welfare consequences of the new good. There is a welfare gain to society in this case, even when the cost elasticity equals one.

The new-goods approach has been implemented for breakfast cereals by Hausman (1997). He finds that the CPI overstates the true cost-of-living index for this category of goods. There is a debate, however, about the use of the market demand curve versus the weighted average of individual demand curves in estimating the reservation prices.[4] I will not pursue these issues in this article; however, they are a potentially important source of bias in current CPI.

PRICE AND QUALITY IN THE CPI

Structure and Objectives

The CPI's working objective is to measure the inflation in prices relative to a *fixed* market basket of goods and services. The rationale for this approach is straightforward: If it costs more to purchase the same goods today compared with last year, prices must have gone up. A simple mea-

[4] See Fisher and Griliches (1995).

Table 1

Relative Importance of Substitutions in the Construction of the CPI*

	1984 (%)	1995 (%)
Link method	1.71	0.57
Overlap method	0.23	0.05
Direct qual adj.	0.30	0.41
Class-mean method	N/A	0.32
Total noncomparable	2.25	1.35
Total comparable	1.70	2.54
Total substitutions	3.95	3.90
Nonsubstitutions	96.05	96.10
Total covered CPI	100.0	100.0

* The price quotes for several item categories are excluded. (See text for list of items, most important of which is residential rent). The class-mean method was not used in 1984.

SOURCE: Moulton and Moses (1997). Total may not equal 100 percent because of rounding.

sure of average prices can be obtained by comparing the expenditure needed to purchase the same items year after year. The implementation of this measure is, however, neither simple nor straightforward. Some 70,000 to 80,000 price quotations on goods and services are collected on a monthly basis from about 21,000 outlets (supplemented by housing market information from 40,000 landlords and renters and from 20,000 homeowners). This information is aggregated in two stages. In the first stage, a Laspeyres index is used to group data into 207 item categories in 44 areas. In the second stage, a Laspeyres index, with weights obtained from the Consumer Expenditure Survey, is used to aggregate the "market basket" into the CPI. The weights are changed periodically.

The objective of measuring price inflation with respect to a fixed market basket is not possible in practice because some items are discontinued or are otherwise not available when the BLS agent goes to a retail outlet. According to BLS data summarized in Table 1, the number of such cases tends to be small in any one month (around 4 percent per month), but it becomes quite large over the course of an entire year. When a discontinuation occurs, another item is introduced as substitute into the index (e.g., Aspirin is replaced by Ibuprofen). This substitution leads to a complicated set of steps, the possibilities of which are indicated in the left column of Table 1. First, replacement items are deemed to be either *comparable* (about 43 percent of items in 1984, but 65 percent of items in 1995), or *non-comparable*.[5] If the new item is comparable, its price can be used in place of the old, and any price increase is treated as inflation.

In the case of noncomparability, the BLS must consider that the new items introduced into the market basket may be systematically superior or inferior to the goods they replace. When new goods appear in the market that satisfy old wants in a better way, or when improvements in old goods lead to greater utility, some of the total price increase must be attributed to quality (as in the price decomposition in Equation 1). The BLS handles this problem in one of three ways.

The *first* option is to estimate the quality component directly. This is done by means of either a price-hedonic regression or an estimate of direct production costs, adjusted for profit margins. Price hedonics are used mainly for housing and apparel. The production-cost approach is mainly for autos, but other items are treated this way as well. In either case, the resulting estimate of θ can be used to strip the quality component out of the "raw" change in price.

The *second* option is to look for models closely related to the missing item (X) and to compare the overlap in the price of the surrogate X to that of the new item (Y) in the same month. In terms of the preceding discussion, this overlap method amounts to estimating the quality component of a price change by using the contemporaneous price ratio $P_Y(t)/P_X(t)$. This method presumes that X remains in the market so that its price can be observed.

5 The estimates of Table 1 are derived from Moulton and Moses (1997). They also provide estimates for the year 1983, but these estimates are omitted from our Tables 1 and 2 for clarity of exposition. They are included in the "bottom-line" estimates of Table 5.

The *third* option is the "link" method, also called the "deletion" method. As Armknecht and Weyback (1989, p.109) explain, "The price change between two observations is imputed as the average change for other items in that item stratum and geographic area." In other words, this method is basically an indirect way of estimating the denominator of the ratio $P_Y(t)/P_X(t)$ that bears no direct relation to the good X. Since 1992, a variant of the link method has also been used; in this method, the average change for related items in the relevant stratum, rather than the average of all items, is used to impute the price change. This variant is termed the "class-mean method."

In the early 1980s, around three-quarters of the noncomparable substitution cases were handled through the link method. As more reliance was placed on the class-mean method and on direct quality measurement, this proportion had diminished to about two-fifths by 1995. The proportion of cases handled through direct estimates of quality change increased from 13 percent to 30 percent.

Table 2 reports the results of the price-quality decomposition for each of the methods BLS uses. A comparison of the seventh and eighth rows of this table reveals that item substitutions occurring during the repricing process account for a disproportionately large share of the total change in the CPI for the periods in question. In 1984, 96 percent of the growth in this part of the CPI came from the 4 percent of items that were substitutions.[6] The proportion fell to 50 percent in 1995, but it is still very much a case of the "tail wagging the dog."

The statistics in Tables 1 and 2 refer to the repricing process. New goods can enter the CPI market basket other than as substitutions. The items in the market basket are subject to sample "rotation," a process in which approximately 20 percent of the sample is redrawn to include new outlets, stores, and so forth. The new sample may contain items that were in the previous market basket, but it also includes new items that did not appear in the old sample. There is, however, no direct matching of old and new items. Moreover, new items are entered in such a way that the CPI is unchanged, thus missing the consumer surplus generated by the introduction of the new good. Only subsequent changes in the price of the new goods lead to changes in the CPI.

Table 2

Relative Importance of Quality and Pure Price Effects in the Construction of the CPI*

	1984 Quality (%)	1984 Pure Price (CPI) (%)	1984 Total (Raw) Price (%)	1995 Quality (%)	1995 Pure Price (CPI) (%)	1995 Total (Raw) Price (%)
Link method	0.99	0.16	1.15	0.99	0.02	1.01
Overlap method	0.14	1.38	1.52	0.00	0.10	0.10
Direct quality adj.	0.10	0.35	0.45	0.11	0.19	0.30
Class-mean method	N/A	N/A	N/A	0.66	0.18	0.84
Total noncomparable	1.23	1.89	3.12	1.76	0.49	2.25
Total comparable	0	1.37	1.37	0	0.60	0.60
Total substitutions	1.23	3.26	4.49	1.76	1.09	2.85
Nonsubstitutions	0	0.14	0.14	0	1.07	1.07
Total covered CPI	1.23	3.40	4.63	1.76	2.16	3.92
Total CPI		3.90			2.50	

* The price quotes for several item categories are excluded. (See text for list of excluded items.)

SOURCE: Moulton and Moses (1997). Figures are rounded.

A Critique of the CPI

The CPI program is constantly under review, and BLS staff has provided a steady flow of studies on various aspects of the index. These studies are a valuable source of information about the CPI and its problems, and most external critiques rely heavily on them. Moreover, both the BLS and external critics agree that the CPI has biases. The consensus of the external critics, summarized in Table 3, is that the upward bias is close to the Boskin Commission's estimate. The BLS itself has not produced an official estimate of the bias.

[6] Total Substitutions in Table 2 are 3.26 percent for 1984, while the "Total Covered CPI" estimate is 3.40 percent. Substitutions thus account for almost all the change in the CPI-U in this year. This is, however, one point on which the result for the year 1983 differs from 1984: Substitutions account for only 61 percent of the growth in the covered CPI in 1983.

Table 3

Recent Estimates of Bias in The CPI

Source	Point Estimate	Interval Estimate
Advisory Commission (1995)	1.0	0.7–2.0
*Advisory Commission (1996)	1.1	0.8–1.6
Boskin (1995)	1.5	1.0–2.0
Congressional Budget Office (1995)	___	0.2–0.8
Darby (1995)	1.5	0.5–2.5
Diewert (1995)	___	1.3–1.7
Gordon (1995)	1.7	___
Greenspan (1995)	___	0.5–1.5
Griliches (1995)	1.0	0.4–1.6
Jorgenson (1995)	1.0	0.5–1.5
Klumpner (1996)	___	0.3–0.5
Lebow, Roberts, and Stockton (1994)	___	0.4–1.5
Pakes (1995)	0.8	___
Shapiro and Wilcox (1996)	1.0	0.6–1.5
Wynne and Sigalla (1994)	1.0<	___

SOURCE: Moulton (1996), Table 1 with addition of Advisory Commission (1996). References can be obtained from Moulton (1996).

The suspected bias in the CPI comes from several sources (Table 4). The very fact that the items in the CPI market basket are meant to remain fixed over time is one source of the problem. Such simplicity is a virtue, but it also has a defect: Because the prices of some items rise more rapidly than other prices, consumers are unlikely to buy the same mix of goods; they will instead shift spending away from the items that are becoming relatively more expensive. Because the CPI tracks the cost of the fixed market basket, it tends to overstate the cost increases consumers actually experience. The term for this phenomenon is "substitution bias." At the lowest level of CPI aggregation, the substitution bias is augmented by a possible bias arising from the formula used in the aggregation, but substitution bias also arises in the second stage of CPI aggregation, and the combined magnitude of these two effects is 0.40 of a percentage point, according to the Boskin Commission. New outlet bias—in which new outlets provide cheaper or more expensive access to items in the CPI market basket—account for another 0.10 of a percentage point. Examples of such new outlets would be discount stores and boutiques.

The rest of the Boskin Commission's total CPI bias comes from underestimating the effects of quality changes and new items (0.6 of a percentage point). Various aspects of the CPI program may lead to these outcomes. First, the objective of the CPI—the pricing of a *fixed* market basket of goods and services—works against the rapid incorporation of items with quality-enhancing innovations. New goods do enter the market basket during the repricing process and during sample rotations, but there is no active search for such items and thus no systematic procedure for identifying innovations as they occur. Moreover, only one-fifth of the sample is rotated each year, so those new goods which do come into the market basket do so with a time lag. Shapiro and Wilcox estimate that, under current sample-rotation procedures, new items attain only 40 percent of their steady-state representation in the CPI index after four years from introduction. And, since new goods enter the market basket without changing the level of the CPI at the time of entry, quality improvements in the "rotated" items tend to be missed.

Another bias may arise from the way judgments are made about item substitutions. When an item is missing from the market basket, BLS commodity analysts must determine whether a new item is comparable to an old item. There are guidelines for this purpose and cases are reviewed, but a considerable amount of individual judgment is still involved. In this regard, it may not be easy for agents to spot the true extent or nature of a model change or to determine what changes are cosmetic (or stylistic) and what changes are substantial. This ambiguity may lead to conservative judgments, but the direction of the bias is not clear. It may indeed be more difficult to spot an erosion in quality than to observe positive improvements, since producers have an incentive to conceal the former and advertise the latter.

Attempts to quantify these potential problems have led to the "consensus" estimates reported in Table 3. However, this consensus is based on evidence from a field in which the state of research is fairly unsettled. As Griliches (1995, p. 129), a member of the Boskin Commission, has observed, "[T]he Committee [on Finance] assumes that we already know that the CPI is overstated. I believe that the CPI does indeed overstate the average price rise during the last several decades, 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." Moulton (1996, p. 160), who cites this remark by Griliches, goes on to observe that "for some of the sources of bias, the evidence is based on case studies of a small number of commodities. The differences between estimates [i.e., CPI versus case study] seem to be largely determined by the willingness of experts to extrapolate from these case studies to estimates for broader categories of goods." Moulton (p. 160) also argues that "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, like computers, prescription drugs, and so on."

The Boskin Commission recognizes the uncertainty involved in its synthesis by providing the range of possible biases (0.8 to 1.6) in addition to its best estimate of 1.1 percent. Whether or not the Commission's range is broad enough to address the issues raised by Griliches and Moulton is unclear, given the complexity of the problem. However, it seems clear that there is room for more analysis.

With these issues in mind, in the next sections I look at the bias problem from a different angle. Rather than starting with a comparison between the CPI and other price estimates that treat quality change differently (and are deemed to be superior) or with an enumeration of the problems that afflict the CPI, I follow Triplett's (1990) suggestion and ask how much quality change is already taken into account in the CPI. I then examine the bias issue in light of the answer. This alternative approach would be unnecessary if reliable price estimates for the broad range of items covered by the CPI were available from other sources. But if such estimates actually existed, there would be little need for a CPI program in the first place.

Table 4

Components of the CPI Bias

Source of Bias	Advisory Commission	Shapiro-Wilcox
Upper-level substitution bias	0.15	0.20
Lower-level substitution bias	0.25	0.25
New-outlet bias	0.10	0.10
New products	} 0.60	0.20
Quality change		0.25
Total	1.10	1.00

SOURCE: Advisory Commission (1996) and Shapiro and Wilcox (1996).

ALTERNATIVE ESTIMATES OF CPI QUALITY BIAS

Quality Change in the CPI

The extent to which the BLS already adjusts for quality differences is surprising in light of the current debate. The price quotes gathered by BLS agents in 1995 indicate a raw, unadjusted price change of 4.72 percent for the items studied by Moulton and Moses (1997). Yet the change in the CPI for these same items was only 2.16 percent. The difference between these numbers is the extent of the adjustment, which Moulton and Moses attributed to "quality" change in the original version of their estimates. They subsequently attributed 0.80 percentage points of their adjustment to changes in the units of measurement, leaving the quality change estimate of 1.76 percent reported in Table 2 of this paper.[7]

[7] The 0.8 percent adjustment for the change in the units of measurement would seem, at first glance, to be too large. Surely there are many instances in which smaller packages (quarts of milk) are substituted for larger packages (gallons) as the other way around. The net size of the net adjustment should therefore have an expected value of zero. However, there is a "formula bias" because of the procedures used to compute percentage changes and this leads to a positive bias.

These estimates cover 72 percent of the items in the CPI-U version of the market basket. The largest component of the missing 28 percent is rent and the rental equivalent for residential housing (25 percent).[8] There is no formal estimate of the quality change imputed to these items in the Moulton-Moses study, but informal discussions with BLS officials suggest that an estimate of 0.5 percent to 1.0 percent per year for housing would be reasonable. Combining these estimates yields an estimate of overall quality of around 1.41 percent. The overall CPI-U (including housing) grew 2.50 percent in 1995, so the overall estimate of quality accounts for about one third of the total raw price change.

A quality change of 1.41 percent may not seem like much in any one year. However, this annual rate implies that an individual who consumed $50,000 worth of goods and services in both 1994 and 1995 was $700 better off in 1995 by virtue of quality improvements in the market basket alone (after adjustment for pure price inflation). The 1.41 estimate also implies that the average quality of the items in CPI would *double* after five decades. And, if the Boskin Commission is correct about this being an under-estimate of the quality improvement in the consumer market basket, another 0.6 percent must be added to the BLS estimate. If product quality improves at a compound rate of 2.01 percent per year, the person consuming $50,000 is better off by almost $1,000 in 1995 than in 1994.

These figures suggest a very high annual dividend from the introduction of new goods and the improvement of existing goods. Most observers seem convinced that the true magnitude of the quality dividend is large, but is it really this large? Or is it possible that some BLS procedures contain additional biases, biases that overstate the true rate of quality improvement? We already know that the biases identified by the Boskin Commission operate in the opposite direction. Given the unsettled state of research on the quality issue, the possibility that some procedures are biased toward overstatement of quality cannot not be dismissed out of hand.

[8] According to Moulton and Moses (1997) p. 48, "Excluded from the studies were price quotes for residential rent and homeowners' equivalent rent (all years), and household insurance, postage, babysitting, and care of invalids (1995) within Housing; used cars (all years) and automobile finance charges (1995) within Transportation; health insurance (all years) within Medical care; magazines, periodicals and books (1983 and 1984) and sports vehicles, including bicycles (1995) within Entertainment."

Link Bias

A closer look at Table 2 reveals that the link method has the practical effect of attributing a very large fraction of the raw price change to quality: 86 percent in 1984 and 99 percent in 1995. Moreover, in 1995 the more precise class-mean variant of the link method (not used in 1984) attributed 78 percent of the raw price change to quality. In contrast, the direct quality-adjustment method attributes a smaller share of the raw price change to quality: 22 percent in 1984 and 37 percent in 1995. This result may simply reflect inherent quality differentials in the items handled with the various methods. It is important to recall, however, that the link method is used only when all else fails; it does not provide a direct estimate of quality change. Its accuracy is therefore called into question when it gives a very different answer from those produced by other methods.

A closer look at the nature of the link method heightens these suspicions. When a BLS agent goes to a retail outlet and finds that an item (for example, aspirin) is missing, the agent first determines whether the item has been discontinued or is temporarily unavailable. If it is temporarily unavailable, a price is imputed and altered when the item reappears in the outlet. On the other hand, if the item has been discontinued, the agent uses a checklist to select a substitute item for pricing (another brand of aspirin, if available, or a related item like ibuprofen, if not). This substitute item is reviewed by a BLS commodity specialist, who determines whether it is comparable to the old item. If the items are comparable, any difference is price is treated as a pure price change. If they are not comparable and direct quality adjustment cannot be used (the typical case), the link or class-mean methods are employed. Pure inflation is assumed to equal the average rate of price increase among similar items and any residual price change is implicitly attributed to quality improvement. Thus, if the price of a package of aspirin was $2.00 last month and the substitute package of ibuprofen costs $3.00 this month, and there was no price change in other related pharmaceuticals, the whole $1.00 price differential would be treated as a quality increase.

This may well be a reasonable outcome. The shift from aspirin to ibuprofen, for example, represents a change in the average quality of the items in the CPI market basket. The link and class-mean methods correct for this change in sample composition and may lead to a large quality adjustment if old and new items are dissimilar. Application of the direct quality-adjustment method would, in principle, give the same result. Moreover, the failure to adjust the CPI for changes in sample composition would lead to the wrong estimate of price inflation.

There is, however, a potential problem when the link method is applied to discontinued items. Some items are discontinued because they are replaced by items of superior quality. In this situation, the link and class-mean approaches may give the wrong answer because, as Reinsdorf, Liegey, and Stewart (1996, p. 1) observe in their study of apparel and autos, "manufacturers often time real price increases to coincide with product redesigns or the introduction of new models." These real price increases will tend to be confused with quality improvements in the link and class-mean methods, with a resulting overestimate of the quality effect.

A comparison of the link and class-mean methods provides additional evidence of bias in the link method. Table 2 shows that, whereas the link method attributed virtually all of the raw price change to quality improvements in 1995, the class-mean procedure attributed 20 percent less of the raw price to improvements in quality.

These considerations led Moulton and Moses (1997) to reconsider their 1995 estimates of 1.76 percent for quality change. Since the 1.76 percent estimate was greatly affected by a relatively small number of items with large price changes, they argued that it made sense to trim or truncate the largest quality effects in each direction. When items with quality effects larger than a 400 percent increase or 80 percent decrease are dropped from the calculation, the estimate of quality change falls from 1.76 percent to 1.10 percent. With a tighter truncation rule—a 100 percent increase or 50 percent decrease in measured quality—the estimate of quality change falls from 1.76 percent to 0.54 percent.

These truncation procedures provide a conservative estimate of quality change. However, the truncated estimates are in line with estimates obtained from a different way of measuring the quality bias. Let us suppose that the true proportion of price that is affected by quality changes in items currently handled through the link and class-mean methods is actually the same as the proportion of quality change under the direct quality adjustment method. If BLS could apply direct quality adjustment to these items, by how much would the link and class-mean estimates of Table 2 be altered?

The answer to this question is given in the top panel of Table 5. When the arithmetic is done, the imputed direct-

Table 5

The Effects of Link Bias and Quality-Cost Bias Implicit in the BLS Estimates of Quality Change

("True" value minus Table 2 value, in percent)

	1983 (%)	1984 (%)	1995 (%)
1. Link-bias only	–0.54	–0.73	–0.97
2. Quality-cost bias only			
Cost elasticity equals 1.00	0	0	0
Cost elasticity equals 0.75	0.37	0.41	0.59
3. Both biases combined			
Cost elasticity equals 1.00	–0.54	–0.73	–0.97
Cost elasticity equals 0.75	–0.36	–0.57	–0.71

SOURCE: Table 2 and my calculations.

quality adjustment value of quality change is less than the corresponding value in Table 2 for each of the years studied. This difference (the "link bias") has grown over time from –0.54 percent in 1983, to –0.73 percent in 1984, and to –0.97 percent in 1995.[9] In other words, if the proportion of price change attributable to quality change in the items currently handled through the link/class-mean method were actually the same as for the items treated with the more direct method, the rate of quality change in 1995 would have been 0.79 percent rather than 1.76 percent. Note that the 0.79 estimate is halfway between the high and low truncation figures of Moulton and Moses (1997).

The case against the link method is circumstantial but nonetheless compelling. It seems reasonable to conclude that current CPI procedures overstate quality change, and that the magnitude of this overstatement is potentially large. Some of this overstatement may arise from pure measurement error, while another part can be attributed to price increases hidden in the quality component. If the hidden price increase were only half of the 1995 link bias of –0.97 percent, the "true" CPI would have grown by 0.485 percent more than the actual CPI.

This hidden price bias operates in the opposite direction of new goods and quality biases identified by the Boskin Commission. Its magnitude is potentially large enough to cancel much of the Commission's estimate (0.60). Moreover, if the hidden price increase were equal to the entire link bias (–0.97), the entire Boskin Commission bias is virtually eliminated. These are not likely outcomes, but they do argue for lowering the bottom of the Boskin Commission range of possible biases from 0.8 toward zero.

The μ Problem: Another Bias

The section on modeling quality change ended with the warning that a bias could result from the failure to establish the correct value for cost elasticity. There, I noted that the presence of the cost-elasticity parameter poses no problem per se for the quality-price decomposition so long as its magnitude is known. But the empirical magnitude of this parameter has never been established, nor can its theoretical value be readily determined. In the absence of the required information or, indeed, general recognition of the problem, the BLS proceeds under the implicit assumption that the value of the cost elasticity is always one.

This assumption is built into each of the imputation methods used in Table 2 to handle noncomparable substitutions. The overlap method, for example, compares the price of a substitute model Y in period t with the price of its predecessor in the same period, $P_Y(t)/P_X(t)$. The entire price difference is attributed to quality change. Remember, this is accurate only if the cost elasticity equals one. If the cost elasticity is actually zero, the overlap price ratio equals one, and no quality change will be attributed to the new variety Y. In general, the overlap method is thus biased whenever μ is less than one.

Similar remarks apply to the direct quality-adjustment method, which comprises the price-hedonic and the direct estimation of cost approaches. The price-hedonic regression may be homogeneous of degree μ, as shown in Hulten (1996), and price hedonics may therefore yield an underestimate of quality change when μ is less than one. As for the cost-estimation method, it treats the entire increase in the production cost of Y relative to X as being the result of quality change. This also leads to an underestimate when the cost elasticity is less than one.

The results of the link method are also affected by the μ problem. In the link method, a surrogate for the price of X must be found in order to construct the ratio $P_Y(t)/P_X(t)$ because X cannot be observed. Even if this proxy were accurate (i.e., there were no link bias), a bias would arise if μ were less than one, for the same reason that a bias occurs in the overlap method.

Potential Bias Magnitudes

How large might the quality-cost bias be? The answer to this question depends on how far below 1.0 the true value of the

[9] The link bias is computed in the following way. In 1984, the average proportion of quality change in items handled through the direct quality adjustment method was 22 percent (0.10/0.45). If this proportion is applied to the raw price change of 1.15 percent for linked items, the quality component associated with the link method becomes 0.26 rather than 0.99, hence an estimated bias equal to –0.73 percent.

cost elasticity lies—an interval that is difficult to judge on theoretical grounds alone, because a case can be made for the proportionality of cost and quality (a unitary elasticity). If the production function for quality is continuous, profit maximization will lead producers to offer that level of quality for which marginal revenue equals the marginal cost of quality. This mechanism drives the incremental cost of quality into line with the benefits (price) and causes prices to move in proportion to quality changes. And if the production function exhibits constant returns to scale, an indeterminately large number of varieties of different quality will result.

However, not all quality changes are achieved with proportional increases in cost. The mechanisms of innovation often lead to non-proportional variations between cost and quality. One such mechanism is knowledge spillover inherent in the process of research and development. Spillovers cause the social return on R&D expenditure to exceed private returns and the benefits to society of obtaining better goods and services are therefore greater than the private costs of innovation. Moreover, innovation is often a discontinuous process, in which inspiration and unexpected opportunities are important, and spending cannot be pushed to the point where costs exactly match expected benefits. Some innovation costs may be close to zero: Griliches (1964) gives the example of a birth control pill (Enovoid) whose dosage had been cut in half, thereby cutting the cost per dose by half.

Because the cost-elasticity formulation is new on the price measurement scene, no reliable empirical estimates of μ exist. A brief survey of existing price hedonic studies reported in Hulten (1996) suggests that the average value of cost elasticity is closer to one than to zero. The studies under review, however, were not designed to measure this parameter, and the resulting evidence about the elasticity is therefore very uncertain.

I use a value of 0.75 as a reasonable alternative to 1.00 to illustrate the potential size of the bias. The resulting estimates in Table 5 are no more than another set of guesstimates. The default assumption of a unitary elasticity is shown in the first row of the middle panel of Table 5. Here, the biases are zero when $\mu = 1$ because this is the assumption already built into the calculations. On the other hand, a value of 0.75 implies that the quality estimates in Table 2 are too low and must be raised by one-third to arrive at the true value. The resulting estimates of the implied bias are positive, and they increase over time. They are also quite large, with a value in 1995 of 0.59 percent. This figure is even larger—1.76—if the value of μ is changed to 0.50 from 0.75. Estimates of this magnitude could significantly extend the top of the Boskin Commission's range of CPI bias.

These estimates assume that the bias in the link method is zero. The bottom panel of Table 5 provides alternative calculations that include the *joint* effects of link and quality-cost bias. The salient result, here, is that the combined bias is negative for all years. This outcome suggests that the net effect of both biases is to overestimate quality change.

I must emphasize that the calculations in Table 5 suggest the possible magnitudes that are at stake in the CPI debate, but these calculations should not be mistaken for the real thing. The total bias obviously depends on the magnitude of a variable—the cost elasticity—whose magnitude is simply not known with any precision.

CONCLUSION

The following points about the treatment of "quality" in the CPI have been raised in this article:

- The CPI currently embodies a large adjustment for quality, although it is not clear how much of the adjustment is the result of improvements in product quality over time.
- The link method may overstate quality change and understate the CPI to a significant degree and may extend the lower end of bias ranged identified by the Boskin Commission to zero.

- The neglected cost dimension of quality change operates in the opposite direction from the link bias and may extend the top of the Boskin Commission's bias range beyond 1.6 percent. The link bias appears to dominate the quality-cost bias, but much more needs to be known about both effects.

The second conclusion has an important implication for policy. Though not likely, it may turn out that, if all the faults in the CPI (not just the faults identified by the Boskin Commission) were fixed, the net effect would be a bias estimate closer to zero than to 1.1 percent. In this situation, the much hoped-for reduction in the federal budget deficit and the improved fiscal health of the Social Security program may be less than anticipated. Congress may, of course, achieve these fiscal results anyway by mandating that annual increases in the CPI be reduced by the Boskin Commission's mean estimate of 1.1 percentage points (or some fraction of 1.1). This is a poor idea even without the uncertainties I have raised here, but it is a *decidedly* bad idea if the true bias turns out to be near zero. The right way to deal with problems in any measurement program, including the CPI, is to address each problem in its own right and apply the appropriate fix. No substitute exists for a real commitment to accurate measurement.

REFERENCES

Advisory Commission to Study the Consumer Price Index. *Toward a More Accurate Measure of the Cost of Living*, Final Report to the Senate Finance Committee, December 4, 1996.

Armknecht, Paul A., and Donald Weyback. "Adjustments for Quality Change in the U.S. Consumer Price Index," *Journal of Official Statistics* (Vol. 5, no. 2, 1989), pp. 107- 23.

Fisher, Franklin M., and Zvi Griliches. "Aggregate Price Indices, New Goods, and Generics," *Quarterly Journal of Economics*, (February 1995), Vol. 110, No. 440, pp. 229-44.

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

Greenspan, Alan. "Prepared Statement" in *U.S. Senate Committee on Finance, Consumer Price Index: Hearings Before the Committee on Finance, United States Senate*, One Hundred and Fourth Congress, first session. Washington: U.S. GPO: 1995, pp. 109-15.

Griliches, Zvi. "Notes on the Measurement of Price and Quality Changes," in *Models of Income Determination*, Studies in Income and Wealth, Vol. 28. Princeton University Press, 1964, pp. 381-404.

_______. "Prepared Statement," in *U.S. Senate Committee on Finance Consumer Price Index: Hearings Before the Committee on Finance, United States Senate.* One Hundred and Fourth Congress, first session. Washington: U.S. GPO: 1995, pp. 129-32.

Hausman, Jerry A. "Valuation of New Goods under Perfect and Imperfect Competition," in Timothy F. Bresnahan and Robert J. Gordon, eds., *The Economics of New Goods, Studies in Income and Wealth*, Vol. 58, Chicago: University of Chicago Press, 1997, pp. 209-37.

Hulten, Charles R. "Quality Change in Capital Goods and Its Impact on Economic Growth," NBER Working Paper 5569, May 1996.

Moulton, Brent R. "Bias in the Consumer Price Index: What is the Evidence?" *Journal of Economic Perspectives* (Fall 1996), Vol. 10, No. 4, pp. 159-77.

Moulton, Brent R. and Karin E. Moses. "Addressing the Quality Change Issue in the Consumer Price Index." *Brooking Papers on Economic Activity*, forthcoming.

Reinsdorf, Marshall, Paul Liegey, and Kenneth Stewart. "Avoiding Downward Bias from Linking in the U.S. Consumer Price Index," Bureau of Labor Statistics Working Paper 276, April 1996.

Shapiro, Matthew D., and David W. Wilcox. "Mismeasurement in the Consumer Price Index," NBER Working Paper 5590, May 1996.

Smith, Adam, *An Inquiry into the Nature and Causes of the Wealth of Nations*, 1776 (Richard D. Irwin Edition, 1963).

Triplett, Jack E. "Hedonic Methods in Statistical Agency Environments: An Intellectual Biopsy," in *Fifty Years of Economic Measurement: The Jubilee of the Conference on Research in Income and Wealth, Studies in Income and Wealth*, Vol. 54, Ernst R. Berndt and Jack E. Triplett, eds. The University of Chicago Press, 1990, pp. 207-33.

_______. "Concepts of Quality in Input and Output Price Measures: A Resolution of the User Value-Resource Cost Debate" in Murray F. Foss, ed., *U.S. National Income and Produce Accounts: NBER Studies in Income and Wealth*, Vol. 47. Chicago: University of Chicago Press: 1983.