Charles R. Hulten

Price Hedonics: A Critical Review

1. Introduction

Price hedonics is a statistical technique developed more than seventy years ago to assess product quality issues. It had enjoyed a quiet and respectable life since coming of age in the early 1960s, but in the past few years, it has gained a degree of notoriety through a series of highly visible assessments of the consumer price index (CPI). This attention prompted a reassessment of price hedonics and its role in the CPI, which in turn has led to important new dimensions in the study of price hedonics. This paper focuses on these developments.

The new debate began in early 1995, when Federal Reserve Chairman Alan Greenspan testified before the Senate Finance Committee that he thought that the CPI was biased upward by perhaps 0.5 to 1.5 percentage points per year. This remark did not surprise specialists who understood the technical difficulties involved in constructing accurate price indexes, but it created a small sensation in the political arena. Here at last was a chance to get around one of the most difficult issues in the debate over balancing the federal budget: what to do about the social security program. Here was a way to reduce expenditures to balance the federal budget and rescue the social security trust fund from insolvency in the next century.

The beauty of it all was that the solution did not involve raising new taxes or changing benefit formulas. Instead, the solution involved "fixing" a biased method of adjusting social security benefits for the effects of price inflation, that is, by fixing the way the U.S. Department of Labor's Bureau of Labor Statistics (BLS) handles problems such as those posed when a new, improved product appears on the market.

These political considerations may seem tangential to the subject of price hedonics, but the events following from Greenspan's remark have linked the two issues. First, the Senate Finance Committee consulted a panel of experts, and that panel reached a consensus supporting Greenspan's estimate. Congress subsequently established the Advisory Commission to Study the Consumer Price Index (better known as the Boskin Commission, after its chairman) to estimate the level of the CPI bias. Boskin et al. (1996) arrived at an estimated bias of 1.1 percentage points per year—a level almost identical to Greenspan's estimate. Furthermore, the report said that about half (0.6) of that bias could be attributed to product innovations that were being overlooked in the CPI. A parallel study by Shapiro and Wilcox (1996) came to the same conclusion, estimating an overall bias of 1 percentage point per year, with 0.45 of that bias coming from quality changes and new goods. The study also observed that this bias was the most difficult to correct, likening the quality-adjustment process to house-to-house combat.

Price hedonics enters this picture because it offers the best hope for dealing with the bias that comes from product innovation. Although Boskin et al. (1996) did not explicitly recommend that the BLS expand the use of this technique in the CPI program (as a report by Stigler [1961] did), the BLS

Charles R. Hulten is a professor of economics at the University of Maryland and a research associate at the National Bureau of Economic Research.

The views expressed are those of the author and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

moved in this direction by increasing the number of items in the CPI treated with price hedonic techniques. In 1998, the BLS also requested that the Committee on National Statistics of the National Research Council (NRC) set up a panel of experts to investigate the conceptual issues involved in developing a cost-of-living index, including the use of price hedonic methods. This committee, chaired by Charles Schultze, released its report in early 2002 (National Research Council 2002). The NRC panel did not provide unanimous support for the underlying philosophy of the CPI as a pure cost-of-living index, and, in its own words, differs from the Stigler and Boskin et al. reports in this regard (National Research Council 2002, p. 3).

The NRC panel was cool to the BLS's expanded commitment to price hedonics. On the one hand, the NRC report endorsed hedonic techniques as a research tool, commenting that they "currently offered the most promising approach for explicitly adjusting observed prices to account for changing product quality." The report's Recommendation 4-2 noted that the "BLS should continue to expand its experimental development and testing of hedonic methods." On the other hand, Recommendation 4-3 of the report cautioned against immediately expanding the use of hedonics in constructing the CPI itself: "Relative to our view on BLS research, we recommend a more cautious integration of hedonically adjusted price change estimates into the CPI." The report explained the apparent disconnect between the two recommendations by pointing to a "concern for the perceived credibility of current methods," adding that "while there is an established academic literature on estimating hedonic functions, researchers are much less experienced using them across a wide variety of goods" (National Research Council 2002, pp. 6-7).

The "perceived credibility" standard is something new in the critique of price hedonic methods and, more generally, in the discussion of price measurement. It asserts a higher standard of acceptability for results that have a significant effect on policy (and, by extension, on the well-being of the public) than it does for "academic" research. This idea has been implicit in policy analysis (and in statistical agency policy) for a long time, and the explicit appeal to the perceived credibility standard may well be the most enduring intellectual contribution of the NRC panel. However, the panel did not spell out what additional requirements were implied by this standard. Its members called for further research, and in Recommendation 4-8 urged the creation of an advisory panel of experts to help guide this research. The goal of this new advisory panel was to "provide an analytic basis for proceeding sensibly in the face of external pressures to proceed quickly in this area" (National Research Council 2002, p. 7).

The absence of explicit criteria is not surprising because the political economy of statistical measurement is largely terra incognita in the practice of economics. However, the NRC panel report forces the debate in this new direction. Accordingly, the main objective of this paper is to make a start in the evaluation of price hedonics from this expanded perspective. In the next section, I describe the hedonic model and review its main uses, because the credibility of price hedonics depends in part on the current state of academic research. This is necessarily a brief overview, and the interested reader is directed to excellent treatments of the subject in Berndt (1991), Triplett (1987), and the extensive expository material in National Research Council (2002). I then turn to some of the standard criticisms of price hedonics and move into the uncharted waters of the political economy of price measurement.

2. The Structure and Interpretation of the Price Hedonic Model

2.1 The Hedonic Hypothesis

Product variety is the raison d'être of the price hedonic model. Certain types of commodities are differentiated into subtypes: different models of autos, different species of petunias, different configurations for personal computers, different brands of toothpaste, and so on. Each subtype could be treated as a good in its own right, with its own price and quantity. This differentiation is appropriate for some purposes (for example, industrial organization studies), but it is inefficient in macro studies of inflation and growth if the number of underlying characteristics or attributes defining the item is small relative to the number of varieties in the marketplace. In this case, a more tractable way of proceeding is to view each subtype in terms of its characteristics, $\chi_{j,t}$, and to define the good by the "quantity" of each of its component characteristics, $X_t(\chi_{1,t}, \ldots, \chi_{n,t})$. This formulation leads naturally to a definition of product quality in terms of the amount of each characteristic that each variety has.

The empirical link between a variety and its constituent attributes is established in the hedonic model through its price, not its quantity. The price of a variety j at time t, $P_{j,t}$, is assumed to be a function of its defining characteristics, $h_t(\chi_{j,t})$, plus a random error term. In econometric applications, the hedonic function is assumed to have linear, log-linear, or semi-

log forms.[1] I use the linear specification as an example of the hedonic function to simplify the exposition, although it is not the best functional form for empirical purposes:[2]

$$(1) \qquad P_{j,t} = \beta_0 + \beta_1 \chi_{1,t} + \ldots + \beta_n \chi_{n,t} + \varepsilon_t .$$

The hedonic weights, β_i, are the portion of an item's overall price attributable to a given characteristic and are usually interpreted as the price of the corresponding characteristic.

There are two basic approaches in the literature to understanding the characteristic price. One tradition relates this price to a consumer's willingness to pay for the characteristic. This utility-based interpretation is reflected in the use of the term "hedonic" to describe the approach, and was the original view of the matter adopted by Court (1939) and other early practitioners. Lancaster (1966) proposed a theory of consumer utility based on characteristics rather than on goods, and Diewert (2001) described the rather restrictive conditions under which the hedonic function can be derived from an underlying utility function.

The second approach, developed by Rosen (1974), has become the generally accepted paradigm of the hedonic approach. Rosen relates the hedonic function to the supply and demand for individual characteristics, that is, the function relates to the demand curves of consumers with heterogeneous tastes for the different combinations of characteristics in each variety, and to the corresponding supply functions for each characteristic. According to this view, the price hedonic equation is basically an envelope linking the various equilibriums, although—as Rosen emphasizes—the link also requires restrictive assumptions. This view was advanced by many authors, including Triplett (1983), Epple (1987), Feenstra (1995), and Pakes (2002).

2.2 Price Inflation and Quality Change

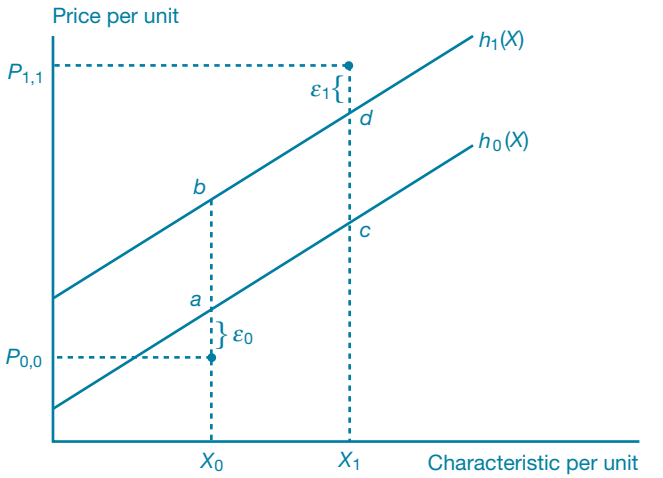
The concepts of price inflation and quality change have a straightforward interpretation in the hedonic model. Inflation leads to an upward shift in the hedonic function because some or all characteristics become more expensive (for example, β "prices" increase). The case of quality change, however, is somewhat more complex. Quality change can arise from two sources: composition change, which brings new varieties into the CPI sample that were technically feasible but were not produced for economic reasons or were produced but not included in the CPI sample; and product innovation, which introduces new varieties to the marketplace that were not feasible in prior years.

Changes in the composition of the varieties seen in the marketplace can occur because changes in income, individual tastes, or demographics dictate a change in the product mix within the feasible set of possible varieties. For example, rising incomes in a particular area may lead some supermarkets to introduce upscale brands of food. A change of this sort is equivalent to a movement along the hedonic function from χ_0 to another χ_1.

Product innovation, however, occurs when technological innovation in product design or production leads to a reduction in the cost of acquiring a given amount of a characteristic (or more characteristics for the same price). Improvements in personal computers fall into this category. This sort of quality change is equivalent to a downward shift in the hedonic function. A variant of this theme occurs when quality innovation leads to the introduction of varieties that have a greater number of one or more characteristics than were previously feasible, without lowering the cost of existing varieties. Aircraft with larger capacity are an example of this possibility. This case can be represented in the exhibit below as an extension of the feasible portion of the existing hedonic function.

The exhibit shows the case of a linear hedonic function with a single characteristic. The hedonic surface for the reference time period $t = 0$ is designated $h_0(\chi)$; the variety sampled in this period has χ_0 units of the characteristic and costs $P_{0,0}$. This price deviates from the hedonic line by the error ε_0. The hedonic surface for the comparison period shifts upward to $h_1(\chi)$, and a new variety is sampled with χ_1 units of the characteristic. It costs $P_{1,1}$, with a deviation from the hedonic

Linear Hedonic Function with a Single Characteristic



line of ε_1. The upward shift in the hedonic function indicates that inflationary pressures dominate any cost-reducing product innovation, but from the data in the exhibit it is not possible to separate the two effects (or even tell if product innovation has occurred).

2.3 Uses of the Hedonic Method

Price hedonics has been applied to a wide range of issues in various economic fields. At the risk of oversimplification, it is useful to put these studies into two broad groups: those that are mainly concerned with adjusting observed prices on the left-hand side of the hedonic regression for changes in product quality, and those that focus on issues relating to the individual characteristics and β-coefficients on the right-hand side of the hedonic regression. Much of the recent debate has focused on the first of these objectives.[3]

Indeed, the main mission of price hedonics has always been to isolate the quality component of price changes to achieve better measures of price inflation. This was the objective of the original Waugh (1928) and Court (1939) studies, and was recognized by Stigler (1961). Price hedonics has influenced official price statistics in two ways: through the decision by the Bureau of Economic Analysis (BEA) of the U.S. Department of Commerce to adjust computer prices for quality change using price hedonic techniques from the work by Cole et al. (1986), and through the quality-adjustment techniques used by the BLS to adjust the CPI and the producer price index (PPI).

The "matched-model" method is the primary procedure used to construct the CPI. A representative sample of consumer goods and retail outlets is drawn and, once a given type of good is selected, the BLS price-taker attempts to find a match for the reference good and price it each month. The individual price matches are aggregated into the CPI using a two-stage procedure. In 1995, an "exact" match was made almost 98 percent of the time each month (see National Research Council [2002, p. 117], based on Moulton and Moses [1997]). In 2.16 percent of the cases where a sample item was replaced, two-thirds of the replacement items were deemed to be comparable substitutes for which no adjustment for quality was necessary. For the remaining one-third, a quality adjustment to price was made by using various techniques, including price hedonics. Hedonics thus played only a small role in the big picture in 1995, affecting about 0.2 percent of the items priced each month (although it had a slightly larger effect on the price index).

These figures do not seem to imply a large enough role to justify all of the attention that hedonics has recently received. However, the BLS expanded the role of price hedonics after the Boskin Commission report and is considering further expansion. This expansion reflects, in part, the technical virtues of the hedonic method, but it is also motivated by dissatisfaction with the other quality-adjustment techniques used in the CPI.

These issues can be illustrated in the context of our exhibit. The matched-model method starts with the selection of a variety (say, χ_0) to price each time period. The expected price change between the reference and comparison periods is simply the ratio $h_1(\chi_0)/h_0(\chi_0)$. If the variety χ_0 remains in the marketplace in a purely static world, the matched-model strategy will continue to price this variety. A problem arises if the variety disappears from the sample. When this happens, a replacement must be found, and if a new variety χ_1 is selected, whose observed price is $P_{1,1}$, then the BLS must consider the possibility that some part of the observed price increase $P_{1,1}/P_{0,0}$ may be because of a change in quality.[4] At this point, the BLS must decide if the new variety is a comparable or noncomparable substitute. If it is comparable, χ_0 and χ_1 are deemed to be equivalent and the observed price ratio $P_{1,1}/P_{0,0}$ is not adjusted for quality. If this is wrong and the new variety is really a noncomparable substitute, the ratio $P_{1,1}/P_{0,0}$ overstates the true rate of pure price increase when $\chi_1 > \chi_0$.[5] More generally, the price ratio is the product of a pure price term and a quality term. This ratio can be written from the standpoint of the comparison period $t = 1$ as

$$\frac{P_{1,1}}{P_{0,0}} = \frac{P_{0,1}}{P_{0,0}} \times \frac{P_{1,1}}{P_{0,1}}, \tag{2}$$

where $P_{0,1}$ is the unobserved price of original variety χ_0 in the comparison period (the price that would have been paid in $t = 1$ for χ_0 had it been available for sampling). In the exhibit, the *expected* price term is the vertical distance between the price $P_{0,0}$ and the point *b*, and the expected quality term is the vertical distance between *b* and *d*.

A parallel quality adjustment can be made from the standpoint of the reference period $t = 0$:

$$\frac{P_{1,1}}{P_{0,0}} = \frac{P_{1,1}}{P_{1,0}} \times \frac{P_{1,0}}{P_{0,0}}, \tag{3}$$

where $P_{1,0}$ is the unobserved price of variety χ_1 in the comparison period (the price that would have been paid in $t = 0$ for χ_1 had it been available for sampling then). In the exhibit, the *expected* price term is the vertical distance between the price $P_{1,1}$ and the point *c*, and the *expected* quality term is the vertical distance between *a* and *c*.

The price-quality decomposition in equations 2 and 3 requires estimates of the missing prices $P_{0,1}$ and $P_{1,0}$. The BLS has several methods for estimating them: the overlap method, where these prices are, in fact, observable somewhere (useful

when the sample is intentionally changed and new items are "rotated" into the sample); the link and class-mean methods, where the missing prices are imputed by averaging the prices of similar products (historically the dominant method); and the "direct" adjustment methods, which impute the missing prices $P_{0,1}$ or $P_{1,0}$ by their cost of production, or by using price hedonics. The hedonic solution is simply $P_{0,1} = h_1(\chi_0)$ or $P_{1,0} = h_0(\chi_1)$. This is the most intellectually satisfying of the various quality-adjustment methods because it appeals to an underlying economic structure rather than to opportunistic proxies. A case for using hedonics can be made on these grounds alone: hedonic regression analysis inevitably involves statistical error, *but so do the other methods*. The current consensus appears to be that the dominant link and class-mean approaches are subject to a greater degree of error, but more research is needed on the accuracy of all methods. Some of the common problems associated with hedonic regressions are reviewed in the next section, but this critique must be viewed with the larger picture in mind.[6]

3. A Critique of the Hedonic Regression Model

3.1 Fact versus Inference in Price Measurement

The portrait of price hedonics painted in the preceding section is rather flattering, particularly when compared with competing alternatives. What, then, accounts for the conservative Recommendation 4-3 from the NRC panel and an ambient skepticism on the part of some users? One of the leading developers and practitioners of price hedonics, Triplett, found it necessary to devote an entire article to the analysis and refutation of common criticisms of the hedonic method (Triplett 1990). I believe that a large part of the problem reflects a lower degree of confidence in data that are imputed using regression analysis. Price estimates collected directly from an underlying population are generally regarded as "facts." When the price is inferred using regression techniques, it becomes a "processed" fact subject to researcher discretion.

It is certainly true that sampling techniques involve a degree of discretion in sample design. In the CPI, decisions are made about which items are included in the matched-model samples, which outlets are visited, the size of the sample, when a substitution is comparable and noncomparable, and so on. The resulting price estimates involve a sampling variance and a potential for bias and are no different in this regard than estimates obtained using regression analysis. There is, however, an important difference from the standpoint of perceived credibility. The CPI sample is constructed directly from the population of consumption goods in retail outlets whose prices are "facts on the ground." Full enumeration of the population is conceptually possible, lending verisimilitude to the sampling process.

The perceived credibility of the researcher discretion involved in regression analysis is not so well anchored. The old saw about statistical regressions applies here: "If you torture Mother Nature long enough, she will ultimately confess to anything you want." This quip reflects a widely understood but seldom emphasized truth about applied econometrics: researchers rarely complete their analysis with the very first regression they try. The first pass-through of the data often produces unsatisfactory results, such as poor statistical fits and implausible coefficient estimates. Rather than stop the analysis at this point, researchers typically use the same data to try out different functional forms and estimation techniques, and drop weak explanatory variables until plausible or satisfactory results are obtained (or the project is abandoned). The NRC panel report cites instances of these practices during the incorporation of price hedonics in the CPI program (National Research Council 2002, p. 142).

This "learning-by-doing" approach has a pragmatic justification: theory is rarely a precise guide to practice, and experimentation with alternative techniques and specifications is both normal and necessary. It is ideal to draw a fresh sample for each new attempt, but resampling is usually expensive and sometimes unfeasible. However, the resulting estimates may lack the statistical power to discriminate among competing models.

3.2 Rounding Up the Usual Suspects

The economics profession has been moving along the price hedonics learning curve for some time, and it may be useful at this point to review briefly the current state of progress (for a more detailed account, see National Research Council [2002, chapter 4]). To that end, I now examine three general issues.

The first general issue is that price hedonics is subject to the problem of all product differentiation models: where does a good stop being a variety of a given product class and become a product on its own? It is intuitively reasonable to group all

Toyota Corollas in the same class and treat different equipment options as characteristics. Is it as reasonable to group included near-substitutes such as Toyota Camrys or all Toyota passenger cars into the same product class? Perhaps the product classes should be functional—subcompacts, compacts, luxury sedans, suburban utility vehicles—regardless of brand.

Theory gives only the following guidance: items should be grouped according to a common hedonic function. For example, if equation 1 is the correct specification, all items in the hedonic class must have the same list of characteristics and the same β-coefficients. This implied grouping seems reasonable for different configurations of a Toyota Corolla, but increasingly less so as the range of included items is expanded. It should be possible to test for homogeneity of items included in a hedonic class, but it is not clear how often this is actually done. Dummy variables for different brands within a given class can be used in some cases, but this is essentially an admission that some important characteristics are missing or that the β-coefficients differ in at least one dimension.

This problem is attenuated in the CPI because the items included in the matched-model design are rather narrowly specified. However, although the narrowness of matched-model item specifications helps with the problem of heterogeneous β-coefficients, it exacerbates the problem of "representativeness." Learning a lot about inflation and quality change in one narrowly defined class like Toyota Corollas may not be indicative of the experience of the broader class of automobiles.

A second general class of issues involves the selection of characteristics. Hedonic theory suggests that a characteristic should be included in the analysis if the characteristic influences consumer and producer behavior. This implicitly assumes that consumers and producers have the same list, which is far from obvious (Pakes 2002). The consumer may be interested in performance characteristics such as top speed and acceleration, while the seller may focus on product attributes like engine horsepower, and the design engineer on technical characteristics like valve design. Furthermore, different consumers may base their spending decisions on different sets of characteristics or assign different weights to them, meaning that the β-coefficients in equation 1 are really not fixed parameters, but weighted averages. As a result, estimated parameters may not be stable over time, and the implied estimates of price and quality may shift simply because of changes in the mix of consumers.

Another concern is the problem of separability and "inside" and "outside" characteristics. The β-coefficients in equation 1 may be unstable over time for another reason: the characteristics defining one good are not separable from the characteristics defining other goods. This is a well-known result in aggregation theory and is hardly unique to price hedonics. But the hedonic hypothesis is a form of aggregation and the stringent conditions for separability may fail. In this case, a change in some characteristic outside the set of "inside-the-hedonic-function" characteristics may cause the relation between the inside elements to shift, leading to a change in the β-coefficients.[7] A similar problem can arise when some of the relevant characteristics are left out of the regression analysis.

The problem of missing inside characteristics and nonseparability with respect to outside characteristics can be subjected to econometric tests. However, the truth is that the selection of characteristics is heavily influenced by data availability, and it is not clear how much progress can realistically be expected to occur when dealing with these conceptual issues.

Choice of appropriate functional form is the third general class of problems often raised in critiques of price hedonics. The three most common forms—linear, semi-log, and log-log—do not allow for a very rich set of possible interactions among characteristics. Important complementarities often exist, for example, between microprocessor speed and storage capacity. One does not substitute for the other at a given price in most applications. Expanding an automobile's performance to racecar levels involves an increase in many characteristics, not just a very large increase in horsepower alone. This suggests the use of more flexible functional forms such as the trans-log. Furthermore, as noted in the preceding section, innovations in product quality can take the form of extensions of the length of the hedonic function over time, and this is hard to capture with the usual functional forms.

3.3 The Pakes Developments and the New Heterodoxy

Many of the problems noted above are generic to many econometric applications and many can be addressed with alternative econometric techniques. However, the recent study by Pakes (2002) suggests that some of these problems are really not problems at all. Pakes' study is a potential paradigm shifter and deserves special attention.

Pakes advances three important propositions, which I call Pakes-I, Pakes-II, and Pakes-III. Pakes-I starts with the usual interpretation of the hedonic function as a locus of supply and demand equilibriums for heterogeneous agents in which the price of each characteristic is equal to its marginal cost—the standard view inherited from Rosen (1974). Pakes observes that this assumes that producers have no market power over

the package of characteristics they offer, and that this is a poor assumption to impose on a world of product differentiation. The product/characteristics space is not continuously dense for most differentiated products, and producers try to differentiate their products to achieve a degree of market power. Moreover, product innovation is part of the product differentiation process, and innovation tends to convey a degree of market power.

Pakes derives an alternative interpretation of the hedonic function in which price equals marginal cost plus a market power term that depends on the elasticity of demand for the characteristic. This is the Pakes-I result, and it is surely correct for many of the goods for which price hedonics is employed. However, the implications of this result are novel to the point of heterodoxy:

> Hedonic regressions have been used in research for some time and they are often found to have coefficients which are "unstable" either over time or across markets, and which clash with naive intuition that characteristics which are generally thought to be desirable should have positive coefficients. This intuition was formalized in a series of early models whose equilibrium implied that the "marginal willingness to pay for a characteristic equaled its marginal cost of production." I hope [the preceding] discussion has made it amply clear that these models can be *very misleading* [author's emphasis]. The derivatives of a hedonic price function should not be interpreted as willingness to pay derivatives or cost derivatives; rather they are formed from a complex equilibrium process (Pakes 2002, p. 14).

This view clashes strongly with the conventional view, which is summarized in the National Research Council (2002) report in the following way:

> Strange-looking variable coefficients could be indicative of larger problems—including omission of key value indicators, characteristic mismeasurement, and functional form issues (p. 142).

Furthermore,

> It is hard to know when a hedonic function is good enough for CPI work: the absence of coefficients with the "wrong" sign may be necessary, but it is surely not sufficient (p. 143).

In the Pakes view of price hedonics, there is no reason to assume that the hedonic function and the β-coefficient should be stable over time, and the "wrong" sign is not necessarily wrong at all. In fact, the price associated with any characteristic may be negative. In other words, the price of a product can go down when it acquires more of a given characteristic. This result is a corollary to Pakes-I, but is so important that it deserves a separate status as Pakes-II. Pakes-II turns conventional wisdom on its head and challenges any notion of perceived credibility based on intuition about parameter instability and "wrong" signs.

Pakes-III is yet another corollary. This result argues that parameter instability and counterintuitive signs are irrelevant if the point of the hedonic analysis is merely to correct observed prices for changes in quality (and not to interpret individual coefficients—recall the two general objectives of price hedonics noted earlier). In terms of our earlier exhibit, Pakes-II implies that two hedonic lines need not bear any close resemblance to each other. Pakes-III implies that estimation of either line is sufficient to make a quality adjustment. All that is needed to impute the terms in the price ratios in equations 2 and 3 are estimates of $h_0(\chi)$ and $h_1(\chi)$.

These results represent a potential paradigm shift in the field of price hedonics. They have yet to be vetted by the specialists in the field, but some or all of each proposition is likely to survive scholarly scrutiny.[8] There are a number of issues to be resolved, such as the problem of cross-sectional stability. The same mechanism that causes the hedonic coefficients to be unstable over time may also cause them to be unstable in a cross-section of consumer prices drawn from different locations and different types of retail outlets. In this case, the movement along the hedonic function at any point in time may not be possible. This, and other issues, await further debate.

4. The Political Economy of Price Hedonics

There is a saying in tax policy that "an old tax is a good tax." This does not follow from any deep analytical insight into optimal tax theory, but from the pragmatic observation that taxation requires the consent of the governed. The public must accept and respect the tax, and this does not happen automatically when a tax is introduced. There is typically a learning curve as people adjust their behavior in light of new tax incentives, and gainers and losers are sorted out. The tax matures as affected groups negotiate changes and as unforeseen consequences become apparent and are dealt with.

A similar argument leads to the proposition that "old data are good data." Old data, like old taxes, involve learning by the public and by policymakers about a new set of facts, and both may involve large economic stakes. In the case of CPI reform, the Boskin Commission estimated that the cumulative effects

of a 1 percentage point per year bias would have added $1 trillion to the national debt between 1997 and 2008. If price hedonics were completely successful in eliminating the Boskin Commission's quality bias, the growth rate of the CPI would fall by about 25 basis points to 60 basis points per year, with an attendant reduction in cost-of-living payments to individuals.[9] In addition, cost-of-living adjustments to social security, federal civilian and military retirement, supplemental security income, and other programs are not the only dimension of policy affected by this line of argument, because the CPI is used to index income tax parameters, Treasury inflation-indexed bonds, and some federal contracts.

Moreover, a revision to the CPI also changes the metric that policymakers use to gauge the rate of inflation. They have to assess how much of the change in measured inflation is the result of underlying inflationary pressures and how much is the result of the new methods. This reflects a fundamental truth about the policy process: policy decisions (indeed, most decisions) must be made with imperfect information. There is learning over time about the nature of the data and the useful information they contain. Chairman Greenspan's 1995 comment about his perception of a bias of 0.5 to 1.5 percentage points in the CPI is a case in point.

The expanded use of price hedonics thus looks different to users who are interested in the "output" of the technique than to expert practitioners who are interested in developing the technique per se. Put differently, there is a policy-user learning curve that is different from the researcher learning curve. However, the two curves are related. The weaker the professional consensus is about a technique, the lower the level of confidence is in the technique's consequences and in its acceptance by the public and policymakers. This is the essence of the "perceived credibility" standard.[10]

This line of argument has implications for the use of price hedonics in the CPI. Perceived credibility is linked to the degree of professional consensus, and Pakes (2002) has pretty much upset whatever consensus had existed. It will doubtless take time to sort out the propositions advanced by Pakes, and this alone justifies the conservatism of the NRC's Recommendation 4-3. More research is needed on the robustness of price hedonic results to changes in assumptions about functional forms and characteristics and about the circumstances under which parameter instability and "wrong" signs occur. Monte Carlo studies, in which the true value of the parameters is known in advance, could be a useful way of understanding the pathology of the hedonic technique and assessing the accuracy of this technique and its ability to forecast the CPI, both in absolute terms and relative to other quality-adjustment methods.

5. Conclusion

Research at the frontier should be innovative and challenging, aimed at convincing peer researchers. However, this is not the way good policy is made. Policy ultimately relies on the consent of the public, not the vision of convinced experts. Changes in official statistical policy therefore should be conservative and credible, and the research agenda must include a component aimed at building confidence that the benefits of change outweigh the costs. Accordingly, the National Research Council panel is right to insist on a conservative approach to the increased use of price hedonics in the CPI. However, the research community is also right to insist that this technique is the most promising way to account for changes in product quality in official price statistics.

Researchers would also be right to point out that part of the credibility issue with hedonics is about the *switch* to the new technique, and not just about the technique itself. Had the BLS used price hedonics more extensively in the past rather than the more commonly used quality-adjustment methods, hedonics would probably have evolved by now to the point of perceived credibility. Indeed, if positions were reversed and the link, overlap, and class-mean methods were offered as substitutes for an entrenched hedonics methodology, the debate would be very different.

Endnotes

1. Berndt (1991) cites the Waugh (1928) study of fresh asparagus in Boston markets as being the earliest known empirical example of the technique. The first hedonic regression analysis is attributed to Court (1939), who studied passenger cars. However, the growth in the field began with the work of Griliches (1961).

2. See, for example, Diewert (2001), who advocates the use of more flexible functional forms.

3. Although this paper is essentially about "left-hand-side" issues, it is worth noting that a number of interesting economic problems are naturally formulated in terms of individual characteristics and their β-coefficient. For example, when the log price of producers' used durable equipment is regressed on two characteristics—the year in which the equipment was sold and its age at the time of sale—the β-coefficient of age can be interpreted as the rate of economic depreciation. Indeed, this is the theoretical definition of depreciation. This approach formed the basis for my own work with Frank Wykoff, which estimated rates of depreciation for a wide variety of business-fixed capital and which has come to be embedded in the national income and product accounts estimates of the capital consumption adjustment. Another example comes from human capital theory. The determinant of wage rates has been studied using price hedonics by putting wages on the left-hand side of equation 1 and worker characteristics on the right-hand side. Other examples include the use of hedonics to study such diverse items as housing values and fine wines.

4. This is one way that quality change affects the CPI sample. Another occurs when the sample is "rotated" to include new items.

5. This is one source of the Boskin Commission's quality bias.

6. This section has focused on the use of price hedonics in the CPI program. However, the most quantitatively important use of hedonics up to now has probably occurred on the "real" side of official statistics through the BEA's computer price adjustment, which is based on Cole et al. (1986). The BEA adjustment redefines the units in which output is measured from computer "boxes" to effective units of computer power, reflecting the fact that new varieties of computers pack more capacity into each box. This, in turn, increased the measured growth rate of real GDP and enhanced the perception of the emerging "new economy."

7. In more concrete terms, the value of extra power in a personal computer may shift as new software or applications become available. Another example is the trade-off between extra performance and additional comfort in automobiles that depends on such factors as the quantity and quality of the highway systems.

8. An active program of research on this subject is currently under way (for example, see Berndt and Rappaport [2001] and Silver and Heravi [2002]). Moreover, the Pakes-II result has precedent in conventional price-quantity analysis. When the price of a good is regressed on its quantity, it is well-known that the underlying supply and demand curves generally cannot be identified separately, and that the regression coefficients will be unstable and can easily have the "wrong" sign. The price hedonic case is somewhat more complex because the hedonic function contains multiple varieties, but it is also a case in which price is regressed on the "quantity" of characteristics.

9. The NRC panel report concludes that the expanded use of price hedonics is unlikely to have a large effect on CPI growth if it is limited to imputing missing prices for noncomparable substitution items. Several recent BLS commodity studies have found that price hedonics did not produce dramatically different results from those of other quality-adjustment methods. However, the impact could be much larger if hedonics was applied more broadly.

10. The "perceived credibility" standard and the notion of "old" data are not well established in the literature on economic measurement. Most discussions focus on "better" or "more accurate" as the appropriate criteria for comparing new measurement techniques with old: if a new method promises more accurate data, it should be adopted without hand-wringing about gainers and losers. The job of the experts, in this view, is to provide the best scientific advice they can and leave politics to the politicians and public. However, this "ivory tower" view of expert knowledge ignores the fact that it is the politicians and the public who asked (and largely paid) for the advice in the first place. Users have a right to demand a quality product from the supplier and to define quality in their own terms. The perceived credibility standard is part of this quality control.

References

Berndt, E. 1991. The Practice of Econometrics. Reading, Mass.: Addison Wesley.

Berndt, E., and N. Rappaport. 2001. "Price and Quality of Desktop and Mobile Personal Computers: A Quarter Century Historical Overview." American Economic Review 91, no. 2 (May): 268-73.

Boskin, M., E. Dullberger, R. Gordon, Z. Griliches, and D. Jorgenson. 1996. "Toward a More Accurate Measure of the Cost of Living." Final report to the Senate Finance Committee from the Advisory Commission to Study the Consumer Price Index.

Cole, R., Y. C. Chen, J. A. Barquin-Stolleman, E. Dullberger, N. Helvacian, and J. H. Hodge. 1986. "Quality-Adjusted Price Indexes for Computer Processors and Selected Peripheral Equipment." Survey of Current Business 66, no. 1 (January): 41-50.

Court, A. T. 1939. "Hedonic Price Indexes with Automotive Examples." In Dynamics of Automobile Demand, 99-117. General Motors Corporation.

Diewert, W. E. 2001. "Hedonic Regressions: A Consumer Theory Approach." Unpublished paper, University of British Columbia Economics Department.

Epple, D. 1987. "Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products." Journal of Political Economy 95, no. 1 (February): 59-80.

Feenstra, R. 1995. "Exact Hedonic Price Indexes." Review of Economics and Statistics 77, no. 4 (November): 634-53.

Friedman, M. 1953. "The Methodology of Positive Economics." In Essays in Positive Economics, 3-43. Chicago: University of Chicago Press.

Greenspan, A. 1995. "Consumer Price Index: Hearings before the Committee on Finance, U.S. Senate." Statement to U.S. Senate Hearing 104-69, 109-15. Washington, D.C.

Griliches, Z. 1961. "Hedonic Price Indexes for Automobiles: An Econometric Analysis of Duality Change." In The Price Statistics of the Federal Government, General Series no. 73, 137-96. New York: Columbia University and National Bureau of Economic Research.

Hausman, J. 1997. "Valuation of New Goods under Perfect and Imperfect Competition." In T. Bresnahan and R. J. Gordon, eds., The Economics of New Goods. Studies in Income and Wealth 58: 209-37. Chicago: University of Chicago Press and National Bureau of Economic Research.

Hulten, C. 2000. "Measuring Innovation in the New Economy." Unpublished paper, University of Maryland.

Hulten, C., and F. Wykoff. 1979. "The Estimation of Economic Depreciation Using Vintage Asset Prices." Journal of Econometrics 15, no. 3 (April): 367-96.

Lancaster, K. 1966. "A New Approach to Consumer Theory." Journal of Political Economy 74, April: 132-57.

Lucas, R. 1976. "Econometric Policy Evaluation: A Critique." Carnegie-Rochester Conference Series on Public Policy: 19-46. Amsterdam: North-Holland.

Moulton, B. 1996. "Bias in the Consumer Price Index: What Is the Evidence?" Journal of Economic Perspectives 10, no. 4 (fall): 159-77.

Moulton, B., and K. Moses. 1997. "Addressing the Quality Change Issue in the CPI." Brookings Papers on Economic Activity, no. 1: 305-66.

National Research Council. 2002. At What Price?: Conceptualizing and Measuring Cost-of-Living and Price Indexes. C. Schultze and C. Mackie, eds. Committee on National Statistics, Panel on Conceptual, Measurement, and Other Statistical Issues in Developing Cost-of-Living Indexes. Washington, D.C.: National Academy Press.

Pakes, A. 2002. "A Reconsideration of Hedonic Price Indices with an Application to PCs." NBER Working Paper no. 8715.

Rosen, S. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." Journal of Political Economy 82, no. 1 (January/February): 34-55.

Shapiro, M., and D. Wilcox. 1996. "Mismeasurement in the Consumer Price Index: An Evaluation." In B. S. Bernanke and J. J. Rotemberg, eds., NBER Macroeconomics Annual 1996, 93-142. Cambridge: MIT Press.

References (Continued)

Silver, M., and S. Heravi. 2002. "On the Stability of Hedonic Coefficients and Their Implications for Quality-Adjusted Price Change Measurement." Paper presented at the National Bureau of Economic Research 2002 Summer Institute, Cambridge, Massachusetts, July 29.

Stigler, G. 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, J. 1983. "Concepts of Quality in Input and Output Price Measures: A Resolution of the User Value-Resource Cost Debate." In M. F. Foss, ed., The U.S. National Income and Product Accounts: Selected Topics. Studies in Income and Wealth 47: 296-311. Chicago: University of Chicago Press and National Bureau of Economic Research.

———. 1987. "Hedonic Functions and Hedonic Indexes." In J. Eatwell, M. Milgate, and P. Newman, eds., New Palgrave Dictionary of Economics, vol. 2, 630-4. New York: Macmillan.

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

Waugh, F. V. 1928. "Quality Factors Influencing Vegetable Prices." Journal of Farm Economics 10, no. 2: 185-96.