Volume Title: The Measurement of Durable Goods Prices

Volume Author/Editor: Robert J. Gordon

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-30455-8

Volume URL: http://www.nber.org/books/gord90-1

Conference Date: n/a

Publication Date: January 1990

Chapter Title: The Methodology of Quality Adjustment

Chapter Author: Robert J. Gordon

Chapter URL: http://www.nber.org/chapters/c8311

Chapter pages in book: (p. 79 - 110)

3 The Methodology of Quality Adjustment

3.1 Introduction

A central theme of chapter 2 was the distinction between the criterion used in choosing quality characteristics for input cost indexes, and the estimator used to adjust durable goods price indexes for differences in characteristics across models. The theoretical discussion identified "user value" as the quality criterion and seller's marginal cost as the quality adjustment estimator. This chapter is concerned with the practical problems involved in computing quality adjustments. It concentrates on general measurement issues applicable to a variety of products; special problems particular to individual products are discussed in the subsequent product chapters.

The primary aim of this book is to use alternative data sources to produce new price deflators for selected types of durable goods. These are compared with existing deflators compiled by BEA, and the implications for BEA measures of output, productivity, and capital stock are then discussed. The great majority of BEA deflators for individual types of consumer durables are based on underlying commodity price components of the CPI, while deflators for individual types of producer durables are based on components of the PPI. Since the BEA is thus largely a "conduit" for the CPI and PPI price data collected by the BLS, this chapter refers mainly to the techniques of the latter agency.

The chapter begins with a description of present quality adjustment techniques used by the BLS. For exposition purposes, the BLS techniques are collectively called the *conventional method.* The advantages and disadvantages of these techniques are then compared with those of available alternatives, the best known of which is the hedonic regression approach. At the outset, it should be emphasized that the choice between estimation techniques does not correspond to the debate reviewed in chapter 2 regarding quality-adjustment criteria. Although some descriptions of government methods(e.g., Jaszi 1971) have ruled out adjustments for the types of quality change called *nonproportional* in chapter 2, in fact the use by the BLS of manufacturers' cost data for making quality adjustments corresponds closely to the general procedure advocated in chapter 2. In that conceptual discussion, the choice of quality characteristics is based on user value, and the estimator used for adjustment is the marginal cost of producing those quality characteristics along a fixed cost schedule. Correspondingly, "the criterion for measuring quality change in a BLS input price index is always user value. . . . The estimator, on the other hand, is sometimes manufacturing cost because that may be the only information available" (Triplett 1983b, 255). Thus, any criticism of BLS methodology, and any suggestion that alternative techniques should be substituted in some cases, involves not a theoretical dispute but rather concentrates on sins of omission and on detailed procedures. Each of the several available techniques has weaknesses, and it may be preferable in some cases to use techniques in combination (partly to serve as cross-checks on each other) rather than to adhere to the mechanical application of a single technique.

This chapter is not a comprehensive review of either the CPI or the PPI (formerly the WPI). For these, the reader is referred to the report of the NBER Price Statistics Review Committee (NBER 1961), or Stigler report, and to the more recent and comprehensive *Wholesale Price Index: Review and Evaluation* (U.S. Executive Office of the President 1977)), or the Ruggles report. These reports take up numerous issues that are not treated here, including sampling methodology, revision of weights, transaction versus list prices, and order versus shipment prices. As will soon become evident, the issues involved in quality adjustment are complex enough to warrant limiting our attention to that single issue.

3.2 Official Price Indexes: Coverage and Procedures

Primary emphasis in this section is on durable goods price information collected in the PPI, the basic source for the national income accounts deflators for PDE. Supplementary information on the CPI and the consumer durable deflator is contained at the end of the next section. The PDE deflator is based (with a few exceptions) on PPI eight-digit "item indexes," weighted by the estimated share of each commodity in gross domestic private investment. The BEA publishes twenty-two separate group deflators for major types of equipment, for example, office machinery and communications equipment. Current-dollar expenditures in the twenty-two groups are deflated separately to yield constant-dollar group expenditure, constant-dollar aggregate PDE expenditure, and the aggregate PDE deflator. Within the twenty-two groups, the individual PPI commodity indexes are combined with unpublished weights based on expenditures in about 150 detailed components of PDE (a list is provided in Appendix A).

Before 1947, motor vehicles, agricultural machinery, and office furniture were the only capital goods included in the PPI; data are available beginning in 1947 for a large number of different types of machinery, with the gradual introduction of new products and replacement of obsolete items throughout the postwar period. One measure of the extent of coverage is the percentage of a four-digit SIC industry's shipments represented by the value of individual seven-digit products that are priced directly. Based on the 1958 Census of Manufacturers, coverage in the PPI in the early 1960s was quite spotty in the machinery and transportation equipment industries (SIC 35, 36, and 37), with less than 10 percent of products priced directly in four-digit industries that contribute 42 percent of total value added, and with coverage of 50 percent or greater in industries contributing only 27 percent of value added. Among the important industries with very low or no coverage were construction machinery, special dies and tools, machine shop equipment, telephone and radio-television communication equipment, aircraft and aircraft engines, and ship building (Searle 1964, 358–59).

By the early 1970s, the situation had improved somewhat as coverage was broadened. Table 3.1 shows computations for six major two-digit durable goods industries based on the value of shipments recorded in the 1972 Census of Production. The first row indicates the number and value of shipments of all five-digit SIC product classes when classified into two-digit SIC major industries. Then row 2 shows the percentage of the five-digit SIC product classes in each two-digit major industry that were directly priced by the BLS by number and value of shipments. By this measure, the extent of coverage seems quite wide, ranging from 44 to 72 percent by value.

Another more meaningful criterion is the percentage of the seven-digit SIC product classes that were "directly priced" within each five-digit SIC product class. This is the criterion used by the BLS for publishing price indexes for the five-digit classes. Here, the extent of coverage is less impressive, as shown on row 3, with the percentages by value ranging from just eight in SIC 38 to fifty-five in SIC 25. In the machinery and transportation equipment industries (SIC 35, 36, and 37), the average percentage had risen from the 27 percent figure quoted above for 1958 to an average of 36 percent for 1972.[1] Thus, one justification for the exploration of alternative sources of price data, as in this book, is to fill in for coverage gaps in the PPI. Since coverage has improved over time, it is important to create alternative indexes over as long a historical time path as possible, in order to determine whether any bias in one direction or another has been created by changing coverage.

The last section of table 3.1 identifies manufacturers' reports as the major source of price information for the PPI. In all the durable goods industries shown, company reports account for between 97 and 100 percent of the total

1. The criterion changed from percentage of four-digit industries in the earlier comparison to percentage of five-digit industries in the later comparison.

Table 3.1 Data on PPI Coverage and Methods, Selected Two-Digit Industry Groups

	Two-Digit Industry and SIC Code					
	Furniture and Wood Products 25 (1)	Fabricated Metal Products 34 (2)	Nonelectrical Machinery 35 (3)	Electrical Equipment 36 (4)	Transportation Equipment 37 (5)	Instruments and Related 38 (6)
1. Total five-digit SIC product classes:[a]						
Number	43	127	190	117	71	48
Value ($billions)	11	51	64	54	103	16
2. Percent directly priced in PPI:[a]						
Percent by number	37	54	60	62	18	29
Percent by value	60	72	62	60	68	44
3. Percent published by BLS:[a]						
Percent by number	30	27	36	33	7	6
Percent by value	55	45	33	36	38	8
4. Sources of price reports:[b]						
Total number of reports	156	672	1771	992	191	140
Percent from trade sources	0	2	1	1	0	1
Percent from government agencies	2	1	1	3	0	0
Percent from companies	98	97	98	97	100	99
Value of shipments directly priced as percent of total	26	16	19	24	22	7

Source: U.S. Executive Office of the President (1977, tables I-2 and I-3).

[a]Computed from the *WPI Weight Diagram for 1972*.

[b]As of March 1975.

reports. Virtually none of the reports come from equipment purchasers, virtually all from equipment sellers. In a sample survey carried out in 1975, it appeared that about two-thirds of the durable goods reports were submitted by company headquarters, and the other one-third by individual establishments.

The essence of the "conventional method" used by the BLS is called *specification pricing.* If a product is specified by all quality characteristics (x) that influence user value, then none of the increase in sales revenue per unit due to increases in x is erroneously considered as a price change. The focus in the PPI has always been on obtaining prices for a given commodity from several producers. Thus, the specifications for the eight-digit item indexes must be detailed enough to eliminate the influence of important types of quality change, while at the same time broad enough to admit price reports from several manufacturers. The level of detail in specifications varies across commodities. For some relatively homogeneous commodities where manufacturers are basically producing identical items, it is possible to make the specification so detailed that it is quite unlikely that quality change slips through unnoticed. An example is the following specification for an eight-digit item index for wire rods:

10–13–01–11.02 Wire rods: 7/32" diameter, coils, hot rolled carbon steel, CI 008 industrial or standard quality, base quantity 20 net tons and over; mill to user, f.o.b. mill, per 100 lbs.

For other commodities where technical innovation is frequent and features vary across manufacturers, commodity specifications are less detailed. Consider this specification for an eight-digit item index for color television sets:

12–52–01–56 Color television receiver, console model, 21", 23", or 25" picture tube, veneer cabinet; manufacturer to dealer or distributor, f.o.b. factory, or warehouse, each.

Since the only specification of the quantity of x here is the size of the picture tube, many other features of color television sets could be ignored and some part of a change in price associated with a new or altered model might be treated as price change instead of quality change. Among the omitted features are remote control capability (and even this varies in the number of functions that can be controlled remotely), capability for cable television, headphone equipment, automatic tuning, type of built-in antenna, and so on.

The NBER review of the PPI carried out a test to determine whether the eight-digit PPI item indexes were defined for homogeneous products by computing the coefficient of variation of individual price reports relative to the average price within each PPI commodity. The coefficients of variation in percentages for the six durable goods categories were 41.3 (furniture), 16.1

(fabricated metal), 24.6 (nonelectrical machinery), 18.7 (electrical equipment), 16.0 (transportation equipment), and 21.6 (instruments). The review also pointed out that these differences persisted over a year between 1975 and 1976, suggesting substantial heterogeneity within eight-digit item groups (U.S. Executive Office of the President 1977, I-24, I-25). This intraitem heterogeneity does not necessarily cause a problem for the specification method, for if the manufacturer alters his product in any significant way, the change is supposed to be reported to the BLS. The BLS then decides whether the change warrants an explicit quantity adjustment or should be ignored. Its standards are sufficiently tight that broad specifications are not allowed to permit dimensional changes to be treated as price changes. For instance, Triplett (1971b, 184) reports that, despite a range in the refrigerator specification of 13.5–16.5 cubic feet, anything over a 0.4-cubic-foot change calls for an explicit quality adjustment. Thus, three different varieties of color television sets sold by three different manufacturers could be included in item index 12–52–01–56, with the top grade costing 50 percent more than the lowest grade, but there would be no problem for the index if the price reports from each of the three sellers were effectively adjusted for quality change.

3.3 Methodology of Quality Adjustment in the BLS Indexes

Since heterogeneity within PPI item indexes is common, the crucial issue becomes the adequacy of the quality adjustments carried out on the individual price reports within the eight-digit commodity classifications. To the extent that quality improvements occur within the eight-digit classes for which no quality adjustments are made, the overall PPI will rise faster than the "true" change in price. When informed by a manufacturer of an alteration on a model, the BLS commodity specialists make a choice among four alternative procedures.

This section first describes the mechanical method of adjustment under each of the four methods, using the formulas suggested by Early and Sinclair (1983, 110). It then turns to the frequency of use for each method, and some of the problems and potential biases that they might introduce.

If there is no model change brought to the attention of the BLS, then the price change between two successive months is simply P_t/P_{t-1}. In each of the four methods that the BLS uses when informed of a model change, we can express the calculated price change as P_t/L_{t-1}, where L_{t-1} is the "link price" of the new model. Each method calculates L_{t-1} by a different procedure. Under "direct comparison," the model change is ignored, so that $L_{t-1} = P_{t-1}$. Under "deletion," the price change between the two months for this particular product is omitted from the index calculation, which has the effect of setting $L_{t-1} = P_t/(P^*_t/P^*_{t-1})$ where the term in parentheses is the average price change recorded for the other price reports in the eight-digit item index. Under the "link, no change" method, then, by definition there is no price change, and so $L_{t-1} = P_t$. Finally, with a producer cost adjustment amounting to Q_t (in dollars), $L_{t-1} = P_{t-1} + Q_t$.[2]

3.3.1 Direct Comparison.

The simplest procedure is simply to neglect the change. This occurs when the alteration is considered to be "small," for example, less than a 0.4-cubic-foot change in refrigerator capacity, or less than 2 inches of length variation for a rug. Within these limits, whatever price change is observed to accompany the model alteration will be treated as price change, not as quality change. As an example of the frequency of direct comparison, in the CPI in 1965, 23.1 percent of all price comparisons of appliances involved substitutions caused by nonidentical models, and 59.2 percent of these substitutions were handled by the method of direct comparison (Triplett 1971b, 187).[3]

The use of the direct comparison method does not necessarily impart a positive bias to the rate of price change recorded by the PPI. Some model alterations may represent a decline in performance quality or durability. Also, Triplett (1971b) has pointed out that a quality characteristic included in a BLS specification, for example, a frost-free freezer or a refrigerator, may "trickle down" over the years to models that have smaller quantities of other quality characteristics that are not included in the BLS specification. If this phenomenon were frequent, it could impart a negative bias to the PPI.

3.3.2 Deletion

In the example cited above for the 1965 CPI appliance indexes, 40.8 percent of the substitutions were not handled by direct comparisons. In most of this remaining group, the price quotation of the new model was simply deleted (39.4 percent), with a small remainder (1.4 percent) related to the previous observation by linking based on information for overlapping models. The choice between direct comparison and deletion depends on the size of the change in price between the old and the new quotations, with the deletion method chosen if the price change exceeds a predetermined but unpublished limit. Since quality improvements are frequently introduced on models with no change in price, the predetermined limit will not be exceeded in this case, and the new and the old models will be compared directly, causing the true decline in price to be missed and the index to be biased upward. When the limit is exceeded and the price quotation is deleted, the

2. Early and Sinclair point out that, since Q_t is valued in dollars as of time t rather than, as should occur, at time $t - 1$, the value of the adjustment is overstated during periods of inflation. The correct formula, apparently used now but not before 1978, is $L_{t-1} = (P_t P_{t-1})/(P_t - Q_t)$.

3. The data reported refer to all reported price quotations for nonfood items for two large cities.

index is based on one fewer observation, so that the average price change on the other included varieties is implicitly applied to the excluded variety.

If the most frequent cause of deletion is a sizable quality change, then the official price indexes consist disproportionately of commodities of constant quality. In some of the detailed product studies later in the book, we find that the quality adjusted price of new models is often below that of the models they replace. If so, the effect of deletion is to reduce the weight given in the index to the price declines relative to price increases and thus to bias the overall index upward. In principle, the opposite effect is also possible, if the typical introduction of a new model involves a mixture of quality improvement and "true" price change, since deletion will therefore reduce the weight given to true price change in the index. If the mixture tilted toward more true price change in periods of high inflation, then it is conceivable that the direction of bias might be inversely proportional to the inflation rate, with a positive bias in low inflation periods and a negative bias in high inflation periods.

3.3.3 Linking

When a model with a new quality characteristic is introduced, and price data are available for models with both quality characteristics at a given date, the new price is "linked" to the old one with a difference between the prices of new and old models on the transition date used to estimate the change in quality. This method is equivalent in principle to the hedonic regression method discussed below but is much less useful in practice, because old models often disappear completely. In the next month's price report, a replacement model often appears containing different quantities of one or more quality attributes, thus preventing the necessary simultaneous comparison of the prices of the two models. As recently as 1965, only 4.8 percent of all nonfood substitutions in the CPI were handled by linking, as compared to 61.4 percent by direct comparison and the remaining 33.8 percent by deletion (Triplett 1971b, 187).

3.3.4 Explicit Adjustment Based on Cost Estimate

The cost-adjustment method is saved for last. While it corresponds most closely to the conceptual framework suggested in chapter 2, it is carried out less frequently than the other methods. As described by Triplett (1983a, 254–55), the adjustment takes place in two steps. First, employing the user-value criterion, the BLS must decide whether the model alteration is to be considered as a quality change at all. Examples of changes not allowed as quality improvements by the BLS in the past include a change from a conventional to a digital clock face, and a redesigned speedometer that did not register extremely high speeds.

The second step is to estimate the value of the change. This is most straightforward when an item previously offered as optional equipment at a published price is made standard equipment. This gradual upgrading process was a common feature of quality change for automobiles in the 1950s and 1960s. However, the option price may exaggerate the cost of adding the item to all units in a standardized production run. When no such option price is available, the BLS must accept an estimate of added cost from the manufacturer. It is unclear how these estimates are monitored or audited. There may be an incentive for manufacturers to exaggerate the extra cost of added equipment; although neither sales nor profits are directly affected, manufacturers may want to understate the extent of price increases for public relations and labor relations purposes.

Because quality change is a pervasive feature of durable goods, and because the requisite overlapping models needed for linking are often not available, the extent to which the official price indexes (both CPI and PPI) adjust for quality improvements appears to depend crucially on cost estimates by manufacturers. Yet published descriptions of official methodology agree that, during most of the postwar period, the cost estimates have been made for only a few types of goods and for only a portion of the quality improvements that occurred in those goods. Before 1960, the method was applied in the CPI only to automobile accessories added to the standard model that were previously optional at a stated price: "For example in the case of automobiles our practice has been to substitute the new model car for the previous model, assuming no quality change except for those features which affect some easily observed difference in operational characteristics and for which a value can be determined. . . . In such a comparison we would make no allowances for such changes as greater length or more wrap in the windshield, because we have no objective standard by which to determine the relationship between quality and price for such features" (Jaffe 1959, 195). In 1961, "the most important items [in the CPI] for which this procedure is used regularly are automobiles, gas for heating and cooking, hospitalization insurance, and contract rents" (Hoover 1961, 1178). By implication, no cost-based explicit adjustments were performed then for any other important consumer durables, for example, washing machines, refrigerators, or radio-television equipment.

For automobiles, published reports describe a clear break in technique in 1959: "The guidelines followed in making adjustments for quality differences between the 1965 and 1966 models of automobiles are the same as those adopted by the Bureau in 1959 and were made possible by quality and cost data supplied by manufacturers. Prior to 1959, most adjustments for quality differences involved changes in optional equipment. . . . For the WPI, some additional adjustments for product changes had been taken into account in earlier years, but the extent of such adjustments was limited by the unavailability of the required information" (Stotz 1966, 178). Yet no explicit description or example is available that indicates precisely how changes in dimensions, for example, length or horsepower, have been handled since 1959. Furthermore, as late as 1965, prices of other consumer durables

Table 3.2 **Classification of Quality Adjustments in Selected Major PPI Commodity Groups**

	Major Group and Two-Digit PPI Code			
	Metals and Metal Products 10 (1)	Machinery and Equipment 11 (2)	Furniture and Household Durables 12 (3)	Transportation Equipment 14 (4)
1. Tally as of July 1975:				
Number of items	373	799	94	104
Number of price quotations	1,225	2,893	431	242
2. 1976 quality adjustments as percent of all price quotations	0.4	0.5	1.3	2.4
3. 1976 percent share by quality adjustment method:				
Direct comparison	32.2	32.4	29.0	11.9
Link, no change	49.1	38.8	56.4	20.9
Cost estimate	18.7	28.8	14.5	67.1
	100.0	100.0	99.9	99.9

Sources by row: **(1)** *BLS Handbook of Methods*, Bulletin 1910 (U.S. Bureau of Labor Statistics, 1976). **(2, 3)** Early and Sinclair (1983, table 2.1).

were still not adjusted explicitly for quality change, and, in one published example for the year 1965, there was not one single explicit quality adjustment based on cost data in the entire nonfood component of the CPI (Triplett 1971b, table 6.2, row 4, p. 187). In another example, for all nonfood items in 1966, the division among the four methods was 59.1 percent by direct comparison, 36.2 percent by deletion, 1.9 percent by linking, and 2.6 percent by an explicit size or quality adjustment (Triplett 1971b, table 6.1, p. 186).

More recently, Early and Sinclair (1983) have summarized the classification of quality adjustments in the PPI for 1976, as shown in table 3.2. Row 2 shows all types of quality adjustments, excluding deletion, as a percentage of all price quotations. For the PPI as a whole, there were surprisingly few quality adjustments, 455 total cases out of 108,756 total price observations for the year. However, as shown in row 3, explicit cost estimates accounted for a much larger share of the quality adjustments than in the CPI examples for the mid-1960s cited above, and direct comparison for a much smaller share. This evidence that the BLS is investing more resources in making explicit quality adjustments with manufacturers' cost data would suggest, ceteris paribus, the possibility that any quality adjustment bias in the PPI might be less in the 1970s than in earlier decades. Nevertheless, the incidence of only 455 cases of adjustment for more than 100,000 price observations also would suggest that some cases of quality change may be missed, or that some types of rapidly changing products (particularly in the electronics area) are not included in the indexes at all. And the incidence of producers' cost estimates outside transportation equipment is still relatively low, ranging from 15 percent for furniture and household durables to 29 percent for machinery and equipment.

Because the greater frequency of explicit cost adjustments in transportation equipment, particularly autos and trucks, still occurs as it has since 1959, the differential rates of price change recorded by different components of the PPI may be partly due to differences in measurement techniques. Compare, for instance, the following recorded percentage changes by decade:

	1950–60	1960–70	1970–80
Agricultural machinery (PPI 11–1)	32.0	31.5	129.0
Motor vehicles (PPI 14–11)	31.2	8.6	84.2
Difference	0.8	22.9	44.8

If one assumes that agricultural machinery and motor vehicles use roughly comparable technology, then one might argue that the slower rate of price change of motor vehicles in the 1960s and 1970s is at least partly due to the greater resources devoted to explicit quality adjustments. This still leaves open the possibility that the quality adjustments for motor vehicles have been too large as a result of manufacturer's overvaluations of the true cost of quality changes.

3.4 The Hedonic Regression Technique: Basic Features

The hedonic regression approach, as developed by Court (1939) and Griliches (1961), can be viewed as an alternative method to manufacturers' cost estimates in making quality change adjustments. It assumes that the price of a product observed at a given time is a function of its quality characteristics (x), and it estimates the imputed prices of such characteristics by regressing the prices of different models of the product on their differing embodied quantities of characteristics. Papers by Muellbauer (1974) and Lucas (1975) relate the hedonic method to Lancaster's (1971) theory of consumer choice. In equilibrium, Rosen (1974) has shown that the hedonic price schedule traces out differences both in user value and in manufacturing cost across models. Thus, the hedonic price approach represents not a new concept in the measurement of quality change, but rather an alternative to cost estimates to be used when practical factors make it more suitable than the conventional method.

The empirical problem is the estimation of the implicit prices of characteristics (x_{ij}) in which the dependent variable is a vector of observations on prices on n different models at a given time, and each of the m independent variables is a vector of data for a quality attribute for the different models.[4] The regression is most often specified in semilog form:

$$(3.1)\qquad \log p_i = a_0 + \sum_{j=1}^{m} b_j x_{ij} + u_i, \quad i = 1, \ldots, n.$$

The regression coefficients (b_j) from (3.1) are combined with the data on changes in quality attributes (dx_{ij}) to yield the estimated change in quality, a calculation that implicitly assumes that all quality changes can be treated as variations in the quality of the x_{ij}, and any change in the quantity of particular product attribute can be converted into an equivalent amount of quantity in the product itself through multiplication by the estimated b_j weights. Then the change in a quality-adjusted price index for the product can be calculated by subtracting for each model the change in its quality index (i.e., the change in each quality characteristic multiplied by its estimated implicit price) from the change in its actual transaction price or unit value.

An alternative approach to the estimation of quality-adjusted price change is to estimate the coefficients on one or more time dummy variables (D_t) in cross-sectional regressions for two or more years:

$$(3.2)\qquad \log p_{it} = a_0 + \sum_{t=1}^{N} d_t D_t + \sum_{j=1}^{m} b_j x_{ijt} + u_{it}, \quad i = 1, \ldots, n;\ t = 0, \ldots, N.$$

An aggregate index of price change then is obtained either from the series of d_t coefficients obtained in one regression like (3.2) run on data for a number of years, or from a string of d_t coefficients obtained from a series of "adjacent year" regressions on data for successive pairs of years. To the extent that the prices of quality characteristics are changing through time, the latter two-year technique allows the regression coefficients on the x_{ijt} to change frequently and is preferable.

There is no difference in principle between the two methods of estimating price change, deflation by a quality index and estimation of coefficients on time dummy variables. Results differ only to the extent that the time period for which the b_j implicit price weights are estimated in (3.1) differs from the time period in the adjacent year regression. For instance, Griliches's (1961) original hedonic study of automobiles contains three separate hedonic price indexes based on deflation by three quality indexes with b_j weights estimated in equation (3.1) for the initial year of the sample period alone, the final year

4. The partial regression coefficients of p_i and x_{ij} are precisely equal to Mp_i/Mx_{ij} only if a linear relation is postulated between the p_i and x_{ij}. If the variables are in semilog form, the coefficients stand for the percentage change in price with respect to a unit change in each quality attribute.

alone, and each year separately. The results differ in each case because of the standard index number problem that occurs when the estimated implicit prices of characteristics (b_j) shift from year to year. The results must be identical to the adjacent year technique if the b_j weights are estimated for the same pair of years.

Most studies have in practice chosen the time dummy variable approach, simply because it conveniently allows the "true" rate of price change to be read directly from the computer printout without any intermediate transformations. As Griliches (1967) has pointed out, however, changing samples of models in a regression like (3.2) for two adjacent years will cause some of the sample variation to be picked up in the time dummy coefficients, unless the sum ($\Sigma\ u_{it}$) of the "model effects" (the effect of left-out qualities) for both groups of models is identical. For instance, if we run a regression for two years, 1959 and 1960, and include only in the latter period models of compact cars that are more expensive per unit of size than full-sized cars, the regression will yield a positive coefficient on the 1960 time dummy, even in the absence of "pure" inflation, because the coefficients on the x_{ijt} variables are constrained to be equal in the next two years. This does not strike me as a major obstacle to the use of (3.2), however, because it is relatively easy, with only a small reduction in sample size (in most cases), to restrict the sample in years t and $t + 1$ to contain the same models and then include the new models in a regression on years $t + 1$ and $t + 2$. In his work with Ohta, Griliches explicitly endorses this view that the number of observations should be reduced where necessary to maintain constancy in the sample of models over time (Ohta and Griliches 1976).

Prices of used capital goods offer two advantages over new good prices for the study of quality change. First, they are by definition transaction prices rather than list prices. Second, the relative prices of two models of differing quality are determined by buyers' evaluations rather than by the decisions of sellers. But used good prices can usefully be studied only in conjunction with the hedonic technique, not by themselves. Any observation in a used asset market can be cross-classified along three dimensions, the time of the observation (t), the age of the asset (s), and its model designation (i). Differences in observed prices can be attributed to changes in "pure" price (P_t), differences in quality (X_{iv}) among models of different vintages ($v = t - s$), and the effect of depreciation on models of different ages (A_s). Maintaining the semilog approach,

$$p_{its} = e^{P_t} e^{X_{iv}} e^{A_s}. \tag{3.3}$$

This parametization of price involves some strong but necessary assumptions, as pointed out by Hall (1971). Equation (3.3) can be converted into a semilog regression equation if the errors (u_{its}) are assumed to enter multiplicatively in (3.5):

$$\log p_{its} = P_t + X_{iv} + A_s + u_{its} . \quad (3.4)$$

In this regression, the right-hand variables are dummy variables corresponding to the appropriate time periods, vintages, and ages, and their coefficients can be converted directly into index numbers. The major problem with (3.4) is that, without further assumptions, the rate of quality change cannot be identified, since observed prices that are consistent, say, with a given depreciation rate, low inflation, and rapid quality change, are also consistent with lower depreciation rates, more rapid inflation, and slower quality change.

Hall's major contribution is the suggestion that (3.4) can be combined with data on the quality characteristics of the different models and vintages in a single regression that generalizes the hedonic regression for new goods written in (3.5):

$$\log p_{its} = a_0 + d_1 D_1 + \sum_{s=1}^{L} a_s A_s + \sum_{j=1}^{m} b_j x_{ijt} + u_{its} . \quad (3.5)$$

The constant term now refers to a new item in the first year of the adjacent year regression, d_1 is the coefficient on the dummy variable for observations recorded in the second year and represents pure price change, and the L separate age coefficients measure the relative prices of models of different ages that are otherwise identical, and their difference is the year-to-year rate of depreciation. In more refined versions, the rate of depreciation, which is assumed constant for all makes and quality dimensions in (3.5), can be allowed to vary along these dimensions.

3.5 The Hedonic Regression Technique: Pitfalls

The hedonic technique has been regarded in some quarters as a replacement for the conventional method and as a general solution to the problem of quality change measurement. The Stigler Committee report endorsed the technique as "potentially of wide applicability" (NBER 1961, 36). But whether the technique is applied to new or used goods, it raises as many problems as it solves; minor changes in specification yield hedonic indexes that differ substantially in both secular trend and year-to-year movements. Existing hedonic studies contain conclusions about price trends that are too divergent to be useful without close critical inspection to resolve differences. In Griliches's original study of automobiles, for instance, the estimated change in "true" price ranges from −18.4 to 14.1 percent over the 1950–60 period, depending on the particular time period used to estimate the b_j implicit price weights in the quality change index (Griliches 1961, 187, table 9). Similarly, more recent studies contain estimates of price change for 1960–65 that range from −0.8 to 6.4 percent (Ohta and

Griliches 1976). The reluctance of the BLS to use the hedonic technique regularly in the CPI and WPI may in part reflect the ambiguous results of past hedonic studies.

There are a number of econometric specification problems that contribute to the ambiguity inherent in the hedonic technique.

3.5.1 The General Excluded Variable Problem

Many of the specific specification difficulties encountered in applications of the hedonic technique are subclasses of the excluded variable problem. Assume that the correct structural model determining the price of two different models in two adjacent years can be completely described with no error:

$$\log p_{it} = a_0 + d_1 D_1 + \sum_{j=1}^{m} b_j x_{ijt} + b_{m+1} x_{i,m+1,t} . \tag{3.6}$$

If a regression is run that erroneously includes only the first m quality attributes, and if the $m + 1$ variable is uncorrelated with the first variables, then the estimate of the rate of true price change (d_1) will be biased as follows whenever the quantity of the excluded quality attribute varies:

$$\hat{d}_1 = d_1 + b_{m+1} \Delta x_{i,m+1} . \tag{3.7}$$

Sometimes a variable is excluded not because of negligence or lack of data, but because in one or more periods it has been observed to be perfectly collinear with one of the included characteristics, let us say the first:

$$b_{m+1} x_{i,m+1,t} = \alpha x_{i1t}, \quad t = 1. \tag{3.8}$$

No problems arise if (3.6) is valid for both of the adjacent time periods, for the estimated value of b_1 will include the effect of the omitted variable, and the estimate of pure price change will be correct:

$$\hat{b}_1 = b_1 + \alpha; \quad \hat{d}_1 = d_1 . \tag{3.9}$$

Trouble arises, however, if, in period 2, the additional quality characteristic yields a marginal product per unit of the first quality characteristic that increases by ϵ over its value in period 1, so that (3.8) is replaced by:

$$b_{m+1} x_{i,m+1,t} = (\alpha + \epsilon) x_{i1t}, \quad t = 2. \tag{3.10}$$

The coefficient of b_1 estimated in a regression that excludes variable $m + 1$ understates the combined influence of characteristics 1 and $m + 1$ in period 2, and the extra "quality" of characteristic 2 is soaked up by the time dummy and is thus interpreted as pure price change:

$$\hat{b}_1 = b_1 + \alpha; \quad \hat{d}_1 = d_1 + \epsilon. \tag{3.11}$$

The preceding simple example generalizes to the statement that the d_1 coefficient is biased whenever there is a shift between one year and the next in the relation between the omitted variables and the included variables.

3.5.2 Performance versus Physical Characteristics

Users of capital goods value quality attributes for their ability to produce output, not for their physical specifications alone. The value of a locomotive depends on the ton capacity of freight cars it can pull at a specific speed, not on the horsepower of its diesel engine. If pulling power is always a constant multiple of horsepower, the hedonic price index will be unaffected by the choice between them as independent variables in the hedonic regression. But if technical progress increases the efficiency of the driving mechanism that transmits locomotive power from engine to wheels, and if horsepower rather than pulling power is the independent variable in the hedonic regression, then the increasing ratio of pulling power to horsepower is an omitted variable that by (3.11) causes an upward bias in the estimated increase in "true" price.

Almost every previous hedonic study has employed data on physical rather than performance data because the former have been more readily available. Two recent exceptions are Ohta and Griliches (1976) and Hogarty (1972); both have experimented with performance variables in automobile regressions. The empirical chapters below report scattered pieces of evidence on performance characteristics and attempt where possible to correct hedonic price indexes in cases where a changing ratio of performance to physical characteristics seems to have occurred.

3.5.3 Make Effects

Most studies have combined price and quality attribute data for all manufacturers. Dhrymes (1971) was the first to test and reject the null hypothesis that pricing behavior by different manufacturers is identical. In a much more extensive exploration, Ohta and Griliches (1976) presented evidence to support make effects, defined as significant coefficients on make dummy variables (M_k) in regressions that identify each model (i) by its make designation:

$$\log p_{itk} = \alpha_0 + d_1 D_1 + \sum_{k=1}^{r} m_k M_k + \sum_{j=1}^{m} b_j x_{ijt} \,. \tag{3.12}$$

In regressions for automobiles containing roughly four models per make, Ohta and Griliches find positive make effects for the three "luxury" makes (Cadillac, Imperial, and Lincoln) that imply prices roughly 25–35 percent higher than can be explained by the physical characteristics included in the regression. In previous studies, the high prices of these makes had been partly explained by high coefficients on the weight variable in regressions without make dummies, partly by upward biased coefficients on other variables, and had partly been reflected in positive residuals for these makes.[5]

Evidence on make effects has been subject to different interpretations. Dhrymes has argued that the make effect invalidates the basic hedonic assumption that the implicit prices of characteristics reflect consumer evaluations and instead implies that prices are marked up over cost by formulas that differ among manufacturers. Ohta and Griliches take the more sensible position that make effects stand as proxies for omitted characteristics, for example, the prestige associated with a luxury car, for which consumers are willing to pay a price. They test for the presence of the make effects in used car markets and discover that the "Cadillac prestige effect" persists permanently throughout the car's life, whereas that of Imperial and Lincoln is transient and evaporates quickly. Although Ohta and Griliches do not comment on the cause of the divergence between Cadillac and Imperial-Lincoln prices in the used car market, the result seems to be rather convincing evidence that the omitted quality variable for which the make effects stands as proxy represents primarily a subjective rather than an objective attribute. If the omitted variable were an objective feature, for example, fancy upholstery or electric windows, it would presumably be valued similarly in the used car market for the various luxury makes. It is possible, however, that the rapid depreciation of Lincoln and Imperial may at least partly reflect their relatively poor repair records.[6]

The make effects serve as a warning that comparisons of individual models of different makes across time must be performed carefully. In his original article on hedonic indexes, Court (1939) compared 1920 models of relatively expensive cars with 1939 models of relatively inexpensive cars that by then had grown to be almost identical in size. But at least part of the dramatic price reductions that he recorded may have been due to the positive make effect of the relatively expensive models in the early year.

3.5.4 New Products

The most intractable aspect of quality measurement has always been the introduction of a totally new characteristic that did not exist before, for example, the invention of television or of the automobile. The hedonic

5. See, e.g., Triplett (1969, 408–17). Triplett's truncated regressions contain estimates that the presence of automatic transmission, power steering, and power brakes raises the average car price by 30 percent, or about $900, although the value of these items as options was only about $350 at that time. The high value of this coefficient stands as a proxy for the "make effect" of the luxury cars. Although Triplett explicitly recognizes this (414), he fails to recalculate his regressions to include a dummy variable for make or price class and computes hedonic price indexes on the basis of his biased coefficients.

6. The frequency of repair record of Cadillac was distinctly superior to the other two luxury makes, as illustrated for the 1966–71 period that concerned Ohta and Griliches. See *Consumer Reports*, April 1972, 237–41.

approach can only evaluate quality improvements that either raise the quantity of characteristics that previously were included in a model or introduce characteristics that were formerly available only on higher-priced models and that "trickle down" to relatively lower-priced makes.

The most important single innovation in the capital goods industry during the postwar period has been the replacement of the clerk working with a calculating machine by the electronic computer. The hedonic technique has been used to evaluate quality improvements in computers, but not to evaluate the relative quality of computers and calculating machines. Similarly, no attention has been given to the saving in time made possible by the replacement of the rotary electric calculating machine by electronic calculators. Hedonic studies that ignore these aspects of technical change omit a variable and create an upward bias in price indexes. One can employ the "user-value" criterion to determine, at least subjectively, whether new products perform operations similar to the products they replace. If so, the relative performance of new and old products can be compared and changes in cost per equivalent operation can be calculated. The reduction in the cost of clerks made possible by the invention of electronic computers could in principle be estimated from time-and-motion studies of the calculations performed per clerk. The value of the invention of television might be estimated by the saving in time spent traveling to and from movie theaters.

3.5.5 Multicollinearity

Even if data were available on all relevant characteristics, multicollinearity usually precludes the estimation of separate implicit price coefficients for each. In comparison with a subcompact model, a full-sized Cadillac has greater weight *and* length *and* horsepower *and* includes as standard equipment automatic transmission, power options, and other accessories. Separate coefficients for each of these characteristics can be estimated only if the regression sample includes models that are short but heavy in weight, powerful but light in weight, and so on. Triplett (1969) showed that length and horsepower were so collinear with weight in his sample that the goodness of fit in automobile regressions was virtually unaffected when the length and horsepower variables were omitted. This poses a dilemma that weight is not desired for itself but only for the comfort that tends to be associated with it, and an innovation that reduces weight without reducing comfort (e.g., thin-wall casting of the engine) will be erroneously interpreted by the hedonic technique as a decline in quality (Triplett 1969, 416). In the absence of adequate data that directly measure performance variables like comfort, then, multicollinearity increases the likelihood that a bias will be caused by excluded variables.

3.5.6 The Treatment of Accessories

In most studies, multicollinearity has prevented the estimation of reliable and sensible implicit price coefficients on accessories like automatic trans-

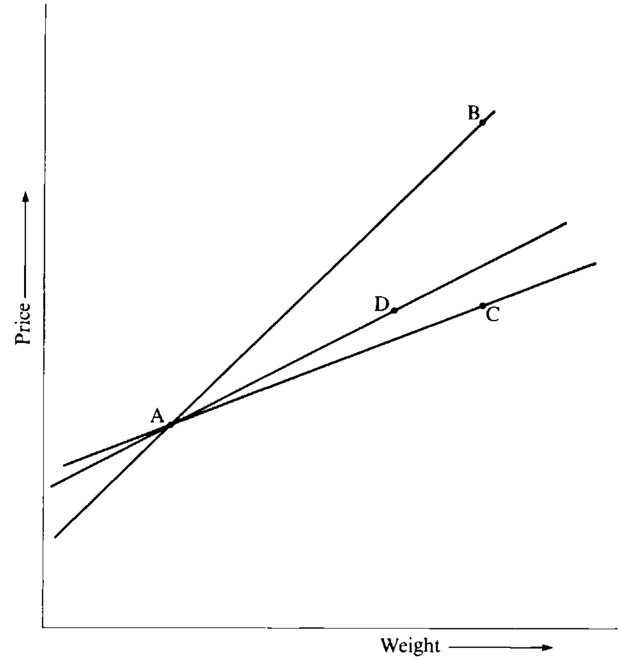

Fig. 3.1 Effect of accessory adjustments on estimated regression coefficients

mission or power steering for automobiles, or central air conditioning in hedonic studies of single-family houses. An alternative technique would be to adjust the price of each model for differences in standard equipment using the prices of accessories published in option lists. Similar adjustments could be made for the addition of accessories that were previously options or unavailable as standard equipment. No published hedonic study has made systematic adjustments of this type for accessories and optional equipment.

A problem is created, however, if a Cadillac is normally sold with air conditioning but its price in a hedonic regression is included net of this accessory. In this case, the weight variable is incorrect, because the published shipping weight includes the weight of the air conditioning equipment. In figure 3.1, point *A* represents the price and weight of a subcompact sold without air conditioning as standard equipment, and *B* represents the published price and weight of a Cadillac sold with this accessory standard. The slope of *AB* represents the coefficient of weight computed in a regression on published data. If the price per pound of the air conditioning unit is higher than the price per pound of the rest of the car, *D* is the Cadillac's price with both the price and the weight of air conditioning excluded, and *AD* is the "true" slope of weight for the car by itself. If the price of air conditioning is erroneously omitted without a weight adjustment, the weight coefficient is underestimated as *AC*.

Most previous studies have taken the approach of estimating the price of accessories without making an adjustment for weight and have thus estimated *AC* in figure 3.1 as the downward-biased weight coefficient. In their more recent investigation, Ohta and Griliches (1976) have argued that a high rate of purchase of power steering on large cars implies a market evaluation that this accessory is a necessary feature. If so, they argue, power steering is part of the price of weight, and the true weight coefficient is *AB*, not *AD*. While the argument may be valid for a single cross section, it is untenable in long historical studies. For instance, the 1948 Buick Roadmaster and the 1971 Chevrolet Impala are virtually identical in length and weight, but no units of the former and virtually all units of the latter model were equipped with power steering.[7] The market did not shift its tastes between 1948 and 1971, but simply responded to an invention that was unavailable in the earlier year. A comparison of true price change between the two years should be based on the prices of comparably equipped models. The 1971 price and weight of the power steering option (and anything else not included on the Buick) must be excluded from the Chevrolet's specifications before a valid comparison can be made. If the price per pound of each accessory is identical to that of the rest of the car, these adjustments make no difference to the estimated coefficient on weight, but in fact the price per pound of an accessory like automatic transmission is relatively high, $2.74 versus $0.85 for the basic car in the case of the 1971 Chevrolet Impala V-8.[8] The Ohta-Griliches procedure leads to a coefficient on weight that is biased upward and a hedonic price index that is biased upward by registering as a price increase the introduction of power steering as standard equipment on a car of a given weight that previously offered power steering as an option.

3.5.7 Functional Form

Almost all studies of automobiles and most other commodities have continued Griliches's original practice of fitting regressions in semilog form. A unit change in a quality characteristic is thus assumed to be associated with a fixed percentage change in price, that is, an increase in weight by one pound raises the price of a 1971 Cadillac by three times as much as that of a 1971 subcompact. The popularity of the semilog form and its acceptance in formal tests by Dhrymes may simply reflect the nonlinear make effect of the luxury nameplates. Another standard assumption is that multiplicative interaction terms are absent; weight and horsepower influence price

7. The weight and length of the 1948 Buick were 4,160 pounds and 217.5 inches, respectively (*Consumer Reports*, May 1948, 209), while the same dimensions for the 1971 Chevrolet were 4,014 and 216.8 (*Automotive News 1971 Almanac Issue*, 46). In 1970 full-sized Chevrolets, 95.3 percent of all units were sold with power steering included (*Automotive News 1971 Almanac Issue*, 62).

8. Information obtained from the General Motors Corp. in letter dated 27 July 1972.

separately but not jointly, even though an increase in weight requires an increase in horsepower to maintain performance.

3.5.8 List Prices versus Transaction Prices

In most hedonic regression studies, the dependent variable is the manufacturer's list price. Discounting that causes variations in the ratio of transaction prices to list prices is not reflected in this dependent variable or, therefore, in the final hedonic price index. The best evidence that list prices are an unreliable guide to actual transactions prices refers to retail sales of automobiles. In his study of the auto industry, White (1982, 105) reports that, "although list prices get all the publicity in any discussion of auto prices, in fact they are only symbolic. Few sales actually get made at list prices." During the early postwar period, transaction prices were above list prices as a "gray market" in automobiles developed, whereas, during most of the period after the Korean War, discounting was common. For instance, a series of studies by Jung (1950, 1960) on the Chicago retail auto market reported an average discount off list price of 16 percent in 1959–60. A Federal Trade Commission survey for 1968 and 1969 reported that half the full-sized models in the survey sold for 15–20 percent below list prices, and that about 60 percent of the intermediates sold for 10–15 percent below list prices (quoted in *Consumer Reports*, April 1971, 203). Because the ratio of transaction to list prices was higher in the early postwar period than after 1953, the secular rate of change of any hedonic price index based on list prices for the entire postwar period is biased upward. And, to the extent that the size of discounts and rebates has varied procyclically, existing hedonic price indexes understate the true variability of auto prices over the business cycle.

In lieu of detailed data on transaction prices of new cars, one possible alternative would be to use prices for one-year-old used cars. The prices of these close substitutes for new cars may reflect to a substantial degree the variations of discounts on new cars. Unfortunately, previous studies have largely ignored the wealth of data available on used car prices. Ohta and Griliches have studied used car prices for the interval since 1961, but not for earlier years, when the main variations in discounts apparently occurred.

3.6 Relative Advantages of the Hedonic and Conventional Methods

The conventional specification method is similar in principle to the hedonic method. A hedonic study that has sufficient data to perform regressions of price on three characteristics, say length, horsepower, and weight, also provides sufficient data to compute a conventional price index for each of several length-horsepower-weight classes. In a period during which no pure price change occurs and any excluded quality characteristics are unaltered, an increase in unit value due to a shift in quality to a higher length-horsepower-weight class is treated as quality change rather than price change in the conventional method because there is no price change for any individual commodity class.

While the conventional method would be able to come close to the results of the hedonic method if more observations were collected, and if exactly the same quality characteristics were used, in practice there could never be enough continually occupied classifications to avert the need for special adjustments when a new higher-quality model is introduced but remains within the same length-horsepower-weight class as the old model. Even in this case, the methods would be identical if a simultaneous overlapping price observation on the new and the old model were available, since the same market price information used to estimate quality change in the hedonic regressions could be used for a special adjustment in the conventional method. But when no overlap is available, the conventional method has no systematic method of adjustment, whereas the hedonic method "creates" an overlap observation by estimating a regression line relating price to quality and calculating the implicit price of a model containing any given amount of the quality characteristics used in the regressions.

The most serious disadvantage of both methods is that neither can measure changes in the relation between excluded and included quality dimensions. A bias will result unless all excluded quality characteristics maintain a fixed relation with included characteristics. For instance, when we estimate the difference between the price of a 1948 Buick and that of a 1948 Chevrolet as depending on differences in length, weight, and horsepower, our coefficients are also picking up the influence of the Buick's carpeting, fancy doorknobs, and plush upholstery. If we discover that a 1970 Chevrolet has the same length, weight, and horsepower as a 1948 Buick, a quality index computed from the 1948 coefficients implies that both cars are identical in every other respect. In fact, the 1970 Chevrolet could be either higher or lower in quality in other respects than the 1948 Buick.

There is no escape from the excluded variable problem with either the hedonic or the conventional methods. An increase in the ratio of excluded to included characteristics creates an upward bias in the coefficient on the time dummy variable in hedonic regressions. Similarly, in the conventional method, an increase in the amount of "excluded quality" relative to the quality attributes included in a product's specifications shows up as an increase in the price index. Only if changes in quality attributes outside the written specification are explicitly adjusted by linking or a manufacturer's cost estimate does the conventional method treat the change as representing an increase in quality. The hedonic technique is superior in its ability to estimate the value of increases in dimensions like length, horsepower, and weight but is prevented by multicollinearity from estimating reasonable coefficients for the value of accessories.

No methodological rulebook states that one technique must be used exclusively. The best procedure for empirical price measurement is to combine the conventional and hedonic methods, taking advantage of the relative strengths of each. To the maximum extent possible, the conventional approach should be used to adjust the prices of different models for discrete options, accessories, and added features. Then the hedonic method should be employed to explain the remaining price difference as a function of basic dimensions or performance characteristics. Care must be taken that the two types of adjustments are consistent by adjusting not only price but also weight and performance characteristics for the effect of added accessories.

One possible procedure would be to proceed in steps as follows:

1. A preliminary hedonic regression should be run to determine the major price-determining quality characteristics. The effects of multicollinearity should be explored and characteristics with coefficients that are unstable, unreasonable, or of incorrect sign should be excluded.
2. Lists of option prices or manufacturers' cost estimates should be obtained to adjust the transaction price of each model for changes in discrete features and accessories.
3. Interactions between the variables in the hedonic regression and the adjustments in 2 should be investigated. If weight is variable in the regression, for instance, a correction should be made for the weight as well as the price of accessories, using weight data obtained from manufacturers or from mail-order catalogs.
4. Then a "final" hedonic regression can be estimated from the observations on adjusted price and adjusted physical or performance characteristics. A hedonic price index can be calculated from the time dummy variables in the final regression.

3.7 Implementation of Quality Adjustments for Changes in Operating Efficiency

A major theme of chapter 2 was the need to make quality adjustments for changes in the operating efficiency of durable goods. This section discusses practical problems in implementing such adjustments. While the analysis is framed in terms of changes in fuel economy, many of the same points could be made about changes in maintenance requirements.

Although little attention has been paid to changing fuel economy in past studies of quality change, it should be possible in principle to make appropriate quality adjustments using either the conventional or the hedonic method. In the conventional method, a manufacturer would report that a new model has improved fuel economy as compared to the model it replaces, owing, say, to electronic fuel ignition. The BLS can then ask for an estimate of the extra cost of the new ignition system and make an appropriate adjustment. Since, as we saw in chapter 2, in equilibrium the value to users of fuel efficiency improvements should be equal to their manufacturing cost, the conventional method should be adequate in this case. A problem is created for the conventional method, however, in another example where the electronic ignition replaces the previous mechanical ignition and yields improved fuel economy with no increase in manufacturing cost. This is what we have previously called a "nonproportional" quality improvement. Under present procedures, no adjustment would be made, even though the cost of producing the quality characteristic "fuel economy" has fallen. In terms of figure 2.5, what needs to be estimated is the marginal cost of producing the extra quantity of input characteristics that would have been required to yield the observed increase in net revenue or consumer value in the absence of the technical change. The hedonic method could in principle cope with the nonproportional improvement if the marketplace provides models with a broad mix of performance and fuel economy characteristics, thus allowing the imputed price of fuel economy to be estimated in the hedonic regression. However, in fact, a single cross section for a given year does not typically provide the required mix of automobiles. Because large cars have poor fuel economy, and small cars have good fuel economy, there is negative collinearity between dimensional variables and fuel usage. It is the absence of large cars with excellent fuel economy, or small cars with poor fuel economy, that prevents the reliable estimation of a coefficient for the imputed price of fuel economy. While the proliferation of foreign cars in recent years has provided a broader product mix, including relatively large cars with relatively good fuel economy, foreign and domestic cars differ in numerous other dimensions as well (frequency-of-repair record, extent of standard accessories).

Widespread interest in the effects of changing fuel prices after the two oil shocks (1973–74 and 1979–80) is reflected in two papers by Kahn (1981) and Ohta-Griliches (1986) that focus on the role of fuel efficiency in hedonic regressions. However, neither paper yields significant estimates of the value of fuel economy changes that could be used to produce suitable quality-adjusted price indexes. Further, their main focus is on the way in which changing fuel prices alter consumer evaluations of dimension and performance attributes in the used car market, not the imputed value of fuel economy in the new car market.

Because the conventional method cannot deal with nonproportional changes in fuel efficiency, and because the hedonic method is plagued by multicollinearity in this case, an alternative technique must be devised. The principle outlined in chapter 2 is to base a quality adjustment on a comparison of the price of old and new models with their respective ability to earn net revenue. In terms of the conventional method, the observed increase in the actual price of the new model is adjusted by the calculated increase in discounted future net revenue. The results of such calculations will always admit to a wide range of possible estimates, depending on the assumed real discount rate and future time path of the real price of fuel. It is possible that changing prices on the market for used cars and airplanes could be used to check the validity of the computed adjustments.

In the study of aircraft prices in chapter 4, the "net revenue" method is used by itself, not in combination with either the hedonic or the conventional methods. Changes in net revenue due to changes either in performance characteristics (plane size and speed) or in fuel usage are treated symmetrically in the calculation of net revenue, and then quality adjustments are computed from a comparison across pairs of old and new models of their ability to earn net revenue. This method cannot be used for automobiles used by consumers, since no net revenue is generated. Here it is preferable to compute quality adjustments by the conventional or hedonic method, and then to make an extra adjustment for the value of fuel savings. As Wilcox (1984) has shown, it is possible to compute an index of "pure" changes in fuel efficiency, by running a regression of fuel economy on other fuel-using characteristics like weight and engine displacement. Then changes in this fuel economy index can be converted into user value terms with suitable assumptions regarding the discount rate and future real fuel prices.

Consider a hypothetical example in which the price increase in a particular model car between 1956 and 1957 is observed to be $150. Let us assume that there are no changes in quality characteristics between the two years besides a modest increase in weight, and that there is no change in observed fuel economy. A hedonic regression across all models yields an imputed value for the increase in weight of, say, $40. Let us also assume that the Wilcox-type fuel economy regression indicates that the observed increase in weight should have reduced fuel economy by, say, two miles per gallon (mpg). The fact that observed mpg did not decrease signals a pure change in fuel economy that requires a quality adjustment. If the computed present value of the fuel economy improvement is, say, $50, then the final quality adjustment is $90 ($40 for weight and $50 for fuel economy), and the observed price of $150 is reduced by the quality adjustment to $60.

The need for a fuel economy adjustment in this example can be stated in another way. The hedonic regression of auto price on characteristics excluding mpg yields a downward-biased coefficient on weight due to a negative correlation between weight and the excluded mpg variable. The estimated coefficient understates the true user value of extra weight (which partly proxies for the other beneficial excluded variables that are positively correlated with weight), since it includes a negative component corresponding to the fuel cost imposed by extra weight. If there were never any technological shift in the Wilcox-type fuel economy regression, then the conventional application of the hedonic technique would properly adjust for the effects of added weight, combining both its positive and its negative aspects. But the shift in the fuel economy regression occurs because the new model does not require as much fuel, given its weight, as would have been expected given the correlation between fuel usage in weight across the old models and across the new models taken separately, and this shift amounts to a nonproportional quality change that must be taken into account.

A further complication occurs in the first half of the 1970s, when there was an adverse shift in fuel economy associated with antipollution regulations. This book accepts the official government treatment of antipollution devices, which is to treat their extra cost as a quality change rather than as a price change, on the grounds that society at large has received benefits that roughly balance the costs. Thus, if we can determine from industry commentaries that the fuel economy deterioration was due entirely to regulation, then we can ignore it and not apply a negative quality correction.

After 1975, fuel economy improved dramatically, but the price of automobiles soared. It is possible that this episode of improving fuel economy represented a proportional quality change, with the marginal cost of achieving higher mpg balancing the user value of the improvements. It will be interesting to determine in chapter 8, which calculates the explicit fuel economy adjustments, whether the BLS practice of ignoring shrinking dimensions in the process of automobile "downsizing" represents an adequate approximation to the "net" quality adjustment called for by the combination of smaller size and improving fuel economy.

To which price index should a fuel economy adjustment be applied? It has been suggested that a fuel efficiency adjustment is acceptable for an "input" price index, that is, one used to measure the purchase price of a capital good by a final user, but not for an "output" price index for use in deflating the production of machines. The skepticism for output price indexes concerns changes in fuel prices. Should the productivity of the machinery industry change when fuel prices change? The proposed method, as set out previously in section 2.9, is free from such objections, because it is based explicitly on a comparison in which the prices of variable inputs (labor, materials, and fuel) are held constant. Net revenue obtainable from new and old models is calculated with the same output price and variable input prices applying to both models.

Thus, a change in fuel prices by itself does nothing to alter the productivity of the machine-producing industry. If a change in fuel prices induces the introduction of a new model that moves up the cost function by introducing greater fuel efficiency at a higher capital goods price, the technique does not register a quality improvement (more precisely, there is no quality change when the capital goods price and net revenue rise in proportion). However, a technical change that allows an improvement in fuel economy with a less-than-proportional increase in the capital goods price *does* count as an improvement in quality. The only ambiguity involved in this technique is a standard index number problem: the value of a nonproportional improvement in fuel efficiency will differ in different fuel pricing regimes. To minimize the sensitivity of my adjustments to this problem, in the aircraft chapter differences in net-revenue generating ability between new and old models are evaluated at several different regimes of output and input prices and then averaged (see sec. 4.5).

3.8 Comparison of "Closely Similar" Models

To cross-check for secular bias in the conventional or hedonic methods, the secular change in price between two widely separated years can be compared. This approach was first used by Court (1939, p. 106) in his original paper on hedonic indexes to compare automobile models of roughly similar dimensions in 1920 and 1939. I used the same approach to compare 1959 and 1970 models in two early papers (1970, 1971b). A particular advantage of the method is that available resources can be invested in a more extensive effort to collect detailed information on specifications, on the presumption that forty specifications of two models manufactured twenty years apart will yield a more accurate index of secular quality change than information on two specifications for two models in each of the intervening years.

A major difficulty with the technique is the standard index number problem that prices of specifications or accessories in widely separated years may be quite different and yield differing estimates of quality change. Another drawback is the difficulty of estimating the implicit price of changed specifications when option or manufacturers' cost estimates are unavailable or unreliable. Finally, items twenty years apart will rarely be identical in size; if we find that a 1948 Buick is 100 pounds heavier and two inches shorter than a 1970 Chevrolet, what implicit prices should be applied to these differences? Presumably, the technique should be used to supplement and check the hedonic or conventional methods, and information from hedonic regressions or any other source can be used to estimate the implicit prices of a change in specifications. Finally, the comparison of a relatively high-priced model in an early year with a relatively low-priced model in a later year that has grown to the same overall size may err if prices are influenced by the "make effects" or "prestige effect" discussed above.

Most of the problems and ambiguities of the hedonic method discussed earlier in this chapter apply as well to the comparison of "closely similar" models. Performance variables should be used in place of physical characteristics where data are available. When we compare the prices of a 1946 DC-6 airplane with a 1971 DC-10, we are interested in the price per unit of seat-mile capacity (adjusted if possible for the saving of time made possible by jet speeds), not in relative length, weight, or horsepower. Comparisons of "closely similar" models also must take into account secular changes in fuel economy and decompose these into changes accounted for by altered quantities of performance characteristics and a residual "pure shift" effect.

3.9 Summary and Conclusion

This chapter examines the conventional and hedonic techniques of quality adjustment and suggests circumstances in which they should be used in combination rather than separately. Then it adds two additional techniques to be used where appropriate, adjustments for "pure shifts" in energy usage not connected with changes in the quantity of other quality characteristics, and comparisons of "closely similar" models.

In the conventional method, one of four explicit quality adjustment procedures comes into play when the BLS is notified by manufacturers that one or more characteristics of a new model are altered from those on the old model. The most satisfactory procedure is an explicit estimate of the cost of the quality change, provided by the manufacturer. While explicit cost adjustments have become more common in the past decade, the other three procedures have normally been used. Under "direct comparison," any change in price between the two models is assumed to be entirely a price change and not a quality change; under "linking," any change in price is assumed to be entirely due to a quality change and not to a price change. The fourth procedure, deletion, imputes to the new model a price increase equal to all the other price reports in the eight-digit commodity classification. One major problem with the conventional method is the possibility of a shift over time toward more accurate measurement as the fraction of explicit cost adjustments has increased. The prevalence in the first half of the postwar era of quality adjustments by the direct comparison procedure, which amounts to no adjustment at all, suggests that price, output, and productivity measures for durable goods indexes might contain a greater bias in the earlier years than more recently.

The hedonic method seems to be more naturally suited than the conventional method to quality adjustments that take the form of changes in dimensions or performance along a continuous dimension, for example, weight, horsepower, or acceleration. The conventional method seems more suited to discrete changes in characteristics than can be valued by manufacturer cost estimates or by option prices, for example, air conditioning and power brakes. Because multicollinearity limits the number of explanatory variables that can be included in a hedonic regression, the potential bias introduced by excluded variables, and the instability of coefficients on variables like air conditioning, is more serious than in the conventional method. An offsetting advantage is that the hedonic technique can be applied uniformly to a long time series of data, for example, 1947–80, as in this book, whereas the prevalence of quality adjustments in the official indexes has not been uniform over the years. The complementary advantages of the two techniques can be utilized in actual measurement by using both, so long as care is taken to avoid double counting.

The chapter suggests two different approaches to quality adjustments for nonproportional changes in energy efficiency. One method, probably well suited for automobiles and electric generating equipment, is to estimate a fuel economy function separately and to compute a quality adjustment for the present value of energy efficiency improvements as a supplement to the conventional and/or hedonic method. Another method, better suited for commercial aircraft where sample sizes are too small for adequate hedonic regressions, is to implement the comparison of selling price with net revenue, as suggested in chapter 2.

A final suggestion in the chapter is that the validity of all quality adjustments, whether by the conventional or the hedonic technique, and whether involving performance characteristics or fuel economy, can be cross-checked by comparisons of "closely similar" models over long time spans. Sometimes, there will be no such models, particularly in industries like computers with pervasive technical change. But in other cases, for example, automobiles and appliances, it is sometimes possible to find models sold ten or twenty years apart that share roughly the same dimensions and differ in other respects to which a value can be attributed. Such comparisons are provided in this book, not only to serve the cross-checking purpose, but also because comparisons of models widely separated in years help provide an interesting chronicle of types and forms of quality change.

This Page Intentionally Left Blank

II Studies of Individual Products