This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: Fifty Years of Economic Measurement: The Jubilee of the Conference on Research in Income and Wealth

Volume Author/Editor: Ernst R. Berndt and Jack E. Triplett, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-04384-3

Volume URL: http://www.nber.org/books/bern91-1

Conference Date: May 12-14, 1988

Publication Date: January 1991

Chapter Title: Hedonic Methods in Statistical Agency Environments: An Intellectu
Biopsy

Chapter Author: Jack E. Triplett

Chapter URL: http://www.nber.org/chapters/c5977

Chapter pages in book: (p. 207 - 238)

7 Hedonic Methods in Statistical Agency Environments: An Intellectual Biopsy

Jack E. Triplett

The Price Statistics Review Committee (1961)—usually referred to as the Stigler Committee after its chairman, George Stigler—recommended that statistical agencies explore hedonic methods, which the committee felt would provide a "more objective" way for dealing with quality change than traditional Bureau of Labor Statistics (BLS) methods. A major hedonic study by Zvi Griliches (1961), was among the staff papers published with the Stigler Committee Report. Griliches' study is—by far—the most often cited portion of the report, and it may fairly be said to have set off the entire modern literature on hedonic functions and hedonic indexes.

The term "hedonic methods" encompasses *any* use in an economic measurement of a "hedonic function,"

$$P = h(c). \tag{1}$$

In this paper I adopt the convention that capital letters designate "goods" variables and lower-case ones refer to characteristics: in (1), P is accordingly a cross section of goods prices—one P_{ijt} for each jth "variety" or "model" of the ith good or service (e.g., the prices of different models of automobiles) available at time t, and the matrix c has a row of "characteristics" for each of the same models.[1]

The first employment of hedonic methods in any official U.S. government price statistic occurred in the Census Bureau's "Price Index of New One-

Jack E. Triplett is chief economist of the Bureau of Economic Analysis, U.S. Department of Commerce.

Valuable comments were received from B. K. Atrostic, Ernst R. Berndt, and Ellen Dulberger, and from members of the 1987 National Bureau of Economic Research Summer Institute, where an earlier version was presented. Views expressed are those of the author and do not represent official positions of either the Bureau of Economic Analysis or the U.S. Department of Commerce.

Family Houses Sold," which was introduced in 1968 with data commencing in 1963. This index was adopted for the construction components of the National Income and Product Accounts (*Survey of Current Business* 54, no. 8 [August 1974]: 18–27). In fact, the single-family house price index is used as a proxy in deflating a variety of construction activities in addition to houses (see U.S. Department of Commerce 1987; and Pieper, in this volume).

The second employment of hedonic methods in a U.S. price index occurred nearly 20 years later. The BEA-IBM computer equipment price indexes (covering four different products) were introduced into the National Income and Product Accounts in the benchmark revision announced in December 1985 (see Cole et al. 1986; Cartwright 1986). Though a substantial amount of research on hedonic methods took place within the BLS or under its sponsorship from the mid-1960s on,[2] the first BLS use in an official price index came in 1988, when an adjustment for aging, estimated with a hedonic function, was put into consumer price index (CPI) housing components (U.S. Department of Labor 1988).[3]

To put it succinctly, hardly any use has been made of hedonic methods in U.S. government price indexes. The same statement holds for price indexes of other countries.[4] Why has this been the case? It is especially remarkable that the once-controversial cost-of-living (COL) index concept has been embraced in the BLS, while hedonic methods have found little role in the price statistics for which they were developed.

The following sections review reasons why statistical agencies have resisted hedonic methods. The list of factors impeding the adoption of hedonic methods is inherently speculative, in that it is based on my perception of how statistical agency operating units viewed hedonic methods as a practical device for use in constructing price indexes. The list is not derived from official agency positions or other documented sources, and the positions I outline may not necessarily all have been held by any agency manager. In particular, these positions should not necessarily be attributed to present managers.

I should note at the outset that I am not entirely unsympathetic to most of the positions I am reporting, even when I disagree with them. In each case, I review how these operating perceptions meshed, or did not mesh, with problems or findings that were present in the research literature. Many concerns of operating mangers had parallels, 15 or so years ago, in the research literature. Most of the points I discuss have now, however, been resolved and are no longer valid reasons for resistance to the use of hedonic methods.

I am also aware that some readers may regard these points simply as "excuses." Perception, or perhaps interpretation, varies with the eye of the beholder. One of the difficulties in the interaction between analytic (read "academically oriented") economists and statistical agency managers is a kind of "two different worlds" syndrome. What one thinks a fatal shortcoming, the other regards as an excuse; and what the second judges vital, the first deems trivial. Nevertheless, there is more parallelism than has usually been appreciated between the hedonic research literature and the concerns of statistical agency managers—or at least that was so 10–15 years ago. This review is intended to promote some communication, as well as to precipitate a current reassessment.

7.1 Hedonic Methods Had No Theory

The statement that statistical agency managers were reluctant to adopt a technique because it lacked theory may seem surprising. However, directly in the research institutions and indirectly in statistical agencies, hedonic methods and research were questioned for theoretical reasons.

Griliches's (1961) revival of hedonic methods arrived when economists were showing increased or heightened interest in index number theory and in more rigorous application of theory to empirical work. Nearly every graduate reading list in the 1960s contained Koopman's (1947) classic "Measurement without Theory," and the 1960s and early 1970s research ethos held that empirical relationships should be derived from theory. For a consumption price measure, this view translated into the demand that the hedonic function and hedonic indexes should be derivable from the utility function (by analogy to deriving empirical demand functions from utility functions and to deriving the form of the COL index from the consumption cost function—see the examples in Christensen and Manser 1976). Hedonic indexes were widely believed within the profession to be empirical constructions that lacked any relation to economic theory.[5]

Within the BLS, theoretical concerns took a parallel but particularized form—the view that hedonic measures could be given no conceptual interpretation within a Laspeyres-formula price index. In the 1960s the CPI was still thought of within the BLS as a separate concept from a COL index. Once the COL index was adopted as the conceptual basis for the CPI, the internal statistical agency concern matched exactly the one in the research literature (though it was not necessarily voiced in the same language).

7.1.1 Filling the Theoretical Gap

By the mid-1970s, however, the gap in theory was well on its way to being filled.[6] The theoretical relation of the hedonic function to utility and production functions was established by Rosen (1974). To outline Rosen's contribution, we first assume that the characteristics of goods, rather than the goods themselves, are the true arguments of the utility function (true inputs to the production function). This is an implication of the "hedonic hypothesis" that heterogeneous goods are aggregations of characteristics. Thus we have

$$Q = Q(c, Z) \tag{2}$$

where Q is utility (scalar output), Z is a vector of other, homogeneous goods (productive inputs), and for expositional simplicity we specify only one heterogeneous good in the system with characteristics (c). It is convenient to suppose that (2) can be written:

$$Q = Q(q(c), Z), \tag{2a}$$

where $q(\cdot)$ is an aggregator over the characteristics (c) that are embodied in the heterogeneous good. A parallel development of the theory on the producer side makes the production of a heterogeneous good the joint output of the set of characteristics it contains.

The economic behavior of buyers and sellers of heterogeneous goods can be described by sets of demand and supply functions for characteristics. These demand and supply functions are derived from the optimization of buyer's and seller's objective functions over characteristics. On the demand side, for example, $q(\cdot)$, above, carries information about preferences (using technology) and the hedonic function—$h(\cdot)$ from equation (1)—provides information about the characteristics price surface.

Rosen (1974) showed that if there are n competitive buyers, with dispersion in tastes (using technologies), the hedonic function, $h(\cdot)$, will trace out an envelope to the set of preferences (using technologies), described by the n aggregator functions, $q_1(\cdot), \ldots, q_n(\cdot)$. As with any envelope, the form of $h(\cdot)$ is thus independent of the form of $q(\cdot)$—except for special cases—and is determined on the demand side by the distribution of buyers across characteristics space. A parallel condition exists on the seller side.

As a consequence, the form of the hedonic function, $h(\cdot)$, is, in the general case, purely an empirical matter.[7] In particular, and despite many statements to the contrary that have appeared over many years, nothing in the theory rules out the semilogarithmic form, which has frequently emerged as "best" in functional form tests in the hedonic literature (Griliches 1971). The hedonic function represents a *price* surface in characteristics space, and empirically it can take on any of a large number of functional forms. It is not analogous to (say) a demand (or supply) function that is derived from a utility (production) function—nor is the hedonic function a reduced-form function of normal demand and supply functions (as is so often, but erroneously, stated in recent literature).

Understanding hedonic *indexes* required extending index number theory into characteristic space (that is, reformulating price index theory using the characteristics of goods as quantities, rather than just the numbers of goods themselves, and the characteristics prices or costs, instead of goods prices). The main reference for this extension is Triplett (1983). A summary follows.

It is well known that a COL index (input cost index) shows the minimum change in cost between two periods that leaves living standards (output) unchanged—that is, the ratio of costs of optimal points on the same indifference curve (production isoquant) under two input price regimes (see Pollak 1983). Such an index is often termed an "exact" index.

For the separable function (2a), an exact "subindex" (Pollak 1975; Blackorby and Russell 1978) can be computed that involves only the characteristics of the heterogeneous good. A subindex might be a price index for, say, automobiles, which would be computed as the ratio of the costs, under two characteristics price regimes, of two constant utility (constant output, where autos are investment goods) collections of auto characteristics (Triplett 1983, 1987). The subindex is a "constant quality" or "equivalent quality" price index because the two collections of characteristics implied by it are equivalent in utility (equivalent in producing output).

The hedonic price index can be thought of as an approximation to the exact characteristics subindex, provided conditions necessary for the exact subindex are met—that is, the utility (production) function can be written as (2a). A parallel exact characteristics price index can be developed on the output side (Triplett 1983), and again the hedonic index provides an approximation.

In view of some confusion that exists in the hedonic literature, one should note that in the general case the hedonic index is not an exact (characteristics) subindex. When the hedonic index is taken from the hedonic function (the "dummy variable" method—see below), the functional form of the hedonic index depends solely on the form of the hedonic function, which is in general independent of the form of both using and producing technologies (see above). The exact (characteristics) index, on the other hand, requires information on the utility (production) function and incorporates the effects of substitution among characteristics as relative characteristics prices change.[8] A similar statement can be made for the output price index case. Special cases exist for which the hedonic index and the exact characteristics subindex for inputs—or the hedonic index and the exact output price index—coincide, but these cases are ignored here.

A recent statement of these developments in the conceptual foundations for hedonic functions and hedonic indexes is Triplett (1987); see also the "Summary of the Theory" section in Triplett (1989). I do not mean to imply there are no unresolved problems. However, by the mid- to late 1970s, to say that hedonic methods had "no theory" was no longer correct.

7.1.2 Empirical Consequences of the Theoretical Gap

Historically, the no-theory perception inhibited empirical research on hedonic methods, both inside statistical agencies and outside them. In academic circles, such work was thought, in some sense, not "respectable," and in fact little hedonic price index research appeared in the journals, particularly after about 1973 or so.[9] With that attitude in the profession at large, one did not find good graduate students in the 1970s choosing dissertation topics on the subject, and, accordingly, there was no stock of such researchers from which to recruit for work inside statistical agencies. The contrast between the availability of complementary research outside statistical agencies for empirical estimation of COL indexes and the absence of it in the case of quality change and hedonic methods is striking.

The perception that hedonic methods had no theoretical foundation also, I am convinced, lowered the quality of the hedonic research that was done: it was all too easy to rationalize a hedonic function whose variables had little or no relation to the technology being investigated, on the grounds that all one wanted was a maximum R^2. That, in my opinion, resulted in a large quantity of poor work.

One should not, however, overemphasize the negative effects of the "no theory" perception. It is always valuable to straighten out theoretical issues. Theoretical thinking about hedonic functions in the late 1960s and early 1970s contained too many theoretical "proofs" that were not, and "impossibility theorems" that were irrelevant. In its present stage of development, the theory of hedonic functions provides a useful guide to empirical research—an implication of the theory, for example, is that the arguments of hedonic functions are technical or engineering variables. Theory also provides guidance on the search for functional form and on the appropriate construction of hedonic indexes.

7.2 The Perception That Hedonic Methods Required That Price Index Calculation Procedures Be Changed

In most early hedonic research, price indexes were calculated by the "dummy variable method"—a time dummy variable, or a series of them, inserted into a "pooled" cross-section regression (that is, eq. [1], with two or more periods of data) estimates the price change that is *not* accounted for by changing characteristics. This remained the dominant approach in the research literature even though, as Griliches (1971,7) remarked, the dummy variable method was "not well articulated with the rest of the index number literature."

Within statistical agencies the perception formed that adopting hedonic methods meant estimating the price index from a regression, as opposed to the traditional calculation of matched-models, matched-outlets price relatives. Indeed, within statistical agencies, the hedonic technique was usually referred to as *the* "regression method," and hedonic indexes as "regression indexes," or indexes computed by "the regression method."

This statistical agency perception was always quite wrong. Griliches (1961), for example, used his automobile hedonic functions in several ways to calculate price indexes; his were not exclusively dummy variable price index estimates. The Census Bureau's price index for new single family houses is not a dummy variable index; it is a price index for characteristics.[10] Triplett (1971) discusses in a general way how to integrate hedonic methods into traditional BLS price index procedures, and Triplett and McDonald (1977) demonstrated empirically how to use a hedonic function within existing producer price index (PPI) methodology to make quality adjustments to the price quotes gathered for the PPI (Early and Sinclair 1983 followed the Triplett-McDonald procedure). The IBM-BEA computer equipment price indexes (Cole et al.

1986) use the hedonic function to impute prices for "missing" computers—those observed in one period but not in another—within an index that is calculated by computing price relatives for matched models and (because they are used as deflators in the National Income and Product Accounts, or NIPAs) a Paasche price index formula.[11] All these calculations fit the definition of hedonic methods (see above), they can all be described as hedonic indexes, and all fit comfortably within established statistical agency procedures for producing price indexes.

It is thus a little puzzling that the perception lingered on so long that hedonic methods required altering index calculation procedures. Hedonic methods affect only the way quality change is evaluated.

Actually, though, the reverse side of the same perceptual shortcoming applied outside statistical agencies. There was far too little concern given in the research literature to the form of the hedonic price index—that is, to the way one goes from the estimated hedonic function to the hedonic price index.

Most researchers constructed a hedonic price index by the dummy variable method and compared the results with some relevant published statistical agency index (and sometimes with some that were not relevant). They called the difference between the two the "effect of quality change." Few considered that the difference, or part of it, might also have been attributable to the fact that the implicit index formula for calculating the (dummy variable) hedonic index differed from the one used by the statistical agency. Equally few researchers considered the robustness of hedonic indexes to variations in the way the hedonic function was employed to create a "quality adjusted" index—despite the good example set by Griliches (1961, 1964). Statistical agency managers were more aware of these points (and even might have overemphasized them) and were accordingly less impressed with the evidence presented by researchers than were many of those researchers themselves.

Thus, statistical agency reservations about dummy variable hedonic indexes had parallels in attitudes of researchers outside statistical agencies. Both groups were wrong in their own ways. The idea that adopting hedonic methods forces a change in calculation procedures has no validity. Quantification of the empirical effects of alternative ways of using hedonic methods, however, is a research topic that has been neglected. Indeed, it is peculiar that, with all of the focus on index number formulas that occurred in the index number literature, the question of index number form was almost entirely ignored when researchers turned to quality change.

7.3 The Perception That Hedonic Indexes Were More Sensitive to Arbitrary Research Procedures Than Were Traditional Approaches

This very widespread perception within statistical agencies drew on some of the evidence from research studies, yet the conclusion is unsupportable.

The lack of robustness in some of the empirical hedonic indexes that have appeared in the economics literature made agency managers very nervous indeed. Cases exist in which different hedonic price indexes were submitted by the same researcher from the same data, using the same dummy variable method, with outcomes that were quite far apart. The suspicion arose that other investigators got similar dispersions and might not have published them. Also, initial, usually unpublished, trials within some statistical agencies obtained unsatisfactory results, including poor fits, wrong signs, and implausible indexes.

On this score, statistical agency managers perceived correctly that empirical hedonic indexes sometimes lacked sufficient robustness to be reliable. Many of the published hedonic studies simply were not very good. Some foundered because researchers did not have access to good quality data. Their cross sections of prices were usually published list prices, with some unknown sets of errors with respect to transactions prices that were probably (from the evidence) correlated with the explanatory variables. Researchers often paid little attention to the selection of explanatory variables and too uncritically accepted published (in trade journals and the like) values on the explanatory variables they chose, without checking the accuracy of published information (against, e.g., manufacturers' information). Even a cursory review of the hedonic literature suggests the need for much more care in the choice of variables to serve as characteristics and the need for more effort on the part of economic researchers to understand the technology of production and use of the product (in order to choose appropriate variables). Moreover, researchers often failed to present the effect of some of their data decisions on their price indexes.

Put another way, research methods for producing valid hedonic functions were not written down anywhere in the literature (and still are not). Perhaps the best examples of such methodological discussion, combining technological knowledge of the product and the economics of hedonic functions, is in the literature on computer hedonic functions (see Fisher et al. 1983; Cole et al. 1986; Dulberger 1989; Flamm 1987; and Triplett 1989).

A related point is that researchers have often presented or used hedonic index variances in inappropriate or irrelevant ways. One usually sees in research studies, for example, a test of the null hypothesis that quality-adjusted (hedonic) indexes do not differ from zero, a test that is obviously not invariant to the true rate of inflation. Because the hedonic technique is a mechanism for adjusting for quality change, it is more appropriate to test the null hypothesis that the hedonic index does not differ from an index, computed from the same data, that has no quality adjustment—in other words, to test the statistical significance of the hedonic quality adjustments, not the statistical significance of the measured rate of inflation. When researchers noted with satisfaction that their hedonic indexes were "significant," they were usually reporting only that the rate of inflation was positive, no matter how measured; the actual "hedonic" part of a good many hedonic studies lacked statistical significance.

It is not surprising that poor hedonic studies tarnished the reputation of the method: poor research can sometimes be as influential as good. Yet, the conclusion that hedonic methods were more arbitrary than conventional approaches to quality change cannot follow from one-sided evaluation of the poorer of the hedonic studies.

For one thing, the robustness of conventional methods for compiling price indexes is not known because there is little or no information on the subject. Consider a possible robustness test of the conventional method: one could assemble alternative teams of BLS commodity specialists, give each team the same information on examples of quality change, and ask them to reach independent judgments about how the examples should be treated in the indexes. I predict that the teams would sometimes reach different outcomes and that the outcomes would, in many cases, produce perceptible effects on the indexes. There is thus a stochastic element to the conventional quality-adjustment process, in the sense that repeated trials yield different outcomes. Those outcomes could be used to produce a quasi-variance estimate for this part of the price index calculation procedure. Though I have no idea whether this variance component would be larger than the comparable variance one gets in a hedonic index, a comparison of the two would prove quite interesting and should be carried out by some statistical agency.

Second, the usefulness of hedonic methods should, of course, be judged on the potential of the best hedonic studies, not on the poorer of them. Though statistical agency managers perceived correctly the inadequacies in a good many published hedonic studies, the better studies show that hedonic methods have great potential for improving measurement.

7.4 Hedonic Functions Need Large Cross Sections of Transactions Prices

This topic reflects another anomaly in the research literature: there has always been much—valid—concern that price quotations gathered for the wholesale price index (WPI), the PPI, and other price indexes might represent list rather than transactions prices. Yet, researchers on quality change were too frequently content to produce a hedonic price index that was nothing more than a quality-adjusted list price series. In this case, statistical agency managers were correct in perceiving the potential error of such approaches, and many outside researchers were too cavalier (Zvi Griliches was, always, an exception; see Griliches 1961).

It is true that on certain assumptions one might form quality *adjustments* from cross sections of list prices and apply them to transactions price quotes obtained for the indexes (this was the approach of Triplett and McDonald 1977). But one never knows the size of biases that ensue when the assumptions do not hold.

Production use of hedonic methods does require cross sections of transactions prices, gathered at least periodically, and such data have seldom been available, even in government price programs. Lack of the required data inhibited use of hedonic methods.

Gathering cross sections of transactions prices, even on a periodic basis, would be both expensive and burdensome to respondents. On the other hand, collecting cross-section information to improve quality adjustment in the price indexes has never received managerial consideration in the BLS, at least since an unsuccessful effort in the mid-1960s.

7.5 The Perception That Automobiles Were the Test Case

Automobiles were the subject in much of the early exploratory work on hedonic price indexes. For a number of reasons they were a poor test case.

A great amount of effort has gone into adjusting government automobile price indexes for quality changes in "new models." The BLS staff faced many problems for which available procedures were recognized as inadequate. However, existing hedonic functions for automobiles contained little potential for resolving the measurement problems that have arisen in automobile price indexes since at least the mid-1960s. It was perhaps therefore too easy—and certainly incorrect—for agency managers to decide that the auto studies proved that hedonic methods were not useful. Because the relation between hedonic measures and the adjustments that were actually performed in the auto indexes is not well understood, and because that relation is important in determining the potential of hedonic methods to improve the indexes, it is worth considering quality change and the automobile indexes in more detail.

7.5.1 Cost-based Quality Adjustments in Automobiles

Since the early 1960s, quality change in automobiles has been handled in BLS price indexes (both the CPI and the PPI are handled similarly) by obtaining production cost information from manufacturers. For example, cost-based quality adjustments for 1988 model cars accounted for 54% of the recorded $400 increase in average auto prices collected for the PPI at new model introduction in October 1987 (table 7.1). This production cost method was instituted (perhaps expanded is a better word) in response to the Stigler committee's judgment that BLS methods for treating quality change were inadequate.

A three-step sequence occurs in the use of manufacturer's cost information. (*a*) For each car included in the indexes, a detailed list of engineering and specification changes is obtained from the manufacturer at the beginning of the model year. (*b*) Information (obtained from the manufacturers and other sources) about each of these changes is used by BLS staff to determine whether each of the changes is to be treated as a quality change. At various times, a set of internal "guidelines" have spelled out the principles governing this stage, but the guidelines are general rules, not specific ones, and in most

Table 7.1 **Values of Cost-based Quality Adjustments in Sample of Automobiles in Producer Price Index, Model Years 1967–88**

	Manufacturer's Level			Retail Level		
Automobile Model Year	Average Value of Quality Change	Of Which, Value of Mandatory Changes[a]	Average Price Increase	Average Value of Quality Change	Of Which, Value of Mandatory Changes[a]	Average Price Increase[b]
1988	$214.94	n.a.	$399.01	$245.56	$65.42	$458.66
1987	37.89	n.a.	694.29	47.13	1.92	776.38
1986[c]	154.55	n.a.	616.86	186.50	27.42	745.52
1985[d]	125.52	n.a.	232.50	151.45	20.02	268.20
1984	91.87	n.a.	183.65	110.08	46.80	221.70
1983	107.66	n.a.	215.55	128.04	64.65	263.92
1982	104.70	n.a.	463.61	126.32	84.68	562.54
1981[d]	438.39	n.a.	664.57	530.85	470.94	536.14
1980	195.19	n.a.	n.a.	241.51	131.33	365.85
1979	37.00	n.a.	n.a.	46.35	17.85	300.30
1978[d]	40.88	n.a.	n.a.	50.12	9.99	424.49
1977[d]	47.05	n.a.	n.a.	59.19	21.21	322.30
1976	12.00	n.a.	n.a.	15.60	21.00	198
1975	102.30[e]	n.a.	n.a.	129.90[e]	147.20[e]	386
1974	91.30	n.a.	n.a.	117.90	109.00	n.a.
1973	95.40	n.a.	n.a.	123.80	113.30[f]	n.a.
1972	n.a.	n.a.	n.a.	20	9.00[f]	19
1971	n.a.	n.a.	n.a.	25[e]	29[e]	220
1970	n.a.	n.a.	n.a.	46	13	107
1969	n.a.	n.a.	n.a.	1	14	41
1968	40.05	$40.75	87.54	57.00[g]	58.00[h]	n.a.
1967	n.a.	n.a.	n.a.	55[i]	—	55

Sources: Annual Bureau of Labor Statistics press releases on quality change in new model automobiles, 1966–87.

Note: n.a. = not available.

[a]cost of changes to meet federal smog, safety, and fuel-efficiency standards.

[b]In most, if not all, cases, average change in manufacturer's suggested retail (list) price.

[c]Calculating procedure changed: under the former system, the five values in the 1986 row are $150.11, $402.68, $181.22, $27.42, and $482.03.

[d]One or more cars in the sample were downsized in this year. In most cases values for these cars are excluded from the quality-change data.

[e]Includes an additional quality adjustment made after new-model introduction, and reported in a subsequent press release.

[f]Includes "voluntarily added" equipment in anticipation of future increases in standards.

[g]Estimated by the author by assuming the −$0.70 manufacturer's value for nonmandatory quality change (col. 2 less col. 1) would amount to −$1.00, retail.

[h]Reported in a subsequent note; press release gave manufacturer-level values for this year.

[i]Press release does not give a value, but rather states that the quality adjustment was equal to "practically all" of the price increase.

cases the final decision about whether a particular engineering change is or is not a quality change rests on staff judgment. As an example, at one point a company's substitution of a digital clock for an analog clock was judged by BLS staff as a styling change, not a quality change, so any price differential associated with the new clock was allowed to pass forward into the indexes. (*c*) *After* the determination in (*b*), the production cost of each of the accepted changes is used to adjust new car prices.

In the early years of the process, the full value claimed by manufacturers became the quality adjustment, but by the early 1970s, if not before, manufacturers' claims were often not fully allowed by the BLS staff.[12] In a number of cases, model changes in autos have been regarded as too extreme to apply the cost procedures. An example was the so-called downsizing that occurred on some domestic cars beginning with the 1977 model year, where the new models were smaller externally, but offered the same or more interior capacity (refer to table 7.1, n. *d* above). In these cases other quality-adjustment methods were substituted (usually, imputing the price change, after cost-based quality adjustments, from another car that was less fundamentally changed).

Engineering changes that occur from one model year to the next, even on cars that are not substantially changed, are complex, perplexing, and multitudinous. They were often hard for a staff without engineering expertise to evaluate, and sometimes were not easy for the manufacturer to fit into price index objectives. In one example, an inexperienced auto company executive had great difficulty locating for BLS staff a brace that was added one year to reduce transmission vibration and also spent some effort searching for the "roll center," which the company claimed it had altered for the new model year (the roll center is an imaginary point—the center of the arc described by the body of the car as it leans into a corner). Some early attempts to obtain evaluations from government agencies involved in automobile regulation failed.

Moreover, it is clear that the cost data provided by manufacturers are frequently not the relevant costs, even for a "resource cost" adjustment (see below). There are thus many reasons for dissatisfaction with cost-based quality adjustments.

7.5.2 Technical Change and Automobile Hedonic Functions

If one were to evaluate most of the automobile quality changes for which the BLS has made a quality adjustment over the past 25 years, using any automobile hedonic function that has appeared in the literature, one would conclude that the individual changes were frequently too small to justify an adjustment. That is, most of those changes in specifications, when introduced as variables into a hedonic function, would have insignificant coefficients. Yet, in total those changes have involved substantial adjustments to car prices, amounting to several hundred dollars in some of the last 20 years (see table 7.1). It is doubtful that one wants to accept conclusions from the hedonic function that these cases should have been ignored. If "quality" encompasses a very large number of characteristics, or the product is complex in its use, the hedonic function suffers from missing variable bias.

The biggest problem with cost-based adjustment, however, centers on the appropriate treatment of mandatory antipollution and safety equipment and, in the second half of the 1970s, engineering changes that were necessary to meet federal fuel economy regulations. These mandatory changes account for a large share of the cost-based adjustments (see table 7.1).

Hedonic studies provided no information to adjust for mandatory changes. In part, the hedonic measures might have been right and the BLS wrong. It is certainly easy to argue that mandatory air pollution devices, for example, should have been treated as a tax on transportation rather than as an improvement in the quality of transportation services that automobiles render to consumers and to business users. Similarly, the right way for improved fuel efficiency to enter a COL index is through a reduction in gasoline consumption, and not by reducing the price of cars by an amount equal to the "cost" of manufacturing fuel-efficient engines.

In summary, the automobile hedonic functions in the *existing* literature did not provide BLS managers with any information at all on the most difficult of their problems.

7.5.3 The Automobile May Be too Complicated for Hedonic Studies

What one might call the 1960s hedonic technology (which really does not differ from Court's initial study in the 1930s) defined automobiles as functions of size and, especially, weight (Griliches 1961; Triplett 1969; Dhrymes 1971). The weight of the car is obviously a proxy for a large number of other characteristics. It has the undeniable advantage that most items of equipment (an air conditioner, say, or a tape deck, or better insulation against noise and vibration) have a weight penalty. Weight can serve as a proxy for a very large number of separate characteristics that could not feasibly be entered into a single regression.

The difficulty is that weight is an unreliable proxy, precisely because weight for its own sake is undesirable (as Court noted in his original paper in 1939). Periodic engineering innovations have reduced the ratio of weight to the characteristics that are truly desired. Automobile hedonic functions based on weight give biased price indexes.

It is not all that hard to improve on the 1960s hedonic technology. Instead of taking weight and the external size of a car, one can specify a hedonic function that is defined on the auto's internal passenger- and luggage-carrying volume, plus a small number of other characteristics. The "fit" for such a function is as satisfactory as for the 1960s version hedonic functions. Measures of cornering, braking, acceleration, ride quality, and even quietness are available from various test programs, and more careful modeling of the automobile as a consumer product (or as an investment good providing input services to production) might yield more believable results that have so far appeared in the literature.

However, the bane of automobile hedonic functions in the past has been the degree of multicollinearity among explanatory variables. Multicollinearity, because it leads to the exclusion of important characteristics, assures that changes in the omitted characteristics can swamp the effect of the included ones, without providing a clear signal to the investigator. One might not care to argue that, from an engineering standpoint, the automobile is more complicated than, say, a computer or an airplane; however, the way automobile characteristics enter the utility function—what the automobile does for its user—is in fact very complicated indeed, and very hard to model, and for this reason the appropriate set of variables is hard to determine. It is in this sense that automobiles may have been the wrong place to start because more credible results can be obtained from hedonic studies on other products.

7.6 The Perception That Hedonic Methods Measured User Value, Not Resource Cost

Actually, I doubt that this factor significantly inhibited the adoption of hedonic methods. It has, however, often been so perceived, and because so much ink has been spilled over the user-value, resource-cost controversy one can hardly review hedonic methods without discussing this set of issues.

Andrew Court himself was probably the originator of the resource-cost, user-value debate. He selected the name "hedonic" because of his belief that his new indexes measured value to the user: "Hedonic price comparisons are those which recognize the potential contribution of any commodity, a motor car in this instance, to the welfare and happiness of its purchasers and the community" (Court 1939, 107).

The view that hedonic indexes carried a user-value interpretation has lingered on through the years. It was adopted by the Stigler Committee, for whom the user-value interpretation was a desirable property (because under this interpretation use of hedonic methods in the CPI would move it in the direction of a COL index). Others who favored the use of hedonic methods have accepted the same interpretation.

The view that hedonic methods represented uniquely user-value measures was also held by some professional critics, who argued that resource cost, not user value, was the appropriate criterion to use for quality adjustments in price indexes for the national accounts. Denison (1957), Jaszi (1964), and Gilbert (1961) were among the economists who took this position.

I emphasize that both proponents and opponents of hedonic methods in the 1960s shared the common view that they represented a user-value approach to measuring quality change. Historically, proponents of hedonic methods almost invariably advocated a user-value quality standard; opponents have favored the resource-cost standard. The resource-cost, user-value debate was spirited and even acrimonious at times, and has often been interpreted as aligning statistical agencies on one side and academic users on the other.

7.6.1 Statistical Agency Positions

There is, however, no evidence that the equation of hedonic methods with user value had any effect whatever on the willingness of statistical agencies to adopt hedonic methods. Under the federal government's decentralized system for producing economic statistics, the BLS has responsibility for price indexes. During the entire period under discussion, the BLS accepted value to the user as the appropriate quality criterion for the CPI and—at least up to 1978—for the WPI as well.[13] Thus, if it were true that hedonic methods measured user value, then the technique fit the BLS's own view of what was conceptually appropriate for its price measures. In my association with the BLS (which began on an intermittent basis in 1968), I never on any occasion heard "hedonic methods equals user value" raised as an objection to their use in BLS price indexes.

When the Census Bureau's new single-family house price index was announced in 1968, it was described as a response to the Stigler Committee's criticism of construction price statistics (Musgrave 1969). So far as I have been able to determine, the user-value, resource-cost controversy was never an issue in the development of the single-family house index, nor was it a reason that the Census hedonic index program was never extended, as originally promised, to other construction activities.[14]

Finally, whenever a hedonic index has become available, and has proven superior to the alternatives, the Bureau of Economic Analysis has used it in the NIPAs. For example, the article that describes the introduction of the Census new house price index for NIPA construction deflation (*Survey of Current Business* 54, no. 8 [August 1974]: 18–27), refers to hedonic methods as measuring "the current price that the purchaser implicitly pays for each of [the] characteristics . . ."; this language is consistent with the then-prevailing user-value interpretation of hedonic methods. Concerns of a conceptual nature, if there were any at the time, appear to have been sublimated to the pragmatic need for improved data.[15]

It is certainly true that a vigorous debate over the proper treatment of quality change in economic statistics was carried on between (roughly) the mid-1960s and the mid-1970s. It is also true that some of the participants in the theoretical debate were identified with statistical agencies. There is little evidence, however, that the theoretical debate had much to do with agencies' willingness to adopt hedonic methods. My belief is that the factors discussed elsewhere in this article were far more important.[16]

7.6.2 The Resolution of the User-Value, Resources-Cost Debate

That this particular conceptual debate has not been a major factor in the adoption of hedonic methods is ironic. Our current understanding of hedonic methods shows that the identification of hedonic measures uniquely with user-value measures is incorrect; they provide, in most cases, approximations to both user-value and resource-cost concepts of quality change. Moreover, resource-cost and user-value concepts are not competitive (in the sense that if the one is right, the other must always be wrong); rather, each of the two concepts corresponds to a particular use of the data. User value is conceptually correct for quality change in a COL index (and thus for the CPI) or for a measure of investment or of productive inputs or their prices. Resource cost is the conceptually correct quality change concept for measures of output or for output price indexes (e.g., the revised PPI is notionally an output price index).

One important theoretical result is Rosen (1974), as noted above, who showed that in the competitive case the hedonic function provides estimates of the incremental acquisition cost of, and also revenues from, characteristics. Accordingly, implicit hedonic "prices" can serve as approximations either to user-value or to resource-cost valuations, an interpretation that is analogous to our normal interpretation of prices of goods. Jaszi (1964) had already noted that the effects of differing mixes of characteristics on the price of the product (he used as an example the proportion in a coffee-chickory mix) would be determined in equilibrium according to the costs of the separate components. Though Jaszi used his example to argue that hedonic methods should give the same measure as conventional approaches, his example also implies that economic forces assure that production cost will be reflected in the coefficients for hedonic function variables. From this, one can make the further argument that in competitive equilibrium hedonic prices will reflect both marginal production costs and incremental user values, so that the interpretation of a hedonic price is similar to the interpretation of any other price under competitive conditions. Thus, the presumption shared by both sides in the user-value, resource-cost debate (that hedonic methods provided uniquely user-side measures) was misconceived.

That the mid–twentieth century debate about the interpretation of hedonic prices parallels mid–nineteenth century debates about the theory of value suggests how difficult it is for economists to shift their mental gears from goods to characteristics. Marshall's two scissors blades cut in characteristics space in exactly the same interdependent mode of operation as they have long been known to function in goods space.

The question of which of the two criteria—user value or resource cost—was theoretically or conceptually *correct* was central to the controversy and was a more difficult issue to resolve. Fisher and Shell (1972) were the first to show that different index number measurements (they considered output price indexes and consumer COL indexes) imply alternative theoretical treatments of quality change, and that the theoretically appropriate treatments of quality change for these two indexes correspond, respectively, to "resource-cost" and "user-value" measures. Triplett (1983) derives this same result for cases where "quality change" is identified with characteristics of goods—and therefore with empirical hedonic methods; the conclusions are that the resource cost of a characteristic is the appropriate quality adjustment for the output price index, and its user value is the quality adjustment for the COL index or input cost index.

Intuitively, these conclusions are appealing. The output price index is defined on a fixed value of a transformation function. The position of a transformation function, technology constant, depends on resources employed in production; accordingly, "constant quality" for this index implies holding resources constant, or a resource-cost criterion.

On the other hand, the COL index is defined on a fixed indifference curve, and the analogous input-cost index is defined on a fixed (user) production isoquant. For these two "input" price indexes, "constant quality" implies holding utility or output constant, or a user-value criterion (an extended discussion is contained in Triplett 1983).

The debate on this subject sometimes generated more heat than light because (*a*) it was not recognized that there were, in effect, two different questions and accordingly two correct answers, not one; and (*b*) as already noted, there was an inappropriate linking, on both sides of the debate, of hedonic indexes with the user-value criterion. It was thus thought, incorrectly, that use of hedonic methods in an economic measurement implied accepting one of the two theoretical positions over the other one.[17]

7.7 The Position That Traditional Methods Should Yield the Same Measurement as Hedonic Methods

I list this one toward the end because it is a relatively recent line of reasoning. In equilibrium, the argument goes, the quality-corrected prices of all varieties ought to be equal (that is, the "quality ratio" should always equal the price ratio of any two varieties); therefore, price movements of varieties or models that have not been changed (matched models) can stand for those models in which quality change has been observed. Traditional "linking" (described below), sometimes termed a "matched models" index—more appropriately, a "matched models only" index—will always give the correct answer and hedonic methods are unnecessary (see also Jaszi 1964).

Linking, as a quality-adjustment method, takes the following form. Suppose that a particular variety or model of the *i*th good, call it Y_{i1}, is selected for pricing accordingly to probability procedures in one of the indexes. Suppose further that Y_{i1} disappears from the outlet from which prices are being collected and is replaced in the second period by a second variety or model, Y_{i2}. In most realistic situations, we have only the price of variety Y_{i1} in period 1 (P_{i11}) and the price of variety Y_{i2} in period 2 (P_{i22}). "Linking" introduces the new price into the index in such a way that the unadjusted price ratio P_{i22}/P_{i11} does not determine the movement of the index.

For historical reasons that seem lost in the mist of time, the procedures differ in the CPI and the PPI. The CPI linking procedure imputes price movements in other (unchanged) varieties—in this example, designated j—to the variety that changed; the implicit "quality adjustment" is

$$(3) \qquad A_1 = P_{i22} / P_{i11} - \Sigma_j W_j (P_{ij2} / P_{ij1}), \quad j \neq 1,2,$$

where the weight, W_j, is the sampling weight for observation j—or just $1/(n - 1)$, ignoring sampling considerations.

The PPI follows the old WPI procedure, in which the entire price change P_{i22} / P_{i11} is attributed to quality change. It is thus assumed that no price change took place in the item whose quality changed. The PPI procedure implies that the quality adjustment is

$$(4) \qquad A_2 = P_{i22}/P_{i11},$$

so that the quality-adjusted price change for this observation is unity.

In either case, the error that quality change puts into the index occurs when price change for new models (i.e., the quality-corrected price ratio for variety $i2$ compared with $i1$) differs from what is implied by the adjustment—or, what is the same thing, when the true quality change, A_T, differs from A_1 (in the CPI) or A_2 (in the PPI). Matched-models-only price changes may be biased because only price comparisons for models that can be "matched" exactly in the two periods are accepted for the index, and the implicit quality adjustments the procedure produces for unmatched varieties does not yield their true price change.

Note that the matched-models-only index is biased by quality change *even if* the statistical agency detects and "links out" correctly every example of quality change that takes place. Moreover, the "tighter" or more narrowly drawn the product specification, the larger the number of price observations that will be rejected for failing the exact-match test and, accordingly, the larger the bias from this source. Note also that the direction of bias is unknown: the sign of the bias depends on the sign of $(A_1 - A_T)$—or the sign of $(A_2 - A_T)$, in the PPI case—and *not* on the sign of A_T. Though quality may be improving (A_T positive), the adjustment implied by equations (3) and (4) may be too large, biasing the index downward.

Whether "linking" invariably works—that is, whether the bias from matched-model-only pricing is small—is an empirical issue, on which there is relatively little evidence. The most careful comparison of hedonic and traditional matched-model linking methods is contained in Dulberger (1989).

Dulberger computed hedonic price indexes for computer processors, using three different methods. As indicated in table 7.2, the three hedonic indexes, though not identical, indicate that computer prices fell about 90% over this interval.[18] A traditional matched-model or linked price index, computed from the same data, fell by two-thirds (67%), considerably less than the price decline recorded by the hedonic computer indexes. This is thus strong evidence that hedonic and linking methods will not always produce the same answers.

Table 7.2 Alternative Estimates Compiled from the Same Data, Price Change for Computers (1972–84 percent change)

	Hedonic Indexes		
Time Dummy[a]	Characteristics Price[b]	Composite (Imputation)[c]	Matched-Model[d]
−92.2	−89.8	−90.3	−66.6

Source: Dulberger (1989), table 8.
[a]Computed from coefficients of time dummy variables inserted in a multiyear hedonic regression (similar to eq. [1] in the text).
[b]Computed from coefficients of the characteristics in a hedonic regression similar to eq. [1] in the text.
[c]Computed from prices of matched models, where "missing" prices (for "new" or "discontinued" models, present in one year but not the other) were imputed from the hedonic function.
[d]Chained index of matched models, no use of hedonic function ("new" models introduced by linking).

On the other hand, the position that *in equilibrium* the two approaches should yield the same measurement is not refuted by Dulberger's research. Her hedonic regression contains a device for testing for equilibrium—defined in her case as failure to reject the hypothesis that the price/performance ratio of "old" computer models had been bid down to equal that of newer computers that embodied the latest technology. When this price/performance definition of equilibrium obtained for computer processors, hedonic and matched-model computer indexes tended to coincide. When the equilibrium hypothesis was rejected, the price movements recorded by the matched-models index often differed greatly from those of the hedonic indexes.

Some reservations about the result should be recorded. The research used list prices. Discounts on older machines are probably more prevalent than on newly introduced ones, so that, even on Dulberger's definition, a smaller disequilibrium would have been measured had transactions prices for computer processors been available; accordingly, smaller differences between hedonic and matched-model methods might have been recorded had computer transactions prices been employed in Dulberger's research.

In a second study, on semiconductors, Dulberger (1988) reports a similar finding (see table 7.3). In this case, the PPI index (produced with conventional methods) moved, approximately, with a measure of the unadjusted average price per chip. Careful matching on a single characteristic of the chip (kilobits) gave an index—the right-hand column of table 7.3—that declined much more than the average price per chip. Note also (see table 7.3) that the differences are *very* large: The price index that controls for kilobits drops *eight to nine times* as much as the PPI index between 1975 and 1982. Though this study was not, strictly speaking, a hedonic one, average price per kilobit can be thought of as approximating a crude one-variable hedonic function, so the results are suggestive.

Table 7.3 Alternative Price Indexes for MOS Memory Semiconductors (Shipments of U.S. Manufacturers)

		Dataquest		
Selected Years	Published Producer Price Index (BLS 11784221)	Average Price/ Chip[a]	Average Price/ Kilobit[a]	Laspeyres Matched Models[a]
1975	212.1	313.5	1,846.2	1,662.5
1978	168.8	147.5	579.5	452.5
1980	129.7	213.9	371.4	344.2
1982	100.0	100.0	100.0	100.0
1985	73.4	80.1	22.8	30.2
1986	61.7	89.4	12.7	23.3
1987	62.8	111.2	11.4	23.2

Source: Dulberger (1988).
[a]Weights are U.S. value shares from Dataquest.

The computer and semiconductor empirical studies demonstrate the danger in relying on equilibrium assumptions in measuring prices, especially for technologically dynamic products.[19] The "traditionalist" or "equilibrium" position amounts to stating that the differences recorded in tables 7.2 and 7.3 should not exist; since they do exist empirically, the traditionalist position does not provide a compelling argument for rejecting hedonic methods.

7.8 Hedonic Methods Give Price Indexes That Fall "Too Fast"

One hears this position more frequently from index users than from price index producers, but it demands consideration here. Government price indexes that are not accepted as meeting user needs are deficient, no matter what their producers believe of them.

Most often cited as "falls too fast" examples are the IBM-BEA computer price indexes (Cole et al. 1986; Cartwright 1986), which are hedonic indexes and show almost unprecedented declines (see table 7.4). Such price behavior is not a recent phenomenon: Research studies on computer equipment show comparable price declines going back to the birth of the electronic computer.[20]

When shipments data are deflated with indexes that drop so far and so fast as do the hedonic computer price indexes, the resulting quantity measures grow very rapidly indeed. Some business economists have reportedly argued that when government "real" quantity data for producer durable equipment are compared with various business records, computer growth in the government data seems high relative to that of other equipment. Note, however, that price indexes for other "high tech" equipment are produced with conventional methods: distortion in the deflated quantity data could arise from upward biases in the conventional indexes as well as from downward bias in the computer hedonic indexes.

A similar assessment comes from compilers of the Federal Reserve Board's Industrial Production Index. The IBM-BEA computer indexes were not used in the IPI because would they allegedly create output growth in computers that is "too large."

Denison (1989) points out that the IBM-BEA computer price indexes create productivity growth in computers that is far larger than in other producer durables. Had computers behaved as other producer durables (that is, had their prices risen) or had their price indexes been compiled as have those for other capital equipment (e.g., the matched-model computer price index in table 7.2), the recent divergence between productivity rates in manufacturing and nonmanufacturing would have been diminished—the manufacturing productivity rate would be lower because output growth would have been lowered.

I do not know any independent data with which to test the view that computer price indexes decline too fast—or that real output or productivity in computers rises too rapidly—when hedonic indexes are employed. The nearly 30 studies reviewed in Triplett (1989) include some conducted by computer scientists and engineering technologists whose objective was to measure the rate of the computer's technical advance, not to produce price indexes. Authors of the computer science studies (which give results consistent with the

Table 7.4 "Fixed-Weight" NIPA Price Index for Computers[a] (1982 weights, 1982 = 100)

Year	Index
1972	566.2
1973	567.1
1974	498.0
1975	448.3
1976	414.4
1977	308.3
1978	186.0
1979	158.5
1980	123.6
1981	108.3
1982	100.0
1983	83.9
1984	70.0
1985	55.7
1986	47.9
1987	41.9
1988:III	36.2(p)

Source: Bureau of Economic Analysis.
[a]"Computers" includes processors and major items of peripheral equipment (see Cartwright 1986) combined with 1982 shipments weights.

economic ones)[21] did not seem to find their results implausible. To cite the very evidence being disputed, even if originating from a technological discipline, is not, of course, compelling.

All the critical positions noted above depend, in some degree, on the idea that the computer indexes are suspect because "nothing else" shows similar behavior. If one uses, for example, conventional measures of computer industry inputs and the BEA-IBM price indexes to deflate the industry's output, then productivity in computer manufacturing is very great indeed.

On the other hand, semiconductor inputs are clearly a major source of technical change in computers. If the PPI semiconductor indexes are replaced by measures like those of table 7.4, the effect is to reallocate part of the measured productivity improvement from computers to semiconductors (because real input measures to computers grow more rapidly when a falling semiconductor price index is used for deflation). It is far from clear that this reconstructed picture is implausible.

Though the use of "quality-corrected" computer price and output measures along with other data that are poorly or inadequately measured may introduce some distortions in the allocation of productivity across industries, it seems doubtful that a better picture would emerge if all the data were consistent but wrong. Additional research on computer prices, output, and productivity would be of considerable interest.[22]

7.9 Conclusions

The adoption of hedonic methods has been impeded by conceptual issues, by doubts about the validity of hedonic indexes (especially when they differ greatly from conventional measures), and by the lack of cross sections of transactions prices needed as dependent variables. Conceptual issues have mostly been resolved, and should no longer pose any barrier. Most of the old validity issues have likewise been disposed of, though some remain unresolved. The data problem remains formidable, but of course a data-gathering exercise is well within the jurisdiction of statistical agencies and could be undertaken if they were convinced that hedonic methods would improve price indexes.

Notes

1. There is no standard terminology in economics for heterogeneity in goods and services. I use "variety" and "model" interchangeably as synonyms. The term "model" is customary for durable goods; a "model" should be understood as, say, a Buick Regal four-door sedan with a specified range of accessories, options, and appointments. Any deviation from the detailed specification is a "new model." For nondurables and ser-

vices, "variety" seems a more natural term than "model"; if a box of Kellogg's Corn Flakes contains 17 ounces where it formerly held 18, it is a different "variety." A similar definition applies to services, where, however, it is sometimes difficult to observe the specification of what is being sold. For purposes of this article, a "good" (or a "service") is defined as a set of models or varieties that "fit" a common hedonic function. "Characteristics" are defined in Triplett (1983).

2. The earliest serious BLS hedonic study appears to be the unpublished work of Gavett (1967). Triplett (1971) was commissioned by the BLS. Gillingham (1975) was a dissertation funded by BLS. Other examples of hedonic research in the 1970s may be found in the list of BLS working papers.

3. Also, Dryden, Reut, and Slater (c. 1987) report that a "regression model" (presumably, a hedonic model was meant) on CPI data for a number of products was used to estimate international price comparisons.

4. However, Cahill (1988) cites a number of regression-based quality adjustments in Canadian price indexes that could be termed hedonic.

5. So far as I know, this "no-theory" position appeared exclusively in oral and workshop presentations and not in any published place, but by the early 1970s such challenges were commonly encountered. The earliest published works on the theory of hedonic indexes (Muellbauer 1974; Lucas 1975) tended to support the no-theory charge because they concluded—incorrectly, I believe—that existing empirical work was inconsistent with theoretical requirements. Later work (see below) has now clarified these issues.

6. The following paragraphs are condensed from the "Summary of the Theory" section in Triplett (1989). A more formal treatment is Triplett (1987). The major basic sources are, as noted, Rosen (1974) and Triplett (1983).

7. Functional form is thus appropriately determined with normal econometric procedures. I have argued elsewhere (Triplett 1989), however, that for theoretical reasons researchers should explore a richer range of functional forms than the limited number that have appeared in most empirical hedonic work to date, and that in some circumstances the functional form question demands being treated as a frontier estimation problem.

8. Thus, estimating an exact characteristics price index requires estimating the hedonic function *and* the utility or production function (properly, the indirect utility function or the production cost function, or else demand functions derived from those functions). The formidable econometric problems in such a task are explored in Epple (1987).

9. An explosion of research using hedonic methods occurred in labor economics ("hedonic wage studies") and in urban economics (the extensive literature on valuing air quality and urban amenities), much of which acknowledged the guidance of Rosen (1974) and, to an extent, the example of Thaler and Rosen (1976).

10. That is, the census housing measure is an index of the hedonic prices of square feet, bathrooms (and so forth), where weights are the average number of square feet, average number of bathrooms (and so forth), in houses constructed in the base period (Laspeyres form) or the current period (Paasche form). Another way of stating it is that the index measures (in its Laspeyres form) the change in the price of a house with (mean) base-period characteristics.

11. The study also compared dummy variable and characteristics price indexes computed from the same data. See Cole et al. (1986) and Dulberger (1989) and also the review of calculation methods for hedonic indexes in Triplett (1989).

12. At one time, e.g., comparison of claims by different manufacturers for the same change or sequence of changes required by federal air-pollution and safety legislation led to a "low ball" rule: the maximum adjustment allowed on any manufacturer's cars was the lowest cost reported by any manufacturer. As federal requirements became

more stringent, however, a wider range of engineering alternatives for meeting the standards developed, so the changes introduced by one company might bear little relation to what was done by another, and the rule was abandoned.

13. For example, the 7 October 1968 BLS press release that provided the data for 1969 model cars used for table 7.1 states: "Quality in an automobile is measured in terms of safety, reliability, performance, durability, economy of operation, carrying capacity, maneuverability, comfort and convenience." Similarly, the internal BLS "guidelines" for adjusting automobiles for quality change (version of 25 July 1980) states: "The basic concept of quality in an automobile or truck is the utility to the user. It is usually thought of in terms of reliability, durability, convenience, safety, economy, speed, acceleration, carrying capacity, maneuverability, comfort, appearance, prestige, etc." This matter is made somewhat cloudy by the fact that the BLS published documentation for its price indexes sometimes used language that was inconsistent with its "quality guidelines" (which were internal working documents).

14. For reasons that are not entirely clear at this writing, exploratory Census Bureau hedonic functions on multifamily dwellings have not produced usable results (Pollock 1987).

15. One of the criteria listed in the article is that "the indexes should measure construction with fixed specifications." Hedonic measures are listed as a variant of the fixed-specification method, a position I believe is correct: the variables in the hedonic function serve as the specification for what is being priced. The (Laspeyres version) Census house price index can be interpreted as the price through time of a house having the mean specification of those built in the base year.

16. It should be noted (see also n. 13 above) that examination of agency documentation can sometimes produce confusing or contradictory passages on the treatment of quality change, and it would be easy to quote selectively from agency documents to challenge the interpretation I have set forth above. For anyone who wishes to pursue this matter for intellectual reasons, I would note that the whole resource-cost, user-value debate was a confused one. Neither the theory of quality change for different economic measurements, nor the theory of hedonic functions and hedonic indexes, was well worked out. Hindsight, from the vantage of a better understanding of the theory, shows error as well as insight on both sides of the debate. See also n. 17 below.

17. An example of the difference of positions on this issue is contained in the report of the Panel to Review Productivity Statistics (1979). Because the report's chapters were authored by different panel members, who ascribed to one or the other of the positions outlined in the text, its chapter on measuring output endorses a user-value quality concept (incorrect for this case), while its chapter on measuring inputs endorses resource-cost for capital (incorrect) and user-value (marginal productivity) for labor inputs (correct).

18. The "composite" estimate in table 7.2 was introduced as a deflator for computer processors in the NIPAs in December 1985. In the composite, price relatives for matched models are used whenever they are available, but when a price is available in one period but "missing" in the other, the missing price is imputed from a computer hedonic function. Indexes for three other items of computer equipment, also calculated by the composite-imputation method, are included in the "computers" category of Producers' Durable Equipment. See Cole et al. (1986) and Cartwright (1986).

19. Fisher, McGowan, and Greenwood remark that "the computer market . . . has never been close to long-run equilibrium in its entire existence" (1983, 149).

20. Some 30 studies are reviewed in Triplett (1989), which presents price indexes for computer processors and peripheral equipment that go back into the 1950s. The computer processor price index, which was compiled by combining "best practice" research studies, declined from 1,320 in 1953 to 14.8 in 1972 (1965 = 100), a decline comparable with that of the post-1972 index in table 7.4.

21. The technical change studies in the computer science literature were readily converted into price indexes since the technological investigators measured performance per dollar, and economists want price indexes to measure the inverse—cost per unit of performance.

22. Rapidly falling computer price indexes also create large changes in relative prices; substantial substitution toward relatively cheaper computer equipment can be observed in investment aggregates such as office machinery, or Producers' Durable Equipment. When these aggregates are combined with fixed-weight index numbers (of price or quantity), the "substitution bias" associated with such index number formulas can be substantial. It appears that some users who have expressed dissatisfaction with the behavior of the computer measures are really upset about the behavior of fixed-weight aggregates that employ the computer data. I am exploring the fixed-weight bias issue in investment categories of GNP, in work to appear elsewhere (using the Time-series Generalized Fisher Ideal, or TGFI, index number that first appeared in the final section of Triplett 1989).

References

Blackorby, Charles, and R. Robert Russell. 1978. Indices and Subindices of the Cost of Living and the Standard of Living. *International Economic Review* 19 (February): 229–40.

Cahill, Leonard. 1988. Quality Change Adjustment Manual. Ottawa: Statistics Canada, Prices Division.

Cartwright, David W. 1986. Improved Deflation of Purchases of Computers. *Survey of Current Business* 66 (March): 7–9.

Christensen, Laurits R., and Marilyn E. Manser. 1976. Cost-of-Living Indexes and Price Indexes for U.S. Meat and Produce. In *Household Production and Consumption*, ed. Nestor E. Terleckyj, 399–446. NBER Studies in Income and Wealth, vol. 40. New York: National Bureau of Economic Research.

Cole, Rosanne, Y. C. Chen, Joan A. Barquin-Stolleman, Ellen Dulberger, Nurhan Helvacian, and James H. Hodge. 1986. Quality-adjusted Price Indexes for Computer Processors and Selected Peripheral Equipment. *Survey of Current Business* 66 (January): 41–50.

Court, Andrew T. 1939. Hedonic Price Indexes with Automotive Examples. In *The Dynamics of Automobile Demand*, 99–117. New York: General Motors.

Denison, Edward F. 1957. Theoretical Aspects of Quality Change, Capital Consumption, and Net Capital Formation. In *Problems of Capital Formation: Concepts, Measurement and Controlling Factors*. NBER Studies in Income and Wealth, vol. 19. Princeton, N.J.: Princeton University Press.

———. 1989. *Estimates of Productivity Change by Industry: An Evaluation and An Alternative*. Washington, D.C.: Brookings.

Dhrymes, Phoebus J. 1971. Price and Quality Changes in Consumer Capital Goods: An Empirical Study. In *Price Indexes and Quality Change: Studies in New Methods of Measurement*, ed. Zvi Griliches. Cambridge, Mass.: Harvard University Press.

Dryden, John, Katrina Reut, and Barbara Slater. c. 1987. Bilateral Comparison of Purchasing Power Parity between the United States and Canada. Ottawa: Statistics Canada.

Dulberger, Ellen. 1988. Price Index for MOS Memories. Presentation at the National Bureau of Economic Research Summer Workshop, Cambridge, Mass., July.

———. 1989. "The Application of an Hedonic Model to a Quality Adjusted Price Index for Computer Processors." In *Technology and Capital Formation*, ed. Dale W. Jorgenson and Ralph Landau. Cambridge, Mass.: MIT Press.

Early, John F., and James H. Sinclair. 1983. Quality Adjustment in the Producer Price Indexes. In *The U.S. National Income and Product Accounts: Selected Topics*, ed. Murray F. Foss, 107–45. NBER Studies in Income and Wealth, vol. 47. Chicago: University of Chicago Press.

Epple, Dennis. 1987. Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products. *Journal of Political Economy* 95 (February): 59–80

Fisher, Franklin M., John J. McGowan, and Joen E. Greenwood. 1983. *Folded, Spindled, and Mutilated: Economic Analysis and U.S. v. IBM*. Cambridge, Mass.: MIT Press.

Fisher, Franklin M., and Karl Shell. 1972. *The Economic Theory of Price Indices: Two Essays on the Effects of Taste, Quality, and Technological Change*. New York: Academic Press. (Essay 1 also published in Zvi Griliches, ed., [1971], below.)

Flamm, Kenneth. 1987. *Targeting the Computer.* Washington, D.C.: Brookings.

Gavett, Thomas W. 1967. Research on Quality Adjustments in Price Indexes, Part III: Experiments in Multivariate Analysis of Quality Change. Bureau of Labor Statistics memorandum (Office of Prices and Living Conditions). Washington, D.C.

Gilbert, Milton. 1961. Quality Changes and Index Numbers. *Economic Development and Cultural Change* 9 (April): 287–94.

Gillingham, Robert F. 1975. Place to Place Rent Comparisons. *Annals of Economic and Social Measurement* 4 (Winter): 153–73.

Griliches, Zvi. 1961. Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change. In *The Price Statistics of the Federal Government*, 173–96. (See Price Statistics Review Committee [1961] below.)

———. 1964. Notes on the Measurement of Price and Quality Changes. In *Models of Income Determination*, 381–418. NBER Studies in Income and Wealth, vol. 28. Princeton, N.J.: Princeton University Press.

———, ed. 1971. *Price Indexes and Quality Change: Studies in New Methods of Measurement.* Cambridge, Mass.: Harvard University Press.

Jaszi, George. 1964. Comment. In *Models of Income Determination*, 404–9. NBER Studies in Income and Wealth, vol. 28. Princeton, N.J.: Princeton University Press.

Koopmans, Tjalling. 1947. Measurement without Theory. *Review of Economics and Statistics* 29 (August): 161–72.

Lucas, R. E. B. 1975. Hedonic Price Functions. *Economic Inquiry* 13 (June): 157–78.

Muellbauer, John. 1974. Household Production Theory, Quality, and the "Hedonic Technique." *American Economic Review* 64 (December): 977–94.

Musgrave, John. 1969. The Measurement of Price Changes in Construction. *Journal of the American Statistical Association* 64 (September): 771–86.

Panel to Review Productivity Statistics. 1979. *Measurement and Interpretation of Productivity.* Washington, D.C.: National Academy of Sciences.

Pollak, Robert A. 1975. Subindexes in the Cost of Living Index. *International Economic Review* 16 (February): 135–50.

———. 1983. The Theory of the Cost-of-Living Index. In *Price Level Measurement: Proceedings from a Conference Sponsored by Statistics Canada*, ed. W. E. Diewert and C. Montmarquette, 87–161. Ottawa: Minister of Supply and Services.

Pollock, Jesse. 1987. Research into a Cost Index for Multiunit Residential Construction. Presentation to the American Statistical Association Census Advisory Committee, October.

Price Statistics Review Committee. 1961. *The Price Statistics of the Federal Govern-*

ment. U.S. Congress, Joint Economic Committee. *Government Price Statistics, Hearings*, pt. 1. 87th Cong., 1st sess. (Also published as National Bureau of Economic Research, General series no. 73.)
Rosen, Sherwin. 1974. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy* 82 (January–February): 34–55.
Thaler, Richard, and Sherwin Rosen. 1976. The Value of Saving a Life: Evidence from the Labor Market. In *Household Production and Consumption*, ed. Nestor E. Terleckyj, 265–97. NBER Studies in Income and Wealth, vol. 40. New York: National Bureau of Economic Research.
Triplett, Jack E. 1969. Automobiles and Hedonic Quality Measurement. *Journal of Political Economy* 77 (May–June): 408–17.
———. 1971. The Theory of Hedonic Quality Measurement and Its Use in Price Indexes. BLS Staff Paper no. 6. Washington, D.C.: U.S. Bureau of Labor Statistics.
———. 1983. Concepts of Quality in Input and Output Price Measures: A Resolution of the User Value–Resource Cost Debate. In *The U.S. National Income and Product Accounts: Selected Topics*, ed. Murray F. Foss, 269–311. NBER Studies in Income and Wealth, vol. 47. Chicago: University of Chicago Press.
———. 1987. Hedonic Functions and Hedonic Indexes. In *The New Palgrave: A Dictionary of Economics*, ed. John Eatwell, Murray Milgate, and Peter Newman, 2:630–34. London: Macmillan.
———. 1989. Price and Technological Change in a Capital Good: A Survey of Research on Computers. In *Technology and Capital Formation*, ed. Dale W. Jorgenson and Ralph Landau. Cambridge, Mass.: MIT Press.
Triplett, Jack E., and Richard J. McDonald. 1977. Assessing the Quality Error in Output Measures: The Case of Refrigerators. *Review of Income and Wealth* 23 (June): 137–56.
U.S. Department of Commerce, Bureau of Economic Analysis. 1987. GNP: An Overview of Source Data and Estimating Methods. Methodology Paper Series MP-4. Washington, D.C.: U.S. Government Printing Office, September.
U.S. Department of Labor, Bureau of Labor Statistics. 1988. *CPI Detailed Report*. January.

Comment W. Erwin Diewert

Triplett has a nice discussion of recent methods for adjusting for quality change that have been used by the Bureau of Labor Statistics. In keeping with the historical nature of this conference, I shall briefly review the ancient literature on methods for quality adjustment.

Some of the early researchers on price measurement were aware of the problem of quality change, but the pace and direction of the change did not seem large enough to warrant an explicit treatment.[1]

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

The author is indebted to R. C. Allen, B. M. Balk, and J. E. Triplett for valuable historical references and discussions over the years. This research was supported by the SSHRC of Canada and by the Summer Institute of the National Bureau of Economic Research Program on Productivity. Neither institution is to be held responsible for any opinions expressed by the author.

However, by the latter part of the nineteenth century, Sidgwick (1883) realized that not only were improvements in the quality of goods leading to a bias in price comparisons, but also the growth of international and interregional trade (due primarily to transportation improvements) led to the systematic introduction of "entirely new kinds of things" and this too led to a bias in price comparisons. As the following quotation indicates, Sidgwick thought that utility theory would play a role in eliminating these biases: "Here again there seems to be no means of attaining more than a rough and approximate solution of the problem proposed; and to reach even this we have to abandon the prima facie exact method of comparing prices, and to substitute the essentially looser procedure of comparing amounts of utility or satisfaction" (1883, 68). Unfortunately, the mathematical apparatus of consumer theory was not sufficiently developed at that time to enable Sidgwick to make any specific progress on the new-good problem.

In a brilliant paper, Marshall (1887) not only proposed the tabular standard, the chain system, and the Edgeworth-Marshall index number formula, he also made the first real progress on the appropriate treatment of new goods, as the following quotation indicates:

> This brings us to consider the great problem of how to modify our unit so as to allow for the invention of new commodities. The difficulty is insuperable, if we compare two distant periods without access to the detailed statistics of intermediate times, but it can be got over fairly well by systematic statistics. A new commodity almost always appears at first at something like a scarcity price, and its gradual fall in price can be made to enter year by year into readjustments of the unit of purchasing power, and to represent fairly well the increased power of satisfying our wants which we derive from the new commodity. (1887, 373)

As the above quotation indicates, Marshall was well aware of the product cycle and he felt that the early introduction of new commodities into the consumer price index in the context of the chain system would capture *most* of the benefits due to the introduction of new commodities. As we shall see later, not quite *all* of the benefits are captured using Marshall's suggested method, since his method incorrectly ignores the new good in the first period that it makes its appearance.

Marshall (1887, 373–74) also realized that improvements in transportation led to the general availability of location-specific goods, such as fish at the seaside or strawberries at a farm. Marshall correctly felt that these "old" goods that suddenly became available at many locations should be regarded as "new" goods and treated in the same way as a genuinely new good. His words on this important observation are worth quoting:

> This class of consideration is of much more importance than at first sight appears; for a great part of modern agriculture and transport industries are devoted to increasing the periods of time during which different kinds of food are available. Neglect of this has, in my opinion, vitiated the statistics

> of the purchasing power of many in medieval times with regard to nearly all kinds of foods except corn; even the well-to-do would hardly get so simple a thing as fresh meat in winter. (374).

Marshall's suggested treatment of the new-good problem was acknowledged and adopted by many authors including Irving Fisher (1911, 204) and Pigou (1912, 47). Divisia (1926, 45), working from his independent perspective, also suggested the use of the chain method as a means of dealing with the new-good problem.

The next important contributor to the discussion of new goods in price measurement was Keynes (1930, 94), who described in some detail one of the most common methods for dealing with the new-good problem: simply ignore any new or disappearing goods in the two time periods under consideration and calculate the price index on the basis of the goods that are common to the two situations. The corresponding quantity index was to be obtained residually by deflating the relevant value ratio by this narrowly based price index. Keynes called this method the *highest common factor method.* This method would be identical to Marshall's chain method if the two time periods were chosen to be adjacent ones. However Keynes (1930, 105–6) advocated his method in the context of a fixed-base system of index numbers, and he specifically rejected the chain method for three reasons: (1) each time a new product is introduced, a chain index does not take into account the benefits of the expanded choice set, and thus over long periods of time, the chain price index will be biased upward and the corresponding quantity index will be biased downward; (2) the chain index fails Walsh's multiperiod identity test (see Diewert 1988, eq. [13]), and (3) the chain method was statistically laborious.

Keynes's last objection to the chain method is no longer relevant in this age of computers. Moreover, Keynes was unable to offer any positive alternative to the chain method for comparing situations separated by long periods of time, as the following quotation indicates: "We cannot hope to find a ratio of equivalent substitution for gladiators against cinemas, or for the conveniences of being able to buy motor cars against the conveniences of being able to buy slaves" (Keynes 1930, 96).

However, Keynes's first objection to the chain method (which was later echoed by Pigou [1932, 72]) was certainly valid (as was his second objection).[2] A satisfactory theoretical solution to Keynes's first objection did not occur until Hicks adapted the analytical apparatus of consumer theory to the problem.

When new consumer goods make their appearance for the first time, say in period 2, their prices and quantities can be observed. In period 1, the quantities of the new goods are all obviously zero, but what are the corresponding prices? Hicks (1940, 114) provided a theoretical solution:

> They are those prices which, in the one situation, would *just* make the demands for these commodities (from the whole community) equal to zero. These prices cannot be estimated, but we can observe that between the two

> situations the demands for these commodities will have increased from zero to certain positive quantities; and hence it is reasonable to suppose that the "prices" of these commodities will usually have fallen relatively to other prices. This principle is sufficient to give us a fairly good way of dealing with the case of new goods.

Of course, in the context of the producer price index, the appropriate period-1 shadow prices for the new goods are those prices that just induce each period-2 producer of the new goods to produce zero quantities in period 1.

Hicks's basic idea was used extensively by Hofsten (1952, 95–97) who dealt not only with new goods but also adapted the Hicksian methodology to deal with disappearing goods as well. Hofsten (1952, 47–50) also presents a nice discussion of various methods that have been used to adjust for quality change, similar to Triplett's (in this volume) discussion of quality-change measurement techniques.

Franklin Fisher and Karl Shell (1972, 22–26) laid out the formal algebra for constructing the first period Hicksian "demand reservation prices" defined in the above quotation by Hicks. Diewert (1980, 498, 501) used the Hicksian framework to examine the bias in the Fisher price index P_F (defined by using vector notation):

$$(1) \qquad P_F(p^1,p^2,q^1,q^2) \equiv [p^2 \cdot q^1 \; p^2 \cdot q^2 / p^1 \cdot q^1 \; p^1 \cdot q^2]^{1/2}.$$

Diewert calculated P_F when the reservation prices were incorrectly set equal to zero and compared this index to the Fisher price index that simply ignored the existence of the new goods in the two periods under consideration (which is Marshall's method).[3] Diewert (1980, 501–3) also made some suggestions for estimating the appropriate Hicksian reservation prices in an econometric framework.

Is the new-good bias large or small? One can only answer this question in the context of the price measurement procedures used by individual statistical agencies. In Diewert (1987, 779), some simple hypothetical examples were given that showed that traditional fixed-base procedures could generate much higher measures of price increase than would be generated using the chain method.[4] However, what is needed is empirical evidence.

Numerical computation of alternative methods based on detailed firm data on individual prices and quantities where new goods are carefully distinguished would cast light on the size of the new-good bias. Another line of empirical work that would be of interest would be to collect industry price and quantity data on various major new goods (e.g., microwave ovens, video recorders, home computers, satellite dishes, etc.) and then attempt to rework the relevant price indexes in the light of this extra data.

Notes

1. Thus Lowe (1823, app. 87) states: "In regard to the quality of our manufactures, we must speak with more hesitation, and can hardly decide whether the balance be in

favour of the present or of a former age; for if our fabrics are now much more neat and convenient, they are in a considerable degree less durable."

2. Pigou (1932, 71) also has a nice criticism of Keynes's highest common factor method, which was later repeated by Hofsten (1952, 59). Pigou also criticized Fisher's (1922, 308–12) later preference for the fixed-base method.

3. The second index has a smaller bias than the first index.

4. Since 1978, the U.S. Bureau of Labor Statistics has used a probability sampling approach in the consumer price index that probably reduces some of this fixed-weight bias, but the bias is not eliminated.

References

Diewert, W. E. 1980. Aggregation Problems in the Measurement of Capital. In *The Measurement of Capital,* ed. Dan Usher, 433–528. NBER Studies in Income and Wealth, vol. 45. Chicago: University of Chicago Press.

———. 1987. Index Numbers. In *The New Palgrave: A Dictionary of Economics,* ed. J. Eatwell, M. Milgate and P. Newman 2:767–80. London: Macmillan.

———. 1988. The Early History of Price Index Research. NBER Discussion Paper no. 2713. Cambridge, Mass.: National Bureau of Economic Research.

Divisia, F. 1926. *L'indice monétaire et la theorie de la monnaie.* Paris: Société anonyme du Recueil Sirey.

Fisher, I. 1911. *The Purchasing Power of Money.* London: Macmillan.

———. 1922. *The Making of Index Numbers.* Boston: Houghton Mifflin.

Fisher, F. M., and K. Shell. 1972. *The Economic Theory of Price Indices.* New York: Academic Press.

Hicks, J. R. 1940. The Valuation of the Social Income. *Econometrica* 7:108–24.

Hofsten, E. von. 1952. *Price Indexes and Quality Change.* London: Allen & Unwin.

Keynes, J. M. 1930. *Treatise on Money,* vol. 1. London: Macmillan.

Lowe, J. 1823. *The Present State of England in Regard to Agriculture, Trade and Finance,* 2d ed. London: Longman, Hurst, Rees, Orme & Brown.

Marshall, A. 1887. Remedies for Fluctuations of General Prices. *Contemporary Review* 51:355–75.

Pigou, A. C. 1912. *Wealth and Welfare.* London: Macmillan.

———. 1932. *The Economics of Welfare,* 4th ed. London: Macmillan.

Sidgwick, H. 1883. *The Principles of Political Economy.* London: Macmillan.

This Page Intentionally Left Blank