Journal of Economic Perspectives—Volume 17, Number 1—Winter 2003—Pages 3–22

The Consumer Price Index: Conceptual Issues and Practical Suggestions

Charles L. Schultze

Committees of economists have long recommended that the U.S. Bureau of Labor Statistics (BLS) should establish a cost-of-living index as the measurement objective for the Consumer Price Index, rather than regarding the CPI as measuring changes in the cost of purchasing a fixed basket of goods, and should undertake the research and operational changes necessary to move the CPI closer to that goal. A committee led by George Stigler made this recommendation back in 1961 (Stigler, 1961). The Boskin commission, appointed by the Finance Committee of the U.S. Senate, echoed this recommendation in its 1996 report (Boskin et al., 1996).[1]

The Consumer Price Index was traditionally based on the concept of measuring the change in a household's cost of purchasing a fixed basket of goods and services in the face of a change in prices between two periods—in shorthand, a cost-of-goods index. A more ambitious objective is to base the index on the concept of measuring the change in the cost of maintaining a household's standard of living at some specified level—a cost-of-living index. In an aggregate CPI, price and expenditure data must be combined to produce an estimate that reflects some

[1] For a discussion, critique and comments on the Boskin et al. (1996) commission report in this journal, see the "Symposium on Measuring the CPI" in the Winter 1998 issue. The Boskin commission estimated that the CPI was overstating the rate of inflation that would be shown in a true cost-of-living index by about 1.1 percent a year. The report attributed 90 percent of this bias to three sources: failure of the index to take into account consumer substitution behavior in the face of relative price changes, and inadequate allowance for improvements in the quality of consumer goods and for the introduction of new goods.

■ *Charles L. Schultze is Senior Fellow Emeritus, Brookings Institution, Washington, D.C. He served as Chair of the Panel on Conceptual, Measurement and Other Statistical Issues in Developing Cost of Living Indexes convened by the National Research Council.*

measure of average change, in either the cost of goods purchased or the cost of maintaining a given living standard for all or for some subgroup of households. But the aggregation of index numbers over the population or over groups is not an issue that separates cost-of-living and cost-of-goods indexes.

Shortly after the Boskin report was issued, the Bureau of Labor Statistics told Congress that it had, in fact, been using cost-of-living theory for some time to make decisions about the index and accepted the cost-of-living index as its measurement objective for the Consumer Price Index (U.S. Bureau of Labor Statistics, 1997). But it also pointed out that "the cost of living is a theoretical construct . . . not a single or straight-forward index formula readily amenable to practical use," and it noted that a wide range of issues had to be confronted in moving the CPI closer to a cost-of-living index.

The Bureau of Labor Statistics asked the Committee on National Statistics of the National Academy of Sciences to establish a panel charged to explore the conceptual and statistical issues that arise in constructing a cost-of-living index; to assess the advantages and difficulties involved in establishing the cost-of-living index as the measurement objective of the CPI; and to make recommendations to the BLS about its operational, data collection and research programs.

This article is not a summary of the panel's report (National Research Council, 2002).[2] Rather, out of a large number of issues that the panel treated, I have elected to discuss four that are both controversial and particularly important: 1) addressing the problem of consumer substitution among goods over time; 2) defining the universe of goods and services that should be included in the CPI; 3) adjusting the CPI to take into account quality changes in existing goods; and, briefly, 4) handling the introduction of new goods. While I will present the panel's recommendations and the reasoning behind them, I add my own commentary and interpretation and in some cases expand on the material in the report.

One feature of the panel that distinguished it from its predecessors—the Boskin commission and Stigler committee—was the inability of panel members to reach unanimous agreement that the cost-of-living index should be the measurement objective for the Consumer Price Index. The panel's report identifies and analyzes some theoretical and measurement difficulties that affect our ability to produce conceptually consistent and accurate measures of a cost-of-living index. Some members of the panel concluded that these limitations are serious enough to

[2] The members of the Panel on Conceptual, Measurement, and Other Statistical Issues in Developing Cost-of-Living Indexes were Charles L. Schultze (chair, Brookings Institution), Ernst R. Berndt (MIT), Angus Deaton (Princeton University), Erwin Diewert (University of British Columbia), Claudia D. Goldin (Harvard University), Zvi Griliches (Harvard University, died in 1999), Christopher Jencks (Harvard University), Albert Madansky (University of Chicago), Van Doorn Ooms (Committee for Economic Development), Robert A. Pollak (Washington University, St. Louis), Richard L. Schmalensee (MIT), Norbert Schwarz (University of Michigan) and Kirk Wolter (National Opinion Research Center). An executive summary is available on the web at ⟨http://www.nap.edu/books/0309074428/html/⟩. The text of the full report is also available at this site, although it is not convenient to read nor download the entire text from the website.

make it infeasible to convert into monetary terms the effects on living standards from changes in the prices and qualities of goods. Other members, myself included, came away convinced that despite these difficulties, it was, on balance, still desirable to aim the CPI at measuring a cost-of-living index, with the recognition that the measure will be an approximation. Despite our differences on this subject, however, the panel reached unanimous agreement on a wide range of recommendations to the Bureau of Labor Statistics aimed at improving the design and construction of the Consumer Price Index.

The Problem of Substitution

A cost-of-goods index starts with a certain basket of goods selected to be representative of total consumption expenditures during some particular point in time and then examines what it would cost to purchase this same basket of goods in a different period.[3] The Laspeyres version of the cost-of-goods index uses a basket of goods that represents the pattern of consumption at some time in the past and then measures what it would cost to purchase that basket of goods up to the present. A Paasche index uses a basket of goods that represents the pattern of consumption in the present, and then projects backward what it would have cost to purchase that basket of goods in some past period.

Because the Laspeyres index neglects the ability of consumers to mitigate the welfare effect of price increases through substitution among goods, it *overstates* the cost of maintaining the consumer's original, or reference period, standard of living. Conversely, because a Paasche index measures how much it would have cost in the past to purchase the basket of goods representing current consumption, it effectively assumes that people had already made the substitutions between goods in the past that they have now made in the present, and thus *understates* the cost of maintaining the consumer's current, or comparison period, standard of living.

These relationships have often led to the statement that a Laspeyres index should always exceed a Paasche index. Empirical studies have shown that Laspeyres indexes do indeed tend to produce a higher measured rate of inflation than Paasche indexes, at least in most years, and this evidence is often viewed as supporting the importance of substitution behavior in explaining changes in the pattern of consumer purchases. But notice that the Laspeyres index overstates the cost of maintaining the *reference* or past period standard of living while the Paasche index understates the cost of maintaining the *comparison* or current period's

[3] The index literature uses a standardized terminology to denote particular periods of time. The *base period* is the one during which the fixed basket of goods was purchased; price changes for particular goods are weighted by the base period quantities purchased. Thus for the official Consumer Price Index during the years 2002–2003, the base period currently is 1999–2000, and the weights are derived from the expenditure pattern in those years. Using the base period weights, price changes are measured between some beginning date called the *reference* period, for example, January 2002, and an ending date, labeled the *comparison* period, like June 2002.

standard of living. When those two standards of living are significantly different—due to perhaps to the size and pattern of relative price shifts or to changes in income—it is at least conceptually possible that a change in the Paasche index could exceed a change in the Laspeyres index.

Superlative Indexes

Clearly, the problem is to find an index that makes some allowance for substitution. If the demand functions for all goods were known, then this task is conceptually straightforward—but the demand functions are typically not known. However, Diewert (1976) showed that a class of "superlative" indexes existed which, under certain assumptions, could provide a close approximation to a cost-of-living index, reflecting the effect of consumer substitution behavior. Knowledge of specific demand functions was not required. These superlative indexes have the common feature that they involve some form of symmetric averaging and a weighting system that utilizes quantity or expenditure data from both the reference (beginning) and comparison (ending) periods covered by the index. One well-known superlative index is the Fisher index, which is calculated as the geometric mean of the Laspeyres and Paasche indexes. Another form of the superlative is the Tornqvist index, the geometric mean of the ratios of prices in the comparison and reference periods weighted by the average expenditure shares of the two periods.

The Consumer Price Index is built up in layers. At the lowest level, prices of individual items are collected and assigned to some 200-plus categories of goods and services, called "strata"—major appliances, televisions, household cleaning products and such. A price index is calculated for each strata. In turn the strata price indexes are combined into indexes for major expenditure categories—like food, apparel and housing—and finally into the overall consumer price index. In 1999, the Bureau of Labor Statistics began to use geometric aggregation of the prices of individual items to calculate about 60 percent of the 200–plus strata price indexes, as a way of approximately taking into account the effect of substitution behavior within those strata. In July 2002, the BLS began to publish a supplemental CPI that aggregated those strata indexes into an overall index using a Tornqvist superlative index technique.[4] However, superlative indexes face some practical difficulties and theoretical concerns.

A superlative index requires knowledge of consumer expenditure weights in both the reference and comparison periods. But acquiring information about current consumption is infeasible under current BLS data collection techniques, which rely upon consumer surveys for detailed expenditure data to use as weights. As a result, the final version of the BLS superlative index is only available after a two-year lag.

[4] The Bureau of Labor Statistics has labeled the new supplemental index the C-CPI-U, the Chained Consumer Price Index for All Urban Consumers. While that index is available in its final version only after a two-year lag, the BLS will publish initial estimates of the index in real time and preliminary estimates with a one-year lag.

Conceptually, the accuracy of the superlative in measuring the change in expenditures that is required to maintain a prespecified standard of living (usually that of the reference period) depends on several factors. If preferences are stable and also homothetic—that is, as income changes, consumers scale their purchases up or down proportionately—all changes in relative quantity weights between reference and comparison periods are due to substitution behavior in response to relative price changes. Only with homothetic preferences will the change in the cost of living brought about by a change in prices be independent of the standard of living or utility at which it is evaluated. But the empirical evidence (via estimates of Engel curves) shows that preferences are not homothetic; relative demands for goods vary as income and living standards change. In that case there are different cost-of-living indexes at different standards of living. And then as Diewert (2000) has shown, when the standard of living in the reference period is different from that of the comparison period, the superlative will measure a cost-of-living index that maintains a standard of living at some level intermediate between the two.[5] The same pattern of absolute and relative price changes can thus produce different superlative indexes depending on what happens to income and the standard of living over the period.

A similar result occurs when there are changes between the reference and comparison periods in one or more outside conditions, such as environmental pollution, the climate, the crime rate, and the provision of public goods. Such changes can alter consumers' preferences among market goods and services. For example, colder winters increase the demand for fuel oil, and higher crime rates raise purchases of home security systems and may lower the demand for downtown restaurant meals. In this situation, the superlative will again produce a cost-of-living index that maintains a standard of living somewhere between the reference and comparison period. Of course, tastes can change for other reasons as well.

Thus, to the extent that income and preferences among market goods change between the two periods, a superlative index loses some of its accuracy as a measure of the cost of maintaining the reference period's standard of living. The bottom line is that an assessment of how accurately a superlative index will in practice capture the effects of consumer substitution behavior depends heavily upon a judgment about the extent to which changes in the pattern of quantities purchased are driven by changes in income and tastes or by substitution responses to changes in relative prices.

The Substitution Issue and the Debate over a Cost-of-Living Index

The panel's discussion of superlative indexes as a way to address substitution problems illustrated a dynamic that emerged a number of times. Even when panel members disagreed on their overall evaluation of pursuing a cost-of-living index as

[5] Diewert (1976, Theorem 2.16) had earlier shown that, at least with translog utility functions, this intermediate standard of living will equal the geometric mean of the standards prevailing in the reference and comparison periods.

an objective for the Consumer Price Index, they found that they could still come to unanimous agreement about specific recommendations—albeit for different reasons.

For example, those who supported the notion that the Consumer Price Index should be designed as a cost-of-living index saw the use of superlative indexes as a step toward this goal. But even members who had reservations about the feasibility of the broad task of measuring the cost-of-living agreed that household substitution behavior moderated the effect of changes in prices on consumer welfare, and that the weighting and aggregation procedure in the superlative index would typically move the CPI in the right direction. Hence, when considering the appropriate index to use in "compensating" Social Security or other income recipients for the effects of price changes, those members joined the others in recommending that the Bureau of Labor Statistics publish a superlative index and that it be used as the basis for compensation payments.

The Appropriate Domain of the Consumer Price Index

The Consumer Price Index has always been confined to the universe of market goods and services. Consumers' standards of living, however, are affected by a wide range of other developments in the physical, social and economic environment, such as changes in the crime rate or the extent of environmental pollution. The panel concluded that restriction of the CPI domain to market goods and services was appropriate and desirable. Even though changes in outside conditions can affect consumers' living standards, a cost-of-living index can be defined that excludes those effects. The result is a "conditional" cost-of-living index, defined as the minimum expenditure ratio necessary to maintain a given standard of living in the face of changes in the prices of market goods and services, when the status of the excluded outside conditions remains unchanged.

Expanding the current universe of goods covered by the Consumer Price Index to include the effect on living standards arising from changes in outside conditions, including the benefits from the provision of public goods, would require analytical and measurement techniques that go well beyond the current state of the art. Some people have, nevertheless, argued that the BLS should undertake a long-range program of research to measure the effects on living costs from at least some of these outside conditions for the purpose of eventually including them in a cost-of-living index. The panel offered two responses to this point of view. First, even if reliable estimating techniques could be developed, a broader definition of a change in the cost of living would not be appropriate for most of the major uses to which the Consumer Price Index is currently put. Second, while additional research on the measurement of a broad standard of living is indeed worth pursuing, such a research effort should be part of a program to produce experimental "satellite accounts" that can supplement the current national income and product accounts. Any research program aimed at exploring the

effect of changes in outside conditions on consumer welfare should be carried out as part of this effort.

Particular Uses of the Consumer Price Index

The Consumer Price Index has a number of particular uses. It provides an overall measure of consumer goods inflation. In that role, it is used for indexing Social Security and other public benefits, as well as income tax brackets, against inflation. It also serves as an indicator for monetary policy. For these particular uses, a broad cost-of-living index that includes changes in living standards resulting from factors like decreases in crime or increases in traffic congestion does not seem appropriate.

By tying Social Security benefits and certain transfers to the poor and disabled to a price index, the Congress (perhaps without full awareness of the fact) did insulate the income of beneficiaries from many positive and negative supply shocks that affect the real wages of the working population, such as changes in productivity and fluctuations in the prices of oil and foreign exchange. But it is hard to believe that there would be significant public support to go well beyond that, distinguishing beneficiaries from other consumers by compensating or penalizing them for beneficial or harmful changes in, say, environmental pollution, the crime rate or the climate.

In the case of inflation, its economic costs all invoke some aspect of its monetary nature and its relationship to market transactions, like the effect of high and variable inflation in raising the risks associated with forward commitments and decreasing the efficiency of forward planning or the effects of inflation in producing unintended increases in marginal tax rates. There is, therefore, no reason we should want the Federal Reserve to tighten monetary conditions to produce deflation in the prices of market goods in an effort to offset the effects on an unconditional cost-of-living index from a worsening in the crime rate or an increase in congestion. Of course the Federal Reserve could, and surely would, strip out the environmental factors from an unconditional price index and look only at the changes in market prices in making its decisions about monetary policy. But that is simply a confirmation of one more use of the Consumer Price Index that would not be met by an unconditional cost-of-living index.

Satellite Accounts

While the members of the panel agreed that the Consumer Price Index should continue to be confined to the domain of private goods and services, we also recognized the potential usefulness of a research program aimed at supplementing the official measures of national output, income and prices with experimental estimates of the effect of various outside conditions on the material well-being of the population. But we concluded that this task should not be undertaken by the Bureau of Labor Statistics on its own, with the aim of producing a more comprehensively defined cost-of-living index. Rather, it should be pursued through the development of experimental measures of expanded national output and income

within an integrated national accounting framework. The reason is that such estimates will typically involve quite different types of analysis and estimation than that associated with the estimation of a cost-of-goods index or a cost-of-living index restricted to the domain of market goods.

For example, the real output of private goods and services is seldom measured by collecting data on physical quantities. Rather, observed nominal expenditure data—for example, the consumption components of the GDP—are deflated with appropriate price indexes (principally the components of the CPI) based on observed market prices, quality adjusted to the extent feasible. But in the case of outside goods, there are no nominal expenditure data to be deflated. Changes in both the "output" and the implicit "prices" of such goods, and in the changes of output and price, have to be estimated independently of each other. This task is especially complex because an important fraction of the "goods" (or "bads") currently excluded from the domain of the Consumer Price Index are intermediate goods, like public highways, the effects of acid rain on building materials and the economic costs of business crime. But the effects of changes in the quantity of such intermediate goods on private costs and prices are already reflected to some extent in the CPI and the value of private output. To avoid double counting, an accurate accounting of these goods and bads in a broad standard of living calculation would require that the intermediate effects be identified and excluded from measures of final consumption output and prices. This task would require a set of consistent national accounts within which to make the estimates.

In sum, research efforts aimed at estimating the effect of selected environmental changes on the nation's economic welfare must be imbedded in a consistent accounting framework that takes account of stocks and flows, output quantities and prices, and distinguishes intermediate from final goods. In the panel's view, this task is not a job to be carried out by the Bureau of Labor Statistics on its own with the aim of significantly broadening the domain of the Consumer Price Index.

Some Complications Created by a Conditional Cost-of-Living Index

All of the members of the panel agreed that the domain of the Consumer Price Index ought to be restricted to the universe of private goods and services. But the advocates of a cost-of-goods index argued that such a restriction raised difficulties for a cost-of-living framework. While an index can be constructed that ignores the direct effects of changes in the environmental factors we want to exclude—for example, the increased insecurity from a rise in the crime rate or any discomfort from a cold winter—changes in outside conditions can sometimes alter the demand pattern among private goods; for example, with a higher crime rate, purchases of downtown restaurant meals may fall while the demand for home security systems rises, and a cold winter raises the demand for heating oil. As a result, the superlative index, which in practice is how we measure a cost-of-living index, will reflect the effects on quantities of private goods purchased induced by changes in outside conditions as well as those made in response to changes in relative prices. The pragmatic question, again, is the extent to which such factors reduce the

accuracy with which the superlative index measures the conditional cost-of-living index that we want.

Specifying that the conditional cost-of-living index be limited to the universe of private goods and services does not automatically provide answers to another difficult set of questions in index design. For example, how comprehensively should we include the effects of changing technology on the standard of living? Arguably, we may want to allow the conditional cost-of-living index to be quality adjusted to reflect estimates of the welfare effects of improvements in identifiable medical procedures that reduce mortality and morbidity. But would we want to include in the cost-of-living index the effects of increases in longevity associated with broad and widely diffused changes in economic conditions and human knowledge like better sanitation, changes in dietary habits, or a higher standard of scientific knowledge? Similarly, how should we treat consumer benefits from the increased networking of information technology?

Because of considerations like these, some members of the panel concluded that conditional cost-of-living indexes are not well-defined because we have no theoretical procedures for deciding whether a particular quality change should be treated as a price change or as an "outside" factor to be conditioned out. On the other hand, other members concluded that such decisions have to be made on the basis of considering the purposes for which the index is to be used (in addition, of course, to measurement feasibility). As I noted earlier, the decision to restrict the conditional cost-of-living index to the domain of private goods and services stems from the need for an index that can be used, among other purposes, to measure inflation, compensate pensioners and index the tax code. Given that decision, the choice about how to treat advances in technology and human knowledge revolves importantly around whether those changes produce specific and measurable quality improvements in one or more private goods and services. While many social and environmental developments have broad effects on the well-being of households—for example, better dietary habits improve longevity—they would not be relevant for inclusion in the domain of a conditional cost-of-living index designed for the purposes mentioned above. And so, according to the supporters of the cost-of-living concept, the fact that the basis for such domain decisions cannot be provided from within the general theory underlying that concept is not a reason to preclude using the conditional cost-of-living index as the framework for the design and construction of the Consumer Price Index.

Quality Change

Given the magnitude of the potential effects of quality changes on consumers' living standards, and the progress that has been made in addressing the substitution issue, the most important challenge facing the Bureau of Labor Statistics at the present time is how to deal with quality changes. The problem with the traditional approach to dealing with quality change in the Consumer Price Index is not that

the BLS fails to make adjustments for such changes. In fact it does, and quite frequently. Rather, the problem is that the traditional techniques for adjustment may often mismeasure the effect of the quality changes on the price index.

The Traditional Approach to Quality Change

For decades, the Bureau of Labor Statistics has made two types of implicit quality adjustments: within-sample item replacement and sample rotation.

The process of within-sample item replacement arises because some 30 percent of the sample of items whose prices the BLS has been collecting as part of constructing the Consumer Price Index disappear from store shelves each year due to natural attrition. As a consequence, other items must be substituted. When a sampled item can no longer be found at a retail outlet, the BLS price agents are instructed to select the *most similar* replacement available within the store. In about two-thirds of the replacements, BLS commodity specialists judge the chosen substitutes "comparable"; they resemble the old good sufficiently so as to be treated the same for purposes of pricing. But one-third of the replacements are classified as "noncomparable"—sufficiently altered that some of the difference between the price of the old and the new good must be attributed to quality differences. In almost all such cases, BLS procedures have assumed that in the month of its introduction into the index, the "pure" price of the new good has risen by the same amount as the average price rise for similar goods. Any remaining difference between its stated price and the price of the good it replaces is therefore assumed to represent a difference in quality and is not counted as a price change.

In addition to the replacement of items that disappear from its current sample through attrition, the Bureau of Labor Statistics also rotates the overall sample; on average, about 25 percent of the BLS sample of items sold in retail outlets are replaced when new stores are rotated into the sample each year. Within each category of goods, the particular models or varieties to be priced in the newly sampled stores are selected to reflect current sales patterns. Since, on average, four years elapse before particular stores and items are replaced, the new sample will contain many items whose characteristics and features are different from those in the old sample. Price indexes for the new sample of stores and items are linked to those of the outgoing sample during an overlap period on the assumption that any contemporaneous differences in the prices of items are assumed to reflect differences in quality and do not enter the index as price changes.

These traditional methodologies for within-sample item substitution and for sample rotation can potentially mismeasure quality change in two important ways. First, as noted above, the quality difference between products with different characteristics is assumed to be exactly equal to their price difference. But when new varieties of products are introduced whose prices, after allowance for their improved quality to consumers, are lower than the prices of older varieties, this approach is likely to understate the value of the improvement and impart an upward bias to the index (Triplett, 2001, chapter 4). The very fact that newer varieties of items have displaced the ones that disappeared creates the presumption

that consumers have found these items to be cheaper, on a quality-adjusted basis, than those that disappeared.[6] In other cases, especially where fashion and fads play a large role, sellers apparently often use the occasion of introducing new models to raise prices on those models. Relative prices then gradually fall until another round of introduction occurs. Here, the current linking technique, which imputes the pure price change in a new model from the changes in prices of the older models, can understate that price change, overstate the implicit quality adjustment and tend to create a downward bias in the index.[7]

Hedonic Techniques

Hedonic techniques employ the assumptions that what consumers value in a good is the set of characteristics that the good possesses and that the analyst can identify and quantify those characteristics. Some examples include the following: screen size and surround sound in TV sets; speed, bytes of random access memory and hard drive capacity in computers; and type of fabric in a dress. In each of these products, the list of relevant and measurable characteristics is, of course, much larger than these few examples. By regressing the prices of different models of a product on measures of their characteristics, one obtains a relationship that explains the price of a product as a function of its characteristics.

Hedonic techniques can be used to make quality adjustments in the Consumer Price Index in one of two ways. First, in what is called the "indirect" approach, a hedonic equation can be fit over a cross-section of the different models or varieties of a particular product during some recent time period. Subsequently, when a noncomparable item is chosen as a substitute for one that has disappeared, the market value of differences in particular characteristics between the old and the new variety can be calculated from the coefficients of the regression and subtracted from the "raw" price change, leaving the residual as the "pure" price change. This indirect approach is the one currently used by the BLS for the hedonic applications it has incorporated into the construction of the CPI.

The alternative and more ambitious "direct" approach essentially treats the price of each variety of a line of products as an aggregate of the prices of its characteristics that are given by the coefficients in the hedonic regression.[8] To produce a monthly index for a given product with a number of quality-differentiated varieties, a hedonic equation is estimated each period. The reference and comparison period coefficients (the implicit prices of the various characteristics) are each weighted by the aggregate quantities of the characteristics of the items in the

[6] See Pakes (2002, pp. 4–5). However, the statement in the text needs to be qualified. Some minority fraction of consumers may still have preferred the older good that disappeared and have been willing to pay at or above the unit costs of producing them at the old volume, but the reduced volume wasn't large enough to cover the fixed costs of continuing their production.

[7] For some quantitative analyses of this phenomenon, see Moulton and Moses (1998) and Triplett (1997).

[8] There are several variations of the direct approach, but for purposes of this discussion, I have concentrated on the one described in the text.

reference period and then averaged. The ratio of the two averages produces a quality-adjusted index for the product line. However, the refitting, review and application of hedonic equations each month for timely incorporation into that month's Consumer Price Index would impose stringent requirements on the data collection and operating system of the Bureau of Labor Statistics.[9]

The Use of Hedonics by the Bureau of Labor Statistics

In 1990, the Bureau of Labor Statistics began employing hedonics in the apparel sector and, in 1999, incorporated hedonic equations for computer item substitutions, based on research done for the Producer Price Index.[10] Several years ago, the BLS developed and introduced into the index indirect hedonic methods for pricing noncomparable substitutions among ten additional products. Nine of these additional products were appliances or electronic products: televisions, VCRs, audio products, camcorders, microwave ovens, refrigerators, clothes dryers, washing machines, and DVD players. The tenth was a subject of particular interest to academics—namely, college textbooks.[11]

For eight of the ten products for which hedonic adjustments were recently introduced, the BLS generated two versions of the various strata indexes within which those products were located, one constructed with the traditional approach and the other with hedonic quality adjustment. In most cases, expenditures on the products to which hedonics were applied represented only a fraction of the total expenditures in the relevant CPI strata. In only three of the eight cases did the use of hedonic techniques make more than a small difference in the rate of price change in the relevant strata indexes during the period of comparison, and in two of those cases the hedonically adjusted indexes showed a smaller rate of price decline than did the published indexes. The switch to hedonic adjustments did produce more substantial differences in the rate of price change for the replacement items themselves. Calculations of the effect on the price changes of replacement items from the use of hedonic quality adjustments in place of the traditional approach were published for only five of the ten products.[12] The differences were positive on average for some of these products and negative for others, and on an unweighted basis, they roughly canceled out across all the products. All in all, the introduction of these hedonic adjustments had little impact on the Consumer Price Index.

[9] Pakes (2002) has suggested an alternative approach that would preserve the basic elements of the direct hedonic methodology while easing somewhat the burden on the Bureau of Labor Statistics (it would no longer be necessary to fit the hedonic equation in current (comparison) month).

[10] Since 1988, the Bureau of Labor Statistics has used a restrictive type of hedonics to adjust contract and owner equivalent rent for the effect of aging in housing units.

[11] Reports on eight hedonic studies can be obtained from the Bureau of Labor Statistics website: ⟨http://www.bls.gov/cpi/#publications⟩. The television study is reported in Moulton, LaFleur and Moses (1998).

[12] For a more detailed analysis of the recent hedonics studies done by the Bureau of Labor Statistics, see Schultze and Mackie (2002).

The television hedonic study, like many of the others, produced no significant difference from the traditional approach when it was applied only for item replacements. But in this study a direct hedonic approach was also employed, which generated an index of the quality-adjusted prices of *all* televisions in the sample, not just item replacements. Over the period 1993 to 1997, that index fell by 1.5 to 2.0 percent a year faster than the traditionally constructed index (with the range arising from the use of alternative base periods and weights). The authors of the television study suggest two possible reasons, both of which may have played some role. First, the direct estimation approach captures the effects of the large number of quality changes picked up during sample rotation, which are missed when only item replacements are hedonically adjusted. Second, the item replacement process is highly conservative, in that it calls for the selection of the replacement that is most similar to the item that disappeared (Moulton, LaFleur and Moses, 1997). One obsolete model is often replaced by another nearly as obsolete, which minimizes the selection of replacements nearer the cutting edge of technological advance where new models are more likely to enter at quality-adjusted prices lower than old models. Similarly, in markets where sellers use the occasion of introducing a new style to raise prices, the most similar replacement is less likely to be one of the new styles.

Concerns About the Current Hedonic Methodology

The panel concluded that even a substantially expanded use of hedonic techniques, if restricted to the current item replacement process, would be unlikely to have a significant effect on the Consumer Price Index. However, if as Moulton, LaFleur and Moses (1997) have suggested, out-of-date items were replaced by those that more nearly reflected changes in consumer buying patterns occurring since the last sample rotation, or if hedonic techniques were applied to quality changes occurring in sample rotation, then the fraction of price quotes receiving explicit and significant quality adjustments would expand substantially, which in turn would importantly increase the potential for the application of hedonic techniques to have an effect on strata indexes.

However, the panel's review of the application of hedonic models in the Consumer Price Index also raised a number of substantive questions about how the technique is currently being applied, including issues about the identification of characteristics, model stability and econometric specification. We concluded that these issues require a good bit of additional research and experimentation before hedonic techniques are further integrated into the CPI and the scope for their application substantially expanded. The reasons for our concern are spelled out in the body of the panel's report (chapter 4, especially pp. 132–145), but a few examples can give some flavor of their content.

A principal issue is the stability of the hedonic regression coefficients. Remember that in the indirect method, a hedonic equation is fit over a cross-section of the varieties of a product, and the resultant coefficients are used *unchanged* in subsequent periods to adjust item substitutions.

Hedonic equations for computers are now refit three or four times a year, because research has shown that the coefficients in such equations can change frequently. But such frequent refitting is exceptional. At least part of the reason is constraints on budget and personnel resources. In seven of the ten hedonic equations discussed above, the current BLS sample size had to be substantially expanded—on average by a factor of three—to obtain reasonably reliable estimates. Also, respecifying the hedonic models and reviewing the results is labor intensive, while the BLS has other research priorities to meet. Whatever the reasons, the Bureau of Labor Statistics, as of October 2002, had only refit equations for three of the ten other products (VCRs, DVD players and televisions), had not refit the remaining equations since they were developed, was considering again refitting the television equation, but otherwise had no current plans or schedule to refit the other equations in the near future.

Ariel Pakes (2002) has argued, convincingly I believe, that at least for some products, rapid technological advance and changes in markups and development strategies among imperfectly competitive firms should be expected to produce changes over time in the hedonic coefficients. Under the indirect approach, with infrequently refit equations, the issue of coefficient stability becomes particularly important. It seems reasonable that the variance over time in hedonic coefficients for a product will depend importantly on the pace of technological advance and on market structure. The individual characteristics of some products may have reasonably stable coefficients over substantial time periods, others not. The key question is which is which.

The use of brand names as characteristics in hedonic regressions raises some important issues. In almost all of the ten hedonic studies recently carried out by the Bureau of Labor Statistics, the regression equation included indicator variables for the brand name of the model. One rationale for the inclusion of brand name is that is serves as a proxy for unobserved qualities, such as quality of service or frequency of repair. But this assumption is not always warranted. In one case—microwave ovens—the study reported that brand coefficients were inversely correlated with *Consumer Reports* rankings for low repair frequency (Liegey, 2000, p. 5). When the correlation between a brand and other important included or excluded characteristics alters, application of an unchanged brand coefficient is likely to yield "wrong" quality adjustments. In this respect, the use of brand names coefficients in the indirect hedonic approach is simply a special example of the coefficient stability problem discussed above.

The Panel's Recommendations on Hedonics

In recent months, a number of the panel members have heard comments to the effect that the panel's report takes a negative view about the potential of hedonic techniques—apparently because the report discusses some of the difficulties with hedonic techniques. Yet our report explicitly concluded: "Hedonics currently offers the most promising technique for explicitly adjusting observed prices to account for changing product quality."

The issue is not whether hedonics is potentially of great usefulness. It is. Rather, what is at stake is essentially a choice between two different ways the Bureau of Labor Statistics could employ its hedonics R&D budget in the near-term future. The BLS could devote the bulk of those resources toward developing a steady stream of hedonic equations and incorporating them into the estimation of the Consumer Price Index. But the panel's analysis suggests that under current operating procedures, the results would not be likely to have much effect on the index. Alternatively, it could, as the panel suggests, channel its efforts principally into analyses, tests and experiments aimed at exploring and resolving some of the methodological issues discussed in the panel's report. The results might well justify the modification of BLS item replacement procedures and an expanded application of hedonics in a way that could make important improvements in the index.

Going beyond the content of the panel's report, my own view is that the research program, among many other goals, could investigate the question of whether evidence about the pace of technological advance or the market structure of the industry could be used to predict the degree of coefficient stability. With some experimentation, it might also be possible to design a regime under which newly developed hedonic equations would initially be refit at short intervals and the results used to help determine the appropriate frequency of future refitting. To the extent that, with sufficient research and experimentation, the Bureau of Labor Statistics can identify products that are likely to have relatively stable hedonic coefficients, the current methodology of indirect hedonics can be applied and expanded with infrequent refitting and reasonable cost. I suspect, however, that the application of hedonic methodology to sample rotation would require the use of direct hedonic methodology, which in turn involves continuous refitting of the equations. The panel recommended that the BLS experiment with the direct method, beginning with a few carefully selected goods.

As explained earlier, fitting hedonic equations typically requires the expansion of the current sample of prices collected by the Bureau of Labor Statistics or the purchase of privately collected data. Under current data collection methodology, frequent refitting and, even more so, the continuous refitting required by the direct method could become very costly. This in turn suggests that research on lowering the costs of data collection through the use of scanners, and perhaps other techniques, could eventually play an important role in enlarging the scope for hedonic methods within the Consumer Price Index.

Cost-of-Living Theory and Hedonic Techniques

One of the most widely cited advantages of cost-of-living theory is its usefulness in dealing with quality changes. It naturally prompts the question "what are the attributes of a good that consumers value" and looks for answers to the standard economic theory of consumption, which tells us that information about relative values can be inferred from their relative prices. But a closer look at the problem of measuring the effect of quality changes with hedonic techniques suggests that

the application of the theory to specific issues of quality adjustment is far from straightforward.

The hedonic coefficients on the characteristics of goods are used to impute a monetary value to the quality difference between two goods on the basis of the differences in their characteristics. In the standard economic theory of consumption, all consumers face the same prices for each good and adjust their purchases accordingly so that the ratios of prices equal their marginal rates of substitution. Price ratios are thus assumed to represent ratios of marginal values received. But in different varieties of a particular good, various attributes or characteristics are combined in a limited number of discrete packages, and hedonic functions are not generally linear. In equilibrium, consumers with different preferences will end up facing different prices for characteristics. Indeed, without this heterogeneity of preferences, individuals at the same living standards would all tend to buy the same variety of a good; all $25,000 automobiles would be the same. One consequence is that changes in income distribution and the demographic mix of consumers can shift the relative market prices of characteristics without any quality changes. The heterogeneity of consumer preferences over the various characteristics of a good, combined with the other aspects of quality comparisons described above, make it difficult to infer welfare interpretations from the properties of hedonic equations.

Zvi Griliches, one of the pioneers in applying hedonics to price index construction, commented in 1976—and cited the comment approvingly 14 years later (Griliches, 1990, p. 189, emphasis supplied):

> What the hedonic approach attempted was to provide a tool for estimating "missing" prices, prices of bundles not observed in the original or later periods. It did not pretend to dispose of the question of whether the various observed differentials are demand or supply oriented, how the observed variety of models in the market is generated, *and whether the resulting indexes have an unambiguous welfare interpretation.*

All the members of the panel agreed that hedonic regressions should be looked upon essentially as devices to estimate the market prices of alternative bundles of characteristics of goods and that hedonics can be applied within either a cost-of-goods or a cost-of-living index framework.

Again, the panel's recommendations on hedonics represent another example in which differences of views about the relative merits of the two index concepts did not prevent agreement about a set of specific recommendations to the Bureau of Labor Statistics.

Dealing with the Introduction of New Goods

The methodology of the Bureau of Labor Statistics does not reflect the gain in consumer welfare (the compensating variation or the consumer surplus) that arises

when new goods are introduced and gain a place in the market. This gain in consumer welfare is measured by the area under the Hicksian compensated demand curve above the current price, and for the consumers purchasing the product, this gain represents a decrease in the cost of living (Hausman, 1997).

To measure the welfare gain from the introduction of a new product, it is necessary to collect in each period data on quantities purchased and to estimate the demand curve for the new product and its "virtual price"—the price sufficiently high to reduce the quantity demanded to zero. A priori, one might expect that only new goods that provide radically improved capabilities would generate significant consumer surpluses. But in a well-known paper, Hausman (1997) estimated a demand curve for what would seem to be a modestly differentiated new variety of Cheerios breakfast cereal—Apple Cinnamon Cheerios—and calculated that its introduction generated substantial additions to consumer welfare. In his companion article in this issue, Hausman argues that the Bureau of Labor Statistics should not only calculate and adjust the Consumer Price Index for the introduction of completely new goods, but that current approaches for dealing with quality change, including the use of hedonic techniques, should be replaced by estimates of the corresponding compensating variations.[13]

The panel recognized that research into the welfare effects associated with new goods is important and should be pursued. But it emphasized the immense practical difficulties in the way of providing estimates of demand curves and virtual prices, especially if done across the large number and wide variety of products that would be required if this methodology were to supplant current methods of adjusting for quality change. In particular, estimating these welfare effects would impose the difficult requirement that the supply and demand factors that interact to generate prices and qualities be disentangled to identify the demand curve itself. Which assumption is chosen for identification purposes, among several competing possibilities, can often make a substantial difference in results. Thus, Hausman's (1997) estimate of the demand elasticity for Apple Cinnamon Cheerios has been disputed on grounds that a key assumption used in identifying the demand curve was open to serious question (Bresnahan, 1997).[14]

Knowledge about the desirability of most new products diffuses gradually throughout the economy, so that the demand curve is, for awhile, shifting rightward. Where fads or fashions play an important role, the demand curves for a new variety may first rise and then recede; consumer surpluses appear and then fade. To capture continuing changes in demand, the demand curve for new products and

[13] In both his cell phone and cereal studies, Hausman (1997, 1999) suggests that to avoid the uncertainties of extrapolating the "true" demand curve backward, outside the limits of observed data one could calculate a conservative lower bound estimate by extending back a tangent to the demand curve from the observed price and quantity.

[14] As pointed out earlier, hedonic equations are designed to estimate the market prices of bundles of characteristics. They do not depend on identifying the demand and supply factors underling price changes and hence generally pose much less rigorous econometric requirements.

varieties must be continuously reestimated, and the forces affecting supply and demand continuously disentangled.

There is an important potential in using scanner data, as well as other commercial electronic databases, to collect real time price and quantity data that could assist in studying new goods. At the same time, however there are substantial practical and conceptual challenges that would have to be overcome to incorporate widespread use of scanner data in the CPI.[15] The panel's report discusses both the possibilities and the challenges and identifies a number of areas that ought to receive high priority for research and experimentation.

The National Academy of Sciences panel concluded that it is unlikely a consensus methodology for producing reliable estimates of demand curves and virtual prices will emerge in the near future. It is impossible for the Bureau of Labor Statistics to attempt to incorporate into the Consumer Price Index measures of the welfare gain from the introduction of new goods or new varieties of existing goods with the economic and statistical techniques available at this time. Some panel members believed that even if reliable estimating methodology were available, the welfare gains from the introduction of new goods should not be treated as equivalent to a price reduction in the CPI. But recognizing that there are no measures of national output growth available that reflect the welfare gains from those events, the panel agreed that research in this area, while not designed to replace the CPI, should be directed toward developing, to the extent feasible, a separate experimental index that did account for such gains.

With All Deliberate Speed

Important progress has been made during recent years in improving the Consumer Price Index, especially from the standpoint of those who favor moving it closer to a cost-of-living index. The potential exists for still further significant progress. But in the process of trying to realize improvements, it is essential to avoid the temptation of moving rapidly to expand the use of potentially valuable techniques before their application across a wide range of areas has been sufficiently developed and tested.

This reasoning underlay the panel's recommendation that research be undertaken to deal with some important methodological problems in the current application of indirect hedonic techniques, not least the issue of coefficient stability, before further major integration of hedonics into the CPI. As one part of that effort, research and experimentation on data collection techniques might make feasible more frequent refitting of hedonic equations where that proves to be necessary, as well as a wider use of direct hedonic techniques.

It is also important that in moving the Consumer Price Index closer toward a

[15] A useful discussion of the promise and the difficulties of using scanner data for the CPI is contained in the NBER Conference volume, *Scanner Data and Price Indexes* (2000).

cost-of-living index, we remember that for the major purposes to which we put the CPI, it must remain grounded on the underlying concept of measuring the change in expenditures needed for a consumer to maintain a given standard of living *in the face of changes in the prices of market goods*, conditioned on stability in the status of conditions outside the market that affect consumers' living standards. The panel recognized that one of the important tasks for research in the area of economic measurement is the conceptual design and practical implementation of experimental measures of selected outside conditions. However, we agreed that the development and improvement of a cost-of-living index is not suitable vehicle into which to cram research about these matters.

In a similar vein, the current level of uncertainty about the accuracy and reliability of available techniques for estimating virtual prices strongly argues that the Bureau of Labor Statistics should not attempt to adjust the Consumer Price Index to take account of such effects. But that does not preclude undertaking research aimed at improving our ability to develop experimental measures of national output growth that take account of the welfare-enhancing effects stemming from the introduction of new goods.

■ *The author thanks Katharine Abraham, Angus Deaton, Erwin Diewert, Christopher Mackie, Van Doorn Ooms, Richard Schmalensee and Jack Triplett for their valuable comments and suggestions. The editors of this journal improved the initial draft and contributed not only in style and organization, but also on matters of substance.*

References

Boskin, M. J. et al. 1996. "Toward a More Accurate Measure of the Cost of Living." Final Report to the Senate Finance Committee.

Bresnahan, Timothy. 1997. "Comment," in *The Economics of New Goods.* T. Bresnahan and R. Gordon, eds. Chicago: University of Chicago Press, pp. 238–41.

Diewert, W. Erwin. 1976. "Exact and Superlative Index Numbers." *Journal of Econometrics.* May, 4, pp. 114–45.

Diewert, W. Erwin. 2000. "The Consumer Price Index and Index Number Purpose." Discussion Paper 00-02, University of British Columbia,

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

Hausman, Jerry A. 1997. "Valuation of New Goods under Perfect and Imperfect Competition," in *The Economics of New Goods.* T. Bresnahan and R. Gordon, eds. Chicago: University of Chicago Press, pp. 209–37.

Hausman, Jerry A. 1999. "Cellular Telephone, New Products and the CPI." *Journal of Business and Economic Statistics.* 17:2, pp. 188–94.

Liegey, Paul. 2000. "Developing a Hedonic Regression Model for DID Players in the CPI." U.S. Department of Labor, Bureau of Labor Statistics.

Liegey, Paul and Nicole Shepler. 1999. "Using Hedonic Methods to Quality Adjust VCR Prices."

Monthly Labor Review. September, 122:9, pp. 22–37.

Moulton, Brent R. and Karin E. Moses. 1997. "Addressing the Quality Change in the Consumer Price Index." *Brookings Papers on Economic Activity.* 1, pp. 305–49.

Moulton, Brent R., Timothy J. LaFleur and Karin E. Moses. 1998. "Research on Improving Quality Adjustment in the CPI; the Case of TVS." Washington, D.C., U.S. Department of Labor, Bureau of Labor Statistics.

National Bureau of Economic Research, Conference on Income and Wealth. 2000. *Scanner Data and Price Indexes.* Available at ⟨http://nber.org/books/criw00/quality12-28-00.pdf⟩.

National Research Council. 2002. *At What Price? Conceptualizing and Measuring Cost-of-Living and Price Indexes.* Charles Schultze and Christopher Mackie, eds. Washington, D.C.: National Academy Press.

Pakes, Ariel. 2002. "A Reconsideration of Hedonic Price Indices with an Application to PCS," in *Hedonic Price Indexes: Too Fast, Too Slow, or Just Right?* Jack E. Triplett, ed. Brookings Workshop on Economic Measurement, Feb 1. Available at ⟨http://www.brook.edu/dybdocroot/es/research/projects/productivity/workshops/20020201_pakes.pdf⟩.

Pollak, Robert A. 1971. "The Theory of the Cost-of-Living Index." Discussion Paper No. 11, Research Division, Office of Prices and Living Conditions, Bureau of Labor Statistics, Reprinted in Robert A. Pollak, *The Theory of the Cost-of-Living Index,* Oxford University Press, 1989, pp. 3–51.

Schultze, Charles and Christopher Mackie. 2002. "The NAS Panels Analysis of Quality Change and the Use of Hedonic Techniques in the CPI," in *Hedonic Price Indexes: Too Fast, Too Slow, or Just Right?* Jack E. Triplett, ed. Brookings Workshop on Economic Measurement, Feb 1. Available at ⟨http://www.brook.edu/dybdocroot/es/research/projects/productivity/workshops/20020201_schultze.pdf⟩.

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

Triplett, Jack E. 1997. "Measuring Consumption: The Post-1973 Slowdown and the Research Issues." *Federal Reserve Bank of St. Louis Review.* May/June, 79:3, pp. 9–42.

Triplett, Jack E. 2001. "Handbook on Quality Adjustment of Price Indexes for Information and Communication Technology." Organization for Economic Cooperation and Development, November.

U.S. Department of Labor, Bureau of Labor Statistics. 1997. Letter from the Commissioner of Labor Statistics to the Chairman of the Joint Economic Committee, U.S. Congress.

U.S. Department of Labor, Bureau of Labor Statistics. Website available at ⟨http://www.bls.gov/cpi/#publications⟩.