

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

Volume Title: Scanner Data and Price Indexes

Volume Author/Editor: Robert C. Feenstra and Matthew D. Shapiro, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-23965-9

Volume URL: <http://www.nber.org/books/feen03-1>

Conference Date: September 15-16, 2000

Publication Date: January 2003

Chapter Title: The Measurement of Quality-Adjusted Price Changes

Chapter Author: Mick Silver, Saeed Heravi

Chapter URL: <http://www.nber.org/chapters/c9739>

Chapter pages in book: (p. 277 - 316)

# The Measurement of Quality-Adjusted Price Changes

Mick Silver and Saeed Heravi

## 9.1 Introduction

A major source of bias in the measurement of inflation is held to be its inability to properly incorporate quality changes (Boskin 1996; Boskin et al. 1998; Diewert 1996; Cunningham 1996; Hoffmann 1998; Abraham, Greenless, and Moulton 1998). This is not to say statistical offices are unaware of the problem. Price collectors attempt to match the prices of “like with like” to minimize such bias. However, comparable items are often unavailable, and methods of implicit and explicit quality adjustment are not always considered satisfactory (Reinsdorf, Liegey, and Stewart 1995; Armknecht, Lane, and Stewart 1997; Moulton, LaFleur, and Moses 1998).

Alongside this is an extensive empirical literature concerned with the measurement of quality-adjusted price indexes at the product level. The

Mick Silver is professor of business statistics at Cardiff Business School, Cardiff University. Saeed Heravi is senior lecturer in quantitative methods at Cardiff Business School, Cardiff University.

This study is part of a wider project funded by the U.K. Office for National Statistics (ONS). The authors are grateful to the ONS for permission to reproduce some of this work in the form of this paper. The views expressed in the paper are those of the authors and not the ONS. Any errors and omissions are also the responsibility of the authors. Helpful advice during the working of this paper was received from David Fenwick (ONS), Adrian Ball (ONS), Dawn Camus (ONS), and Pat Barr (GfK Marketing), and valuable programming assistance was received from Bruce Webb (Cardiff University). The paper has also benefited from useful comments from Ernst Berndt (Massachusetts Institute of Technology), Erwin Diewert (University of British Columbia), Rob Feenstra (University of California, Davis), John Greenlees (Bureau of Labor Statistics), Christos Ioannidis (Brunel University), Matthew Shapiro (University of Michigan), and Ralph Turvey (London School of Economics). In particular, the authors had sight of a draft Organization for Economic Cooperation and Development manual by Jack Triplett (Brookings Institution), and this proved most useful for section 9.5, in which a variety of approaches are used. The usual disclaimers apply.

main approach is the use of hedonic regressions (but see Blow and Crawford 1999 for an exception) in which the price of a model, for example, of a personal computer is regressed on its characteristics. The data sources are often unbalanced, panel cross-sectional time series from catalogs or web pages. A hedonic regression is estimated that includes the characteristics of the variety and dummy variables on time, the coefficients on the time dummies being estimates of the changes in price having controlled for changes in characteristics. This is referred to as the *time dummy variable hedonic method*. The quality-adjusted price index is taken from the coefficients on the time dummies in the hedonic regression. There is usually little by way of data on quantities, and thus weights, in these estimates. Yet estimates from hedonic regressions have been used to benchmark the extent of bias due to quality changes in consumer price indexes (CPIs; Boskin 1996; Boskin et al. 1998; Hoffmann 1998).

In this first part of the paper (sections 9.2 to 9.4) we argue against the use of this widely adopted time dummy variable approach. It is set against theoretical developments in the measurement of exact hedonic indexes by Fixler and Zieschang (1992), Feenstra (1995), and Diewert (chap. 10 in this volume) and superlative index number formulas by Diewert (1976, 1978). The exact hedonic approach also uses hedonic regressions, but it differs from the time dummy hedonic approach in two ways. First, the coefficients on the quality characteristics are not restricted to be the same over time, as is the case with the time dummy variable method. Use is made of repeated cross-section regressions in each period, rather than a single panel-data regression with dummy variables. Second, a formal sales weighting system is used in the exact approach, as opposed to implicit, equally weighted observations. The exact approach also provides estimates of, and bounds for, cost-of-living indexes (COLI) based on economic theory. Cost-of-living indexes measure the ratio of the minimum expenditure required to maintain a given level of utility. The dummy variable approach is shown to be a restricted version of the exact hedonic approach. Concordant with the development of the theory for the exact hedonic approach has been developments in data availability. Use is made of scanner data from electronic-point-of-sale bar code readings, which provide a sufficiently rich source to implement the exact hedonic approach and compare it with results from the dummy variable method.

There are of course other variants of the time dummy variable method. A sales-weighted least squares estimator could be used, or estimates could be made on a chained basis with, for example, a comparison between January and February being based on hedonic regressions for these months only, with a time dummy for February, and similarly for February and March, March and April, and so on. The estimates of price changes over these binary comparisons would be linked by successive multiplication to form a chained estimate over the whole period.

Against all of this Turvey (1999a,b) has proposed, on pragmatic grounds, a matched method akin to that adopted by statistical offices. The availability of scanner data with information on the relative expenditure of each product variety allows the compilation of matched indexes using exact and superlative formulas. The matched approach identifies the price of particular varieties and compares this with the prices of the same varieties in subsequent periods. It thus reduces the need to use regressions for the measurement of quality-adjusted price changes since “like” is being compared with “like.” It is shown here how the exact hedonic approach based on Feenstra (1995) and the matched approach are related, having respective pros and cons for the measurement of quality-adjusted COLI. Also provided are estimates, using scanner data, of quality-adjusted price indexes for washing machines using all three of these approaches.

It is worth noting that scanner data are now available in Europe and North America for a wide range of consumer durables and fast-moving goods. The coverage of the data is often quite extensive, being supplemented by store audits for independent stores without bar code readers (see Hawkes and Smith 1999). Market research agencies including ACNielsen and GfK Marketing Services collate and supply such data. Their use for validation and other purposes is now recommended for the compilation of consumer price indexes by Boskin (1996) and for direct use by Diewert (1993) and Silver (1995).

There have been a number of studies using matching on this rich data source in which prices of items with a particular specification are compared with their counterparts over time. These include Silver (1995), Saglio (1995), and Lowe (1999) for television sets and Reinsdorf (1996), Bradley et al. (1998), Haan and Opperdoes (1998), Dalen (1998), and Hawkes and Smith (1999) for selected food products. The matching used is often at a highly disaggregated level, matching individual item codes with their counterparts over time. There have been fewer studies that compare the results of alternative methodologies: these include Silver’s (1999) study on TVs, which uses the dummy variable and exact hedonic approaches, and Moulton, LaFleur, and Moses’ (1999) and Kokoski, Waehrer, and Rozaklis’s (1999) studies on TVs and audio products, which use the dummy variable and hedonic quality-adjusted matching approaches. Studies, especially those using the dummy variable approach, invariably focus on a single methodology with little interest in the relationship between methods. An early and notable exception comparing the results from hedonic regressions and matching, although not based on scanner data, was Cole et al. (1986). In this study we show how all three approaches are related and contrast the results for the case of washing machines.

In section 9.2 we outline the three methods of measuring quality-adjusted price indexes and show how they are related. Section 9.3 provides a description of the data, the application in this study being to monthly data

on washing machines in the United Kingdom in 1998. The implementation of the three methods and their results are also outlined in section 9.3. Conclusions on the appropriate method to measure quality-adjusted price changes using scanner data are in section 9.4.

The first part of the paper is concerned with how best to measure quality-adjusted price changes given scanner data. The second part is an initial attempt to replicate the practice of statistical offices with regard to quality adjustment. The same scanner data are used matching prices between product varieties in a base month, January, with their counterparts in February, and similarly for January with March, January with April, and so on. When a product variety is missing in the current month, different variants of implicit and (hedonic) explicit adjustments are undertaken, as would be used by a statistical office, and the results from these methods are compared. The methodology is explained and the results are presented and discussed in section 9.5.

## 9.2 Quality-Adjusted Price Indexes: Three Approaches Using Scanner Data

This section outlines three methods for measuring quality-adjusted price changes using scanner data: the time dummy variable hedonic method, an exact hedonic approach, and a matching technique, which can utilize exact and superlative formulas. It is reiterated that

- both the time dummy and exact hedonic methods use hedonic regressions, the former using a single panel-data regression, whereas the latter uses repeated cross-sectional ones;
- the time dummy hedonic method implicitly weights each observation equally in the regression, whereas the exact indexes have weighted formulations;
- the need for hedonic regressions is reduced when matching is effective.

### 9.2.1 Time Dummy Variable Hedonic Method

The hedonic approach involves the estimation of the implicit, shadow prices of the quality characteristics of a product. Products are often sold by a number of manufacturers, who brand them by their “make.” Each make of product is usually available in more than one model, each having different characteristics. A set of ( $z_k = 1, \dots, K$ ) characteristics of a product is identified, and data over  $i = 1, \dots, N$  product varieties (or models) over  $t = 1, \dots, T$  periods are collected. A hedonic regression of the price of model  $i$  in period  $t$  on its characteristics set  $z_{tki}$  is given by

$$(1) \quad p_{it} = \beta_0 + \sum_{t=2}^T \beta_t D_t + \sum_{k=1}^K \beta_k z_{tki} + \epsilon_{it}$$

where  $D_t$  is a dummy variable for the time periods,  $D_2$  being 1 in period  $t = 2$ , zero otherwise;  $D_3$  being 1 in period  $t = 3$ , zero otherwise, and so on.

The coefficients  $\beta_t$  are estimates of quality-adjusted price changes, that is, estimates of the change in the price between period 1 and period  $t$ , having controlled for the effects of variation in quality (via  $\sum_{k=1}^K \beta_k z_{tki}$ ).

The theoretical basis for the method has been derived in Rosen (1974), in which a market in characteristic space is established (see also Triplett 1987; Arguea, Haseo, and Taylor 1994). There is a plethora of studies of the above form as considered by Griliches (1990), Triplett (1990), and Gordon (1990), but subsequently including Nelson, Tanguay, and Patterson (1994); Gandal (1994, 1995); Arguea, Haseo, and Taylor (1994); Lerner (1995); Berndt, Griliches, and Rappaport (1995); Moulton, LaFleur, and Moses (1999); Hoffmann (1998); and Murray and Sarantis (1999). An issue of specific concern is the choice of functional form to be used. There has been support for, and success in, the use of the linear form, including Arguea, Haseo, and Taylor (1994); Feenstra (1995); Stewart and Jones (1998); and Hoffmann (1998). The semilog formulation has also been successfully used in studies including Lerner (1995); Nelson, Tanguay, and Patterson (1994); and Moulton, LaFleur, and Moses (1999). Studies using, and testing for, more complex functional forms have been advocated by Diewert (chap. 10 in this volume) and generally applied to housing (Rasmussen and Zuehlke 1990; Mills and Simenauer 1996) with some success, such studies for consumer durable goods (using flexible functional forms and neural networks; Curry, Morgan, and Silver 2001) being more limited.

The data sources used may be scanner data but are often specialist magazines or mail-order catalogs. The approach as conventionally used is not without problems. First, it implicitly treats each model as being of equal importance, when some models will have quite substantial sales, whereas for others sales will be minimal. If data are available on sales values, a weighted least squares estimator may be employed (Ioannidis and Silver 1999). Second, the prices recorded are not the transaction price averaged over a representative sample of types of outlets, but often a single, unusual supplier.

A final problem arises with the manner in which the time dummy variable method takes account of changing marginal values (coefficients) over time. It is the usual practice that the coefficients are held constant and thus not allowed to reflect changes in the marginal worth of the characteristics. Dummy slope coefficients on each characteristic for each period would relax the constraint. Yet this would render the estimate of quality-adjusted price changes, the coefficient on the dummy (time) intercept, dependent on the values of the performance characteristics (Silver 1999; Kokoski, Waehrer, and Rozaklis 1999). We will see that the above problems are dealt with in the exact hedonic formulation, the dummy variable hedonic method being a restricted version of the exact hedonic approach.

## 9.2.2 Exact (and Superlative) Hedonic Indexes

Konüs (1939) and Diewert (1976) define a theoretical COLI,  $P_c$ , as the ratio of the minimum expenditure required to achieve a given level of utility,  $U$ , when the consumer faces period  $t$  prices compared with period  $t - 1$  price,  $p_t$ , and  $p_{t-1}$ ; that is,

$$(2) \quad P_c(p_t, p_{t-1}, U) = \frac{E(p_t, U)}{E(p_{t-1}, U)}$$

The above does not recognize that changes may occur in the quality mix of the items compared. Fixler and Zieschang (1992) and Feenstra (1995) define an analogous hedonic COLI:

$$(3) \quad P_c(p_t, p_{t-1}, z_t, z_{t-1}, U) = \frac{E(p_t, z_t, U)}{E(p_{t-1}, z_{t-1}, U)}$$

that is, the ratio of the minimum expenditure required to maintain a given level of utility when the consumer faces  $p_t$  and  $p_{t-1}$  prices and quality characteristics  $z_t$  and  $z_{t-1}$ .

The construction of such indexes requires the existence of a representative consumer whose expenditure functions are defined over the space of “characteristics,” prices, and utility. When goods differ in their characteristics and consumers are heterogeneous in their preferences, only a specific class of functions describing the behavior of agents can be aggregated to some “representative” agent.

Theoretical frameworks are given by Feenstra (1995) and Diewert (chap. 10 in this volume). Feenstra uses aggregation results from McFadden (1983) to show that a representative agent formulation indeed arises from a discrete choice model, in which the individual consumers are deciding which of a discrete number of alternative varieties to choose. Feenstra proposes a reasonably broad class of utility functions for the individual consumers, which has two components: a subutility function over characteristics  $z_i \in R_+^K$ , which is the same across consumers; and an additive term obtained from each variety chosen, which differs across consumers. The latter additive terms are modeled as random across consumers, with a general “extreme value” distribution. Any pattern of correlation in the utility obtained from different models is allowed for, so this framework is much more general than the multinomial logit model, for example (in which the additive errors obtained from each variety are independent).

In this context, there exists an expenditure function for the representative consumer,  $E(p_t, z_t, U_t)$ , where  $p_t = (p_{1t}, \dots, p_{Nt})$  is the vector of prices for the  $N$  varieties, and  $z_t = (z_{1t}, \dots, z_{Nt})$  is the  $NK$ -dimensional vector of characteristics over all the product varieties. Social welfare,  $U_t$ , is interpreted as the sum of utilities over the individual consumers (i.e., utilitarian social welfare), and  $E(p_t, z_t, U_t)$ , measures the minimum expenditure summed over all

consumers to obtain  $U_t$ . For each variety, we can also define the marginal value of characteristics to consumers,  $\beta_i \in R_+^k$ , which is the same across consumers. As characteristics change over time, bounds for the exact index can be constructed using these values. The current (Paasche) period and base (Laspeyres) weighted quality-adjusted bounds for a COLI, for an arithmetic aggregation using a linear hedonic equation, are given by Feenstra (1995) as

$$(4a) \quad \left( \frac{\sum_{i=1}^N x_{it} p_{it}}{\sum_{i=1}^N x_{it} \hat{p}_{it}} \right) \leq \frac{E(p_t, z_t, U)}{E(p_{t-1}, z_{t-1}, U)} \leq \left( \frac{\sum_{i=1}^N x_{i,t-1} \hat{p}_{it}}{\sum_{i=1}^N x_{i,t-1} p_{i,t-1}} \right)$$

where  $E(\cdot)$  denotes the expenditure function, at periods  $t$  and  $t - 1$ , evaluated at a fixed level of utility, and the arguments in the index are given by

$$(4b) \quad \begin{aligned} \hat{p}_{it} &\equiv p_{it} - \sum \beta_{kt} (z_{ikt} - z_{ikt-1}) \\ \hat{p}_{i,t-1} &\equiv p_{i,t-1} + \sum \beta_{k,t-1} (z_{ikt} - z_{ikt-1}) \end{aligned}$$

where Laspeyres and Paasche in equation (4a) are upper and lower bounds on their “true”, economic theoretic COLIs:  $x$  is quantity sold,  $p$  is price, and  $z$  is a vector of characteristics with associated marginal values  $\beta_{kit}$  derived from a linear hedonic regressions over  $i = 1 \dots N$  product varieties (models) for each period  $t$ . Changes in the quality of models are picked up via changes in their characteristics ( $z_{kt} - z_{kt-1}$ ), which are multiplied by estimates of their associated marginal values  $\beta_{kt}$ . With sales data available, the vector  $z$  can be the sales-weighted average usage or mix of each characteristic in each period. Note that  $\hat{p}_{it}$  corrects the observed prices  $p_{it}$  for changes in the characteristics between the two periods, corresponding to the “explicit quality adjustment” described by Triplett (1990, 39).

Equation (4) has a simple intuition. In equation (4a) matched prices are being compared using current period quantities (weights) on the left-hand side and base period quantities on the right-hand side of the equation. However, the matching may not be perfect in that for each  $i$ , the quality may change over time. Consequently, predicted values are generated in equations (4b) to correct for such changes. For the left-hand side they adjust the base period prices for changes in the characteristics taking place between  $t - 1$  and  $t$ : maybe the goods are getting better over time. This change for each  $i$  is  $\Delta z_{ik} = (z_{ikt} - z_{ikt-1})$ , where  $z_{ikt}$  is the sales-weighted average of each  $k$  quality characteristic: say, the average spin speed or load capacity of washing machines has increased. However, some characteristics are more important, in a price-determining sense, than others, so each  $\Delta z$  is weighted by an estimate of its marginal value from a hedonic regression. For the left-hand side the hedonic regression is estimated using base period  $t - 1$  data to correct  $p_{t-1}$ , and for the right-hand side current period  $t$  data are used to correct  $p_t$ . It will be shown later how the  $i$  has to be defined in practice as product groups in which the quality mix changes but the intuition remains.

Economic theory provides further help with the choice between index number formulas. Cost-of-living index number formulas are defined in economic theory as exact for particular types of preferences if they equal the ratio of expenditure required to maintain constant utility for consumers with those types of preferences. Different index number formulas have been shown to have an exact correspondence to the functional form of the consumer's expenditure function. Laspeyres and Paasche price indexes (equation [4a]) correspond to fixed coefficient Leontief forms and act as upper and lower bounds on a true COLI. Base and current period weighted geometric means indexes could also be calculated, these being exact for (corresponding to) utility-maximizing consumers with constant elasticity of substitution (Feenstra 1995 and footnote 1). Diewert (1976, 1978) found that symmetric averages of these bounds provide index number formulas that correspond to flexible functional forms for the expenditure function, which are much less restrictive. He defined such index number formulas as being *superlative*. Fisher's index is the geometric mean of Laspeyres and Paasche and is superlative. The Törnqvist index<sup>1</sup> uses a symmetric mean of the weights of the bounds in equation (4a) and is superlative and exact for (corresponds to) a flexible translog utility function. Fisher's and Törnqvist indexes are thus quite special in that they are superlative, although Diewert (1995, 1997) has also shown the two formulas to be superior to many others from an axiomatic approach, with Fisher's in particular satisfying more "reasonable" tests than its competitors. The exposition here has been for arithmetic aggregation as opposed to a geometric one, although the results for geometric bounds and a Törnqvist index are noted in section 9.4 and are available from the authors.

The advantages of the exact hedonic approach are threefold. First, it utilizes the coefficients on the characteristics in an unconstrained manner to adjust observed prices for quality changes. Second, it incorporates a weighting system using data on the sales of each model and their characteristics, rather than treating each model as equally important. Finally, it has a direct correspondence to a constant utility index number formulation defined from theory.

### 9.2.3 Matching

We finally consider the process of matching. It compares the prices of matched identical varieties over time, so that the pure price changes are not tainted by quality changes. The aim is to compare only like with like. This is akin to the process used by price collectors for statistical offices in the

1. The geometric current and base-period bounds are given by  $\Pi_{i=1}^N (p_{it}/\hat{p}_{it-1})^{s_{it}} \leq [E(p_t, z_t, U)]/[E(p_{t-1}, z_{t-1}, U)] \leq \Pi_{i=1}^N (\hat{p}_{it}/p_{it-1})^{s_{it-1}}$ , where  $\hat{p}_{it-1} \equiv p_{it-1} \exp[\sum \beta_{ikt-1}(z_{ikt} - z_{ikt-1})]$  and  $\hat{p}_{it} \equiv p_{it} \exp[-\sum \beta_{ikt}(z_{ikt} - z_{ikt-1})]$ , and the Törnqvist index is given by  $\Pi(p_t/p_{t-1})^{(w_t + w_{t-1})/2} = [\Pi(p_t/p_{t-1})^{w_t} \times \Pi(p_t/p_{t-1})^{w_{t-1}}]^{1/2}$ , where  $w_t = (p_t x_t / \sum p_t x_t)$  and  $w_{t-1} = (p_{t-1} x_{t-1} / \sum p_{t-1} x_{t-1})$ .

**Table 9.1** Illustration of Matching and Approaches to Quality Adjustment

Outlet type	Variety	Weight	January	February	March	April
Multiple	1	$w_1$	$P_{11}$	$P_{12}$	$P_{13}$	$P_{14}$
	2	$w_2$	$P_{21}$	$P_{22}$		
	3				$P_{33}$	$P_{34}$
	4			$P_{42}$	$P_{43}$	$P_{44}$
Mass merchandiser	5	$w_5$	$P_{51}$	$P_{52}$	$P_{53}$	$P_{54}$
	6	$w_6$	$P_{61}$	$P_{62}$		
	7				$P_{73}$	$P_{74}$

compilation of CPIs, but the matching is electronic using scanner data. Scanner data have a code to describe each model of a good. The code can be extended to include the type of outlet in which it is sold, in order that a particular model of a good in a particular type of outlet is matched against its counterpart in successive periods. Since individual retailers often have unique codes for the same model, the matching is in practice closer than by “model and outlet type.” The problem with such matching is missing observations. For scanner data they arise when there is no transaction in that outlet (type) in a period, possibly because the item is no longer being sold or is on display but has not been bought.<sup>2</sup>

Turvey (1999a) proposed the use of chained matched indexes whereby the aggregate price change between, for example, January and February is spliced to that for February and March and so forth by successive multiplication. For example, the chained index between January and December,  $CI_{J,D}$ , is given by the product of the individual successive binary comparisons:

$$(5) \quad CI_{J,D} = I_{J,F} \times I_{F,M} \times I_{M,A} \cdots I_{N,D}$$

Table 9.1 illustrates the matching procedure. There are seven varieties each with assumed equal weights ( $w_i$ ), the distinction between outlet types being ignored for now.

The price index for January compared with February (J:F) involves price comparisons for varieties 1, 2, 5, and 6. For (F:M) it involves varieties 1, 4,

2. It is worth contrasting this with the way missing observations are recorded by price collectors. Price collectors may collect a display price even though the item is not sold in that particular month. Scanner data only pick up actual transactions. Alternatively, price collectors sampling from only some outlets may record a missing value if the model is not on display, even if the same model is being displayed and sold in other outlets. Scanner data match model numbers in types of outlets. Price collectors may not look at the number but use their own description of the main features of the item—for example, “a Bosch washing machine with 1400 spin speed”—which may be matched with a new or different model with similar, but not the same, characteristics. Price collectors match display prices of similar items from specific outlets; scanner data match unit values of all sales for identical items from types of outlets. Finally, the price collector has, within this context, an idea of replacement when a similar item is found to be almost taking the place of the old one. With scanner data this is something that can be explored, even automated, but it is not the subject of this study.

and 5, for (M:A) the same three varieties and, in addition, varieties 3 and 7. The sample composition changes for each comparison as varieties die and are born. The results for each comparison are chained to provide a single index for the whole period. Turvey advocates a chained geometric mean of matched observations.

Where wholly new products reflecting rapid technical improvement are introduced into a market, overlap price ratios between old and new products usually change from month to month. Instead of proceeding as above, arbitrarily selecting just one month's overlap price ratio between a replacement product and the replaced product, this procedure takes into account the ratio during all overlap months so that the prices of both the old and new products enter into the index computation. When new products arrive on the market their prices should be brought into the index, the prices of old products only being removed from it when they disappear from the market. Thus a chained geometric index of matched observations will be used with a sample size which varies through time. (Turvey 1999a, 13)

The method involves some loss of information. No use, for example, is made of  $p_{a2}$  for (J:F), and of  $p_{22}$ ,  $p_{33}$ ,  $p_{62}$ , and  $p_{73}$  for the (F:M) price comparisons, which is naturally to be regretted. This is to allow constant quality comparisons. However, it is on the birth and death of a product that price changes are unusual, and these are the very ones lost.<sup>3</sup> Some care, however, is needed in such statements. Table 9.1 illustrates how the loss of a matched observation takes place. For variety 2 in table 9.1,  $p_{22}$  is used for the January to February comparison, so it is not lost here. However, it is lost in the matching for the February to March comparison. It is tempting to argue that this loss is unimportant. It relates to a meaningless comparison because it does not exist in March and there is thus no basis for a price comparison. However, economic theory would assert otherwise. The economics of new goods is quite clear on the subject. If a new good is introduced, it is not sufficient to simply wait for two successive price quotations and then incorporate the good. This would ignore the welfare gain to consumers as they substitute from old technology to new technology. Such welfare gains are inseparably linked to the definition of a COLI defined as indexes, which measure the expenditure, required to maintain a constant level of utility (welfare). There exists in economic theory and practice the tools for the estimation of such effects (Hicks 1940; Diewert 1980, 498–503). This involves setting a “virtual” price in the period before introduction. This price is the one at which demand is set to zero. The virtual price is compared with the actual price in the period of introduction, and this is used to estimate the

3. A parallel issue arises for indexes of industrial production, especially in less developed countries, where new products are often new industries and ignoring their contribution to production when they are set up may seriously understate growth (Kmietowicz and Silver 1980).

welfare gain. Hausman (1997) provides some estimates for the introduction of a new brand of Apple-Cinnamon Cheerios. He concludes:

The correct economic approach to the evaluation of new goods has been known for over fifty years since Hicks’s pioneering contribution. However, it has not been implemented by government statistical agencies, perhaps because of its complications and data requirements. Data are now available. The impact of new goods on consumer welfare appears to be significant according to the demand estimates of this paper; the CPI for cereal may be too high by about 25 percent because it does not account for new cereal brands. An estimate this large seems worth worrying about.

Notwithstanding this, a curious feature of scanner data is that a missing transaction may arise simply because a model has not been purchased but is still on sale. A price collector would pick up the missing sales price not knowing it is above its reservation price, and there will thus be no need to estimate a virtual price.

### 9.2.4 Correspondence between the Methods

#### *Matched versus Exact Hedonic Indexes*

There is an interesting and useful correspondence here. Consider equation (4b). The  $\hat{p}_i$  is the price (or unit value) of model  $i$  (in a given outlet) in period  $t$ , having been adjusted by changes in its quality characteristics between period  $t - 1$  and  $t$ , the change in each characteristic being weighted by its associated marginal value in period  $t$ . If we are matching, no such adjustment is necessary. Matching does, however, have its failings in that information is lost. The exact hedonic formulation, as undertaken here in practice, aggregates not over each model, but over a subset of meaningful characteristics. For washing machines, for example, we might use makes and outlet types. The Laspeyres formulation on the right-hand side of equation (4) may be approximated by

$$(6a) \quad \frac{\sum_{j=1}^J x_{jt-1} \bar{p}_{jt}}{\sum_{j=1}^J x_{jt-1} \bar{p}_{j,t-1}}$$

where we define a narrow set of  $G_j$  characteristics—say, dummy variables for makes and outlet types—that are present in most models of the product in each period, where  $j = 1 \dots J$  combinations of makes and outlet types. The  $\bar{p}$  and  $x$  in equation (6a) are now the average prices and total quantities for each make in each outlet type, for example, for Zanussi washing machines sold in multiples, *within* a make and outlet type  $j$  for each period  $t$ . The adjusted average price for models  $i$  in each  $j$  in period  $t$  is given by

$$(6b) \quad \bar{p}_{jt} = \bar{p}_{jt} - \left[ \sum_{i \in G_j} \beta_{1i} \Delta \bar{z}_{j1t} + \sum_{i \in G_j} \beta_{2i} \Delta \bar{z}_{j2t} + \dots + \sum_{i \in G_j} \beta_{ki} \Delta \bar{z}_{jk t} \right]$$

where  $\bar{p}_j$  is the sales-weighted average price for each  $j$  make in a particular outlet type,  $\Delta \bar{z}_{jkt} = (\bar{z}_{jkt} - \bar{z}_{jkt-1})$ , where  $\bar{z}_{jkt}$  and  $\bar{z}_{jkt-1}$  are sales-weighted averages (in each period) of the  $k$  characteristics other than makes and outlets (e.g., spin speed) and  $\beta_k$  is their estimated marginal value. In equation (6a) for, say, group  $j = 1$ , Zanussi washing machines in multiples, the average price in period  $t$ ,  $\bar{p}_j$ , is compared with that in period  $t - 1$ ,  $\bar{p}_{j,t-1}$ . However, the quality of such machines may have changed over the period. This is adjusted for in equation (6b). For each  $j$ , say,  $j = 1$ , the quality-adjusted average price  $\bar{p}_j$  in period  $t$  is the (sales-weighted) mean price of varieties in that group minus (adjusted for) the change in quality. For example,  $\Delta z_{1t}$  might be the change in the (sales-weighted) average spin speed of washing machines for that make or outlet type, multiplied by an estimate in period  $t$ , of the marginal value of a unit of spin speed from an hedonic regression. The summation is over the  $i$  varieties that are members of the  $j = 1$  group. These adjustments continue for other quality characteristics  $z_2 \dots z_k$  in equation (6b). The  $\bar{p}_j$  are thus quality-adjusted within each of the groups being aggregated in equation (6a). The more quality characteristics we aggregate over in the body of equation (6a), the fewer characteristics are used in determining  $\bar{p}$  in (6b). Equation (6a) should collapse down to the matched method, when aggregating over all characteristics. So why restrict the aggregation in (6a) to only makes and outlet types? The answer is that in doing so we use all the data. On matching in table 9.1 we lost  $p_{33}$  for the February to March comparison but regained it for March to April.

Note that there is a minimal loss of information in the exact hedonic formulation given by equation (6) because each model has a make and outlet type. There is an efficiency gain to the estimate akin to that from stratified sampling in which we correct for quality changes by matching for price changes between strata and use estimates within strata. If we aggregated over all characteristics in equation (6a), with no adjustments in (6b) we would have a matching process with some models having no price data for either period  $t$  or  $t - 1$  in any two-way comparison. These would be excluded. However, when we allow the aggregation over a limited number of characteristics, which include models available in both periods, no information is lost. The adjustment for the variables not included in this weighted aggregation takes place in  $\bar{p}$ . There is a trade-off. The more quality variables in the weighted aggregation, the more chance of losing information. We consider this in the empirical section.

Both the exact hedonic and matching approaches allow all forms of weighting systems, including Laspeyres, Paasche, and the superlative Fisher, to be used to gain insights into such things as substitution effects. The exact hedonic formulation uses statistical estimates of product "worth" to partial out quality changes for the characteristics excluded in the aggregation in equation (6a), rather than the more computational, and accurate, matching. The differences between the methods depend on the reliability of

the  $\bar{p}$  adjustment process, in terms of both the extent of changes in characteristics ( $\Delta z$ ) and the values of  $\beta$ , and the relative loss of observations in bringing these characteristics into the aggregation process. Its extent is an empirical matter, and we will investigate this.

The equivalence of the two methods requires that the exact hedonic index take a chained formulation, as is the case for the matched approach. The chained approach has been justified as the natural discrete approximation to a theoretical Divisia index (Forsyth and Fowler 1981). Reinsdorf (1998) has formally determined the theoretical underpinnings of the index, concluding that in general chained indexes will be good approximations to the theoretical ideal—although they are prone to bias when prices changes “swerve and loop,” as Szulc (1983) has demonstrated (see also Haan and Opperdoes 1998).

#### *Direct versus Exact Hedonic*

We include in the analysis results of the time dummy variable method, given its use in many studies and the taking of such estimates as indicators of potential errors due to lack of quality adjustment in CPIs (Boskin 1996; Hoffman 1998). However, as argued in Silver (1999), it is but a limited form of the exact hedonic approach, the limitations naturally arising from the limited catalog data upon which the estimates are often based and the absence of sales weights. Consider the time dummy variable method if we, first, used weighted average prices on the left-hand side of equation (1) and a sales-weighted least squares estimator and, second, introduced dummy slope variables for each characteristic against time to allow for changing marginal values. The improved specification would require estimates of the change in quality-adjusted price change to be conditioned on the change in characteristics. If we take the value-weighted mean usage of each characteristic as the average usage upon which the change in quality-adjusted prices is conditioned, we have a framework akin to the exact hedonic one. Each of the modifications outlined above is just a relaxation of a restrictive assumption of the time dummy variable approach. We nonetheless include in this study estimates from the time dummy variable approach in order to identify the extent of errors arising from its conventional use.

### **9.3 Data and Implementation**

#### 9.3.1 Data

##### *Scope and Coverage*

The study is for monthly price indexes for washing machines in 1998 using scanner data. Scanner data are compiled on a monthly basis from the scanner (bar code) readings of retailers. The electronic records of just about every transaction include the transaction price, time of transaction, place of

sale, and a code for the item sold—for consumer durables we refer to this as the model number. Manufacturers provide information on the quality characteristics, including year of launch, of each model that can then be linked to the model number. Retailers are naturally interested in analyzing market share and pass on such data to market research agencies for analysis. By cumulating these records for all outlets (which are supplemented by visits to independent outlets without scanners) the agencies can provide comprehensive data on a monthly basis for each model for which there is a transaction, on the following: price (unit value), volume of sales, quality characteristics, make, and outlet type. Agencies are reluctant to provide separate data for a given model in a given outlet. This would not only allow competitors to identify how each outlet is pricing a particular model, and the resulting sales, but also allow manufacturers and governmental and other bodies to check on anticompetitive pricing. Data are, however, identifiable by broad types of outlets, and model codes often apply to specific outlets, although they are not identifiable.

It should be noted that the data, unlike those collected by price collectors, possess the following characteristics:

- They cover all time periods during the month.
- They capture the transaction price rather than the display price.
- They are not concerned with a limited number of “representative” items.
- They are not from a sample of outlets.
- They allow weighting systems to be used at an elementary level of aggregation.
- They include data on quality characteristics.
- They come in a readily usable electronic form with very slight potential for errors.

The data are not without problems, in that the treatment of multibuy and discounts varies between outlets and the coverage varies between product groups. For example, items such as cigarettes, which are sold in a variety of small kiosks, are problematic. Nonetheless, they provide a recognized alternative, first proposed by Diewert (1993) and used by Silver (1995) and Saglio (1995), but see also, for example, Lowe (1998) for Canada and Moulton, LaFleur, and Moses (1999) and Boskin (1996) for the United States. As Astin and Sellwood (1998, 297–98) note in the context of Harmonised Indices of Consumer Prices (HICP) for the European Union:

Eurostat attaches considerable importance to the possible use of scanner data for improving the comparability and reliability of HICPs ([European Union] Harmonised Indices of Consumer Prices), and will be encouraging studies to this end. Such studies might consider the various ways in which scanner data might be used to investigate different issues in the compilation of HICPs for example . . . provide independent estimates as a control or for detection of bias in HICP sub indices; . . .

analyse the impact of new items on the index; [and] carry out research on procedures for quality control.

Our observations (observed values) are for a model of the product in a given month in one of four different outlet types: multiples, mass merchandisers, independents, and catalog. We stress that we differentiate models as being sold in different types of outlets. This is a very rich formulation since it allows us to estimate, for example, the marginal value of a characteristic in a particular month and a particular type of outlet and apply this to changes in the usage of such stores. Not all makes are sold in each type of outlet. In January 1998, for example, there were 266 models of washing machines with 500 observations; that is, each model was sold on average in 1.88 types of outlets.

The coverage of the data is impressive in terms of both transactions and features. For the United Kingdom, for example, in 1998, there were 1.517 million transactions involving 7,750 observations (models or outlet types) worth £550 million. The coverage of outlets is estimated (by GfK Marketing Services) to be “well over 90%” with scanner data being supplemented by data from price collectors in outlets that do not possess bar code readers.

### *The Variables*

The variable set includes the following:

- *Price*: the unit value = value of sales or quantity sold of all transactions for a model in an outlet type in a month.
- *Volume*: the sum of the transactions during the period. Many of the models sold in any month have relatively low sales. Some only sell one of the model in a month or outlet type. Showrooms often have alongside the current models, with their relatively high sales, older models, which are being dumped, but need the space in the showroom to be seen. For example, 823 observations—models of washing machines in a month (on average) differentiated by outlet type—sold only one machine each in 1998. There were 1,684 observations (models in outlet types) selling between two and ten machines in a month, on average selling about 8,000 machines: so far, we have a total of 2,407 observations managing a sales volume of about 8,800. Yet the twelve models achieving a sales volume of 5,000 or more in any outlet or month accounted for 71,600 transactions.
- *Vintage*: the year in which the first transaction of the model took place. With durable goods, models are usually launched annually. The aim is to attract a price premium from consumers who are willing pay for the cachet of the new model, as well as to gain market share through any innovations that are part of the new model. New models can coexist with old models; 1.1787 million of the about 1.517 million washing machines sold in 1998 were first sold in 1997 or 1998—about 77.7 percent—leaving 22.3 percent of an earlier vintage coexisting in the market.

- *Makes*: the twenty-four different brands for which transactions occurred in 1998. The market was, however, relatively concentrated, with the three largest-selling (by volume) makes accounting for about 60 percent of the market. Hotpoint had a substantial 40 percent of sales volume in 1998. This was achieved with 15 percent of models (observations). Zannusi, Hoover, and Bosch followed with not unsubstantial sales of around 10 percent each by volume.

The *characteristics* set includes the following:

- *Type of machine* (out of five types): top-loader; twin tub; washing machine (WM; about 90 percent of transactions); washer dryer (WD) with computer; WD without computer;
- *Condensers*: with or without (for WD; about 10 percent with);
- *Drying capacity* of WD: a mean 3.15 kg and standard deviation of 8.2 kg for a standard cotton load;
- *Height* of machines in centimeters (cm)—(about 90 percent of observations being 85 cm tall);
- *Width and Depth* (94 percent being about 60cm wide; most observations taking depth values between 50 and 60 cm inclusive);
- *Spin speeds* of five main speeds: 800 rpm, 1,000 rpm, 1,100 rpm, 1,200 rpm, and 1,400 rpm (which account for 10 percent, 32 percent, 11 percent, 24 percent, and 7 percent, respectively, of the volume of sales);
- *Water consumption*, which is advertised on the displays as “not a measure of efficiency since it will vary according to the programme, wash-load and how the machine is used.” It is highly variable, with a mean of about 70 liters and standard deviation of 23liters;
- *Load capacity*, defined on the display as “a maximum load when loaded with cotton”—a mean about 50 kg with a standard deviation of about 13kg;
- *Energy consumption* (kWh per cycle), which is “based on a standard load for a 60 degree cotton cycle”—a mean of about 12 kWh with, again, a relatively large standard deviation of about 6 kWh.;
- *Free-standing*, built-under, and integrated; built-under, not integrated; built-in and integrated.
- *Outlet-type*: multiple, mass merchandiser, catalog, or independent.

### 9.3.2 Implementation of Each Method

The aim of this section is to compare the results of the three methods of measuring quality-adjusted price changes using scanner data.

#### *The Time Dummy Variable Approach*

Both linear and semilog formulations were considered. Results for the linear model are considered here, although those from a semilogarithmic model are referred to later and are available from the authors. The  $R^2$  for the respec-

tive forms were relatively high, at 0.83 and 0.82. A Box-Cox transformation was used for testing functional form, the estimated  $\lambda$  being 1.003 with  $SE(\lambda) = 0.024$  favoring the linear form.<sup>4</sup> A Bera-McAleer test based on artificial regressions was, however, inconclusive (the  $t$ -statistics for  $\theta_1$ ,  $\theta_2$  were 13.18 and 36.1, respectively).<sup>5</sup> The  $F$ -statistics for the null hypotheses of the results of coefficients all equal to zero were rejected for both functional forms at 314.6 and 297.2, respectively, for linear and semilog and  $p$ -values of 0.0000.

The results for the linear form are given in table 9.2. The coefficients were almost invariably statistically significant with appropriate signs and magnitudes. An additional spin speed of one rpm, for example, had a price premium of £0.30. Between December and January, other quality characteristics held constant and prices fell from the mean of £405.71 by £25.33, to £380.38—that is, a fall of 6.2 percent. There was some evidence of heteroskedasticity with the Breusch-Pagan test statistics of 27.7 and 9.0 for the linear and semilog forms respectively, both exceeding the critical value of Chi-squared (3 degrees of freedom) = 7.815. However, the estimator remains unbiased, and the standard errors were adjusted to be heteroskedastic-consistent using a procedure by White (1980).

The regressions were estimated on a data set that excluded models with sales of thirty or less in any month and a minimal number of models with extreme prices arising from variables not included in the data, such as stainless steel washing machines. A failing of the dummy variable approach is that models with only one transaction are given the same importance in the regression as a model with, say, 10,000 transactions. The choice of thirty was based on some experimentation. The loss in the number of observations was quite severe for washing machines from 7,750 to 3,600, whereas the loss in terms of the *volume* of sales was minimal, from 1.517 million to 1.482 million. An alternative approach is to use a sales volume weighted least squares estimator, as considered later.

### *Exact (and Superlative) Hedonic Indexes*

First, it is necessary to decide which quality-related variables are used in the aggregation in equation (6a) and which for the adjustment in equation

4. With this approach a variable  $Z$  is transformed to  $(Z^\lambda - 1)/\lambda$ . Since the limit of this as  $\lambda$  approaches zero is  $\log Z$ , it is defined to be  $\log Z$  when  $\lambda = 0$ . If all variables in a linear functional form are transformed in this way and then  $\lambda$  is estimated (in conjunction with the other parameters) via a maximum likelihood technique, significance tests can be performed on  $\lambda$  to check for special cases. If  $\lambda = 0$ , for example, the functional form becomes Cobb-Douglas in nature; if  $\lambda = 1$  it is linear. A confidence interval on  $\lambda$  can be used to test whether or not it encompasses 0 or 1.

5. The Bera-McAleer test involves obtaining predicted values  $\log(\hat{z})$  and  $(\hat{z})$  from a semi-logarithmic and a linear formulation, respectively. Artificial regressions are then computed using  $\exp\{\log(\hat{z})\}$  and  $\log(\hat{z})$  on the left-hand-side, and the residuals from each of these regressions,  $\hat{v}_1$  and  $\hat{v}_0$ , are included in a further set of artificial regressions:  $\log(z) = \beta_0 + \beta_1 X_1 + \mu_0 \hat{v}_1 + \epsilon_1$ ;  $z = \beta_0 + \beta_1 X_1 + \mu_1 \hat{v}_0 + \epsilon_2$ . Using  $t$ -tests, if  $\theta_0$  is accepted we choose the log-linear model, and if  $\theta_1$  is accepted we choose the linear model, the test being inconclusive if both are rejected or both accepted.

Table 9.2

**Regression Results for Linear Hedonic Equation for 1998**  
(dependent variable: price)

Variable	Estimated Coefficient	Standard Error
C	-981.290***	115.562
Months (benchmark: January)		
February	-0.214244	5.64587
March	6.71874	5.37602
April	0.220101	5.36765
May	-1.60710	5.37089
June	-5.02523	5.2291
July	-10.9314*	5.2195
August	-10.6569*	5.2936
September	-15.6103**	5.2193
October	-16.2413**	5.2970
November	-17.9056***	5.3397
December	-25.3343***	5.3068
Characteristics		
Height (cm)	-0.347469	0.291905
Depth (cm)	6.12143***	0.47413
Width (cm)	6.26849***	0.69974
Water consumption (liters)	-1.14070***	0.06825
Load capacity (kg)	-0.287457**	0.096536
Spin speed (rpm)	0.304251***	0.006965
Drying capacity—washer/dryer (kg)	-0.335703	0.368737
Condensor—washer/dryer	35.9352***	6.2530
Energy consumption (kWh per cycle)	0.331592**	0.104301
Vintage	4.24294***	0.90397
Type of machine (benchmark: front loader washing machine)		
Top loader	228.876***	12.125
Twin tub	-704.998***	20.251
Washer/dryer	64.6312***	9.1877
Washing machine with computer	127.455***	8.856
Washer/dryer with computer	129.682***	16.406
Installation (benchmark: free-standing)		
Built-under integrated	238.908***	10.1389
Built-under	-61.3298	42.8550
Built-in integrated	293.221***	27.349
Outlet type (benchmark: multiples)		
Mass merchandisers	23.1155***	2.9010
Independents	31.2458***	2.5882
Catalogs	71.8786***	3.3023
Makes (benchmark: Bosch)		
AEG	66.8428***	6.1338
Siemens	46.9125***	7.2096
Hoover	-68.0069***	3.65656
Miele	165.895***	10.316
Candy	-98.8340***	6.0759
English Electric	7.99810***	0.82048
Ariston	-21.9183*	8.5417
New Pol	-113.529	60.062

**Table 9.2** (continued)

Variable	Estimated Coefficient	Standard Error
Beko	-134.695***	10.558
Zanussi	5.16116	4.12916
Electro	0.7362	11.8086
Indesit	-68.7762***	5.4285
Neff	109.284***	16.897
Philco	-108.286***	25.939
Ignis	-22.4469***	22.9162
Creda	-67.3200***	7.4762
Tricity/Bendi	-58.2687***	6.1059
Hotpoint	-32.5816***	3.9348
Servis	-76.1764***	5.7801
Asko	164.781**	60.226
$N^a$	3,600	
Mean of price	405.713	
Standard error of regression	59.8572	
Adjusted $R^2$	0.8299	
Breusch-Pagan	27.7	
$F$ -statistic (zero slopes)	314.626	

<sup>a</sup>Volume of sales greater than thirty in a month or outlet type.

\*\*\*Statistically significant at the 0.1 percent level for two-tailed  $t$ -tests.

\*\*Statistically significant at the 1 percent level for two-tailed  $t$ -tests.

\*Statistically significant at the 5 percent level for two-tailed  $t$ -tests.

(6b). The  $\beta_i$  estimates are then derived from *monthly* hedonic regressions and multiplied by changes in the sales-weighted change in the mix of quality characteristics to provide an adjustment to average prices ( $\bar{p}$ ) for use in the main body of equation (6a).

The  $\beta$  coefficient is required for each  $K$  quality-related variable in equation (6b) *in each month*. The specification of the regression equations estimated for this purpose used all variables, to avoid omitted variable bias, with only the relevant  $\beta_i$  coefficient being used to generate  $\bar{p}$ . The specifications were therefore similar to the time dummy variable method except that separate regressions were estimated for each month.<sup>6</sup>

A weighted least squares estimator was used, the weights being the volume of sales. The mean  $\bar{R}^2$  over the twelve monthly hedonic regressions was 0.842.<sup>7</sup> The coefficients were almost invariably statistically significant and

6. It is noted that observations with sales of 30 and less were not used for estimating the individual coefficients, but all the data were used for the average prices, quantities, values, and sales-weighted mix of qualities in formulas.

7. Test statistics here are illustrative, being based on semilog and linear models for the data as a whole, although they are indicative of the results for individual months (available from the authors). Estimates for log-log models were not feasible, given the large number of dummy variables on the right-hand side of the equation.

of reasonable magnitude with the appropriate sign.<sup>8</sup> As with the dummy variable method, the regressions were estimated using the linear and semi-log forms. The coefficients from the linear form were used to derive quality-adjusted prices for use in an arithmetic framework—that is, for Laspeyres, Paasche, and (superlative) Fisher hedonic indexes (equation [4]). The coefficients from the semilog form were used to calculate base and current period weighted geometric means and (superlative) Törnqvist exact hedonic indexes given in footnote 1 and available from the authors. It is noted that the estimation of semilogarithmic functions as transformed linear regressions requires an adjustment to provide minimum variance unbiased estimates of parameters of the conditional mean. This involved the standard errors, which in any event were very small, although the adjustments were undertaken (Goldberger 1968).

As explained in section 9.2, the exact hedonic approach has the advantage over matching of minimal loss of data. However, the more variables included in the aggregation in equation (6a), the greater the information loss, as either  $\bar{p}_{it-1}$  or  $\bar{p}_{it}$  becomes unavailable in any period for comparison. In the limiting case of all variables being included, the method collapses to the matched approach. If we aggregate over makes, or even makes and outlet types, there is very little loss of data in terms of the number of observations and volume of sales. Aggregating only over the 21 makes leaves 99.67 percent of observations and 99.97 percent of sales volume. Extending the aggregation to the 21 makes and 4 outlets, 84 combinations still has little loss of data—99.08 percent of observations and 99.92 percent of sales volume. Any manufacturer operating in a particular outlet type continues to do so on a monthly basis. Extending the aggregation further to 24 spin speeds (i.e., over 2,016 combinations) reduces the coverage to 95.9 percent of observations and 99.6 percent of sales volume.

### *Matching*

The matching procedure used incurred further loss of data: Only 83 percent of observations were used, although the missing ones were models in outlets that were being discarded with low sales, the volume of sales used in the matching being 97.8 percent.

The extent of the matching is illustrated for washing machines in 1998 in table 9.3. There were for example, 429 matched comparisons of a particular model in a specific outlet type in February 1998. These were selected from 500 and 488 observations available in February and January 1998, respectively. In total there were 6,020 matched comparisons for 1998, which compares with 7,750 available in 1998 or, more fairly,  $7,750 - 500 = 7,256$  to exclude the January figures because the matched comparisons are over eleven monthly comparisons as opposed to twelve months' data.

8. Results are available from the authors.

**Table 9.3** Data on Matching for Washing Machines, 1998

Matched	Number of Observations		Volume of Sales (thousands)	
	Unmatched	Matched	Unmatched	Matched
January	500		126.2	
February	488	429	111.4	115.1
March	605	425	134.0	118.6
April	625	510	113.3	120.7
May	647	527	112.5	111.3
June	711	555	137.1	122.5
July	744	620	116.3	124.9
August	711	627	123.0	118.3
September	717	606	150.4	135.1
October	695	602	129.2	138.5
November	643	566	124.8	125.8
December	664	553	138.6	129.8

*Note:* Matched comparisons are between each month and the preceding one for chained indexes. There were, for example, 429 matched comparisons between February and January 1998 taken from 488 observations (model in a specific outlet type) in February and 500 in January 1998. Similarly, there were 425 matched comparisons between March and February 1998. Although the number of matched observations will not exceed those unmatched, the volume of sales may do so.

This difference of 1,236 observations is price data that exist in either period  $t$  or period  $t + 1$  but do not have a counterpart to enable a comparison.<sup>9</sup> Since they are models just born or about to die, they should have low sales, and thus their omission should not unduly affect the index.<sup>10</sup> The total sales volume of matched comparisons was 1.3605 million, compared with 1.3906 million (unmatched but excluding January)—a difference of about 30,000 sales or about 2 percent of sales. From table 9.3 the monthly variation can be deduced. The worst loss of information was in the March to February comparisons: from  $(111.4 + 134.0)/2 = 122.7$  thousand to 118.6 thousand—a loss of 3.3 percent. For the September to October and October to November comparisons the losses were less than 1 percent. A unit value index is given by

$$(7) \quad \left[ \frac{\sum p_t x_t}{\sum x_t} \right] / \left[ \frac{\sum p_{t-1} x_{t-1}}{\sum x_{t-1}} \right],$$

which is a weighted measure of price changes *not* adjusted for changes in the quality mix. It is included in the analysis for comparison.

9. If, for example, the matched item had sales in period  $t$  of 100 and in period  $t + 1$  of 50, and a new model was launched in period  $t + 1$  with sales of 10, the matched volume would be  $150/2 = 75$  and the unmatched 60 in period  $t + 1$ .

10. The data are transactions over the month, so recently born or dead models may only have been available for part of the month in question and have relatively low sales.

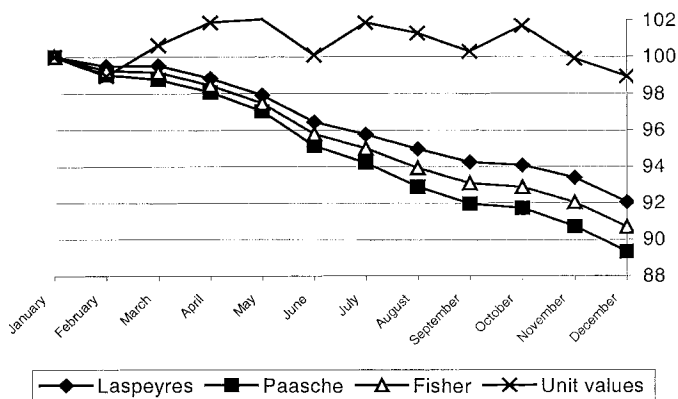
### 9.4 Results on the Three Methods for Scanner Data

Table 9.4 and figure 9.1 provide results using the matched approach for several different formulas.

First, the Laspeyres and Paasche provide outer upper and lower bounds, respectively, to the superlative Fisher index, the extent of the substitution being about 1.35 percent over the year, as consumers substituted away from machines with relatively high price increases. Note that because these indexes are chained on a monthly basis, they are different from those that arise from their fixed-base index counterparts. They allow the basket to be updated each month, the substitution in each month being compounded over the year. Second, figure 9.1 and table 9.4 show the unit value index, defined by index equation (7), to be unaffected by changes in the quality mix of models. Although the index shows only a slight overall fall in prices over

**Table 9.4** Matched Quality-Adjusted Price Indexes by Formulas

	Laspeyres	Paasche	Fisher	Unit Values
January	100	100	100	100
February	99.49	99.02	99.25	98.96
March	99.52	98.77	99.15	100.61
April	98.82	98.08	98.45	101.86
May	97.91	97.05	97.48	102.04
June	96.46	95.13	95.79	100.09
July	95.78	94.21	94.99	101.86
August	94.97	92.89	93.92	101.28
September	94.24	91.95	93.09	100.28
October	94.06	91.71	92.88	101.67
November	93.39	90.70	92.03	99.89
December	92.06	89.34	90.69	98.93



**Fig. 9.1** Washing machines: Alternative formulas

the year of about 1 percent, and increases in other months compared with January 1998, quality-adjusted price changes have fallen by just under 10 percent over the year. The superlative matched index effectively adjusts for changes in the quality mix of purchases being based on computational matching as opposed to statistical models. Finally, not reported here but available from the authors, the results for the superlative Törnqvist index in footnote 1 are similar to those for the superlative Fisher, the Törnqvist in December being 90.711, compared to Fisher's 90.69. The geometric base and current-period weighted indexes can be seen in footnote 1 to be expected to be upper and lower inner bounds, respectively, on the superlative index because they incorporate some substitution effect (Shapiro and Wilcox 1997)—again, results are available from the authors. All of this is as predicted by economic theory. These indexes are the results of matched computations with different weighting systems. The matching, however, loses 2 percent of the data by sales volume. We consider below exact (and superlative) hedonic indexes, which lose only 0.4 percent of sales volume.

Figure 9.2 and table 9.5 provide results for different approaches to measuring quality-adjusted price indexes using an arithmetic formulation. The estimates from a linear model using the time dummy variable approach show a fall of only 6.0 percent, falling, in December, outside the Laspeyres and (what would be) Paasche bounds of the matched and exact approaches. In section 9.3 we found the hedonic regression to have a relatively high  $\bar{R}^2$  with signs and values of the coefficients being as expected on a priori grounds. The linear formulation was supported by a Box-Cox test, although the results from a semilog formulation are very similar, the index falling to 93.85—by 6.15 percent. By conventional standards, these estimates are quite acceptable. The difference between the results from other approaches is more likely to be a result of the absence of a weighting system for the time dummy variable approach. If prices of more popular models are falling faster than those of unpopular ones, the weighted matched and exact approaches will take this into account, whereas the time dummy variable method will not. The concern here is with the time dummy variable approach as it is usually employed. However, a sales-weighted least squares estimator should in principle bring these estimates closer to the exact hedonic and matched results, although in practice the ordinary least squares and weighted least squares results were quite similar, a fall of 6.0 percent and 5.5 percent, respectively (and for semilog 6.0 percent and 5.7 percent, respectively). The results from the exact (and superlative) hedonic approach and matched estimates are not too dissimilar, a difference of about 2.0 and 1.7 percentage points for Laspeyres and Fisher over the year. In this case study, the loss of data for the matching, at about 2 percent by volume, was relatively low, giving confidence to the matched results.

Finally, figure 9.3 shows the results for the exact (superlative) hedonic approach at different levels of aggregation for Fisher indexes. As we expand

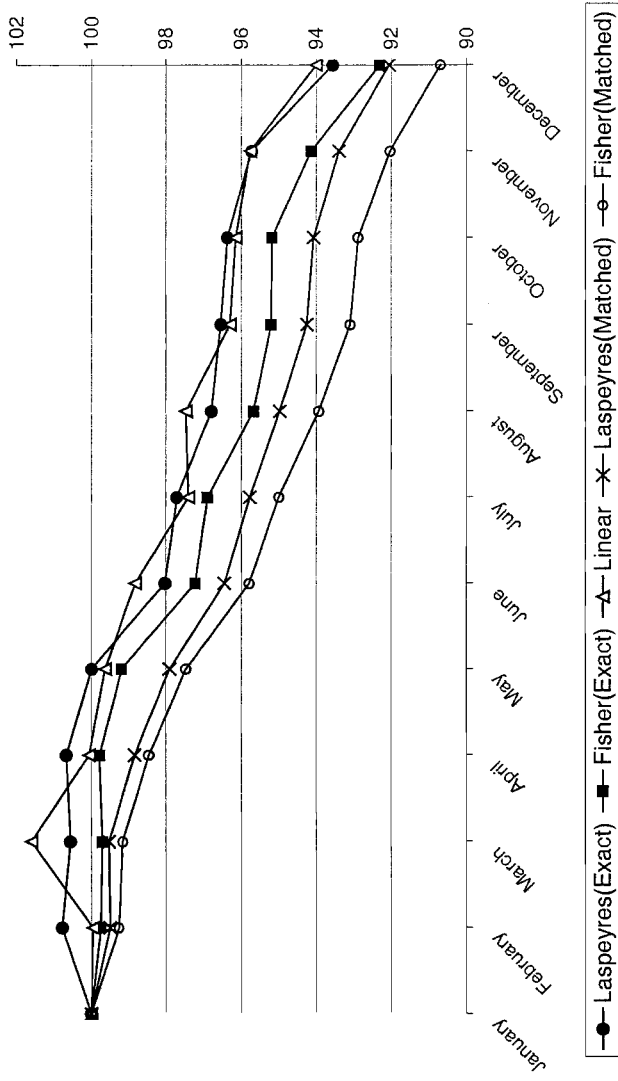
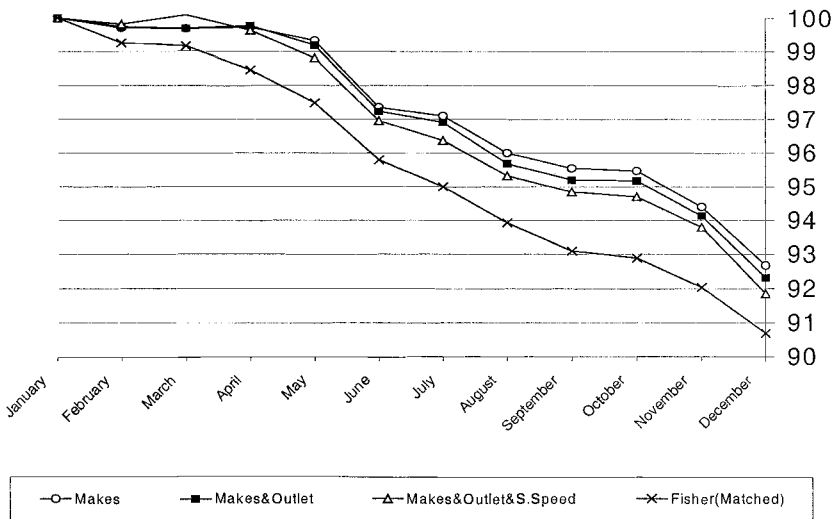


Fig. 9.2 Quality-adjusted price indexes using arithmetic means

**Table 9.5** Quality-Adjusted Price Indexes Based on Arithmetic Means

	Exact Hedonic by Make and Store Type		Time Dummy Variable: Linear	Matched	
	Laspeyres	Fisher		Laspeyres	Fisher
January	100	100	100	100	100
February	101.04	99.93	99.95	99.50	99.25
March	101.26	100.23	101.59	99.52	99.15
April	100.99	99.86	100.05	98.82	98.45
May	100.25	99.11	99.62	97.91	97.48
June	98.41	97.27	98.81	96.46	95.79
July	98.19	97.02	97.41	95.78	94.99
August	97.10	95.61	97.47	94.97	93.92
September	96.96	95.24	96.30	94.24	93.09
October	97.05	95.23	96.15	94.06	92.88
November	96.14	94.18	95.76	93.39	92.03
December	94.12	92.42	93.99	92.06	90.69



**Fig. 9.3** Exact hedonic indexes at different levels of aggregation

the weighted price changes in the body of the formula in equation (6a) from makes to makes within each outlet type, and then further by spin speed, the exact (superlative) hedonic index approaches the matched index.

In summary, this part of the paper uses scanner data to show how to measure quality-adjusted price changes using scanner data. It casts doubt on the use of the time dummy variable approach. It also argues for a matched approach as a special case of the theoretically based exact hedonic approach, the matched approach being based on computational matching and

not being subject to the ideosyncrasies of the econometric estimation of hedonic indexes (Griliches 1990; Triplett 1990). Caution is, however, advised when the loss of data in matching is severe. In such a case an empirical investigation into the trade-off between including variables in the aggregation and the resulting loss of data is advised.

## **9.5 Quality Adjustment and Consumer Price Index Practice: An Experiment**

### **9.5.1 Alternative Methods**

The above account concerned the measurement of quality-adjusted price changes using scanner data. However, the problems of quality adjustment for the practical compilations of CPIs by statistical offices are quite different. In general, display prices are recorded by price collectors on a monthly basis for matched product varieties in a sample of individual stores. When a variety is missing in a month, the replacement may be of a different quality, and “like” may no longer be compared with “like.” Its quality-adjusted price change has to be imputed by the statistical office. There is no problem of quality adjustment when varieties are matched. It is only when one is unavailable and its price change has to be imputed that there is a problem. The purpose of this experiment is to attempt to replicate the practices used by statistical offices in CPI compilation in order that the effects of different quality adjustment techniques can be simulated.

It should be noted that this is not a trivial matter. Moulton, LaFleur, and Moses (1999) examined the extent to which price collectors were faced with unavailable varieties of TVs in the U.S. CPI. Between 1993 and 1997, 10,553 prices on TVs were used, of which 1,614 (15 percent) were replacements, of which, in turn, 680 (42 percent) were judged to be not directly comparable. Canadian experience for TVs over an almost identical period found 750 of the 10,050 (7.5 percent) to be replacements, of which 572 (76 percent) were judged to be not directly comparable. For international price comparisons the problem is much more severe (Feenstra and Diewert 2000).

We should stress that sections 9.2–9.4 were concerned with how best to measure quality-adjusted price changes using scanner data. Here, the use of scanner data is to simulate CPI practices to help judge the veracity of alternative quality adjustment procedures that might be employed to supplement matched models procedures. A number of well-documented options are available and are outlined in Turvey (1989); Reinsdorf, Liegey, and Stewart (1995); Moulton and Moses (1997); Armknecht, Lane, and Stewart (1997); Armknecht and Maitland-Smith (1999); and Moulton, LaFleur, and Moses (1999), although the terminology differs; these options include

- *Imputation.* Where no information is available to allow reasonable estimates to be made of the effect on price of a quality change, the price change in the elementary aggregate group as a whole, to which the variety belongs, is assumed to be the same as that for the variety.
- *Direct comparison.* If another variety is directly comparable, that is, if it is so similar it can be assumed to have the same base price, its price replaces the missing price. Any difference in price level between the new and old is assumed to be due to price changes and not quality differences.
- *Direct quality adjustment.* Where there is a substantial difference in the quality of the missing and replacement varieties, estimates of the quality differences are made to enable quality-adjusted price comparisons to be made.

For illustration, consider a product variety with a price of £100 in January. In October, a replacement version with a widget attached is priced at £115. The direct comparison would use the price of an essentially identical variety. The direct quality adjustment method requires an estimate of the worth of the widget. For example, if it was found that the widget increased the product's flow of services by 5 percent, the £115 in October could be compared with an adjusted base-period price of £100  $(1.05) = £105$ . The imputation approach would use the index for the relevant product group. If this was 110.0 in October (January = 100.0), the replacement item would have a revised base-period price of  $£115/1.10 = £104.55$  to compare with the new price of £115; that is,  $115/104.55 \times 100 = 110.0$ , a 10 percent increase in price, the residual 5 percent being assumed to be due to quality differences.

Moulton, LaFleur, and Moses (1999) note that the direct or explicit quality adjustment approach can be used with data from manufacturers on the cost of quality changes or coefficients from hedonic regressions as explicit estimates of quality differences. Valuable insights into the validity of such methods can be gained by using a number of such methods on actual CPI data and comparing the results; seminal work in this area includes Lowe (1999) and Moulton, LaFleur, and Moses (1999). The alternative approach adopted here is to attempt to replicate such procedures using scanner data.

### 9.5.2 The Experiment

The purpose of this experiment is to replicate CPI data collection using scanner data to provide a means by which different CPI procedures can be emulated. The formulation here is relatively crude, being an initial attempt at the exercise. However, we hope it will be useful for illustrative purposes. The data are the same monthly data on washing machines in 1998 used in sections 9.3 and 9.4. We start by taking a January fixed basket of washing

machines comprising all varieties for which there was a transaction in January. Our varieties are for a model in one of four outlet types: multiples, mass merchandisers, catalogs, and independents. Since many models are only sold in chains of particular outlets, the classification is in practice closer to a given model in a specific chain or even individual outlet, which is the price observed by a price collector. The unit value of each variety in January is treated as the average display price collected by the price collectors. Since the volume of transactions is known for each variety, the January sample is taken to be the universe of every transaction of each variety. This January universe is the base-period active sample. We can, of course, subsequently modify this by using different sampling procedures and identify their effects on the index.

If the variety in each outlet type continues to exist over the remaining months of the year, matched comparisons are undertaken between the January prices and their counterparts in successive months. Consider again for illustration table 9.1, the case of four varieties existing in January, each with relative expenditures of  $w_1$ ,  $w_2$ ,  $w_5$ , and  $w_6$  and prices of  $p_{11}$ ,  $p_{21}$ ,  $p_{51}$ , and  $p_{61}$ . A Laspeyres price index for February compared with January = 100.0 is straightforward. In March the prices for varieties 2 and 6 are missing. Each of these was collected from different outlet types, multiples and mass merchandisers in this example. To enable Laspeyres price comparisons to be undertaken in such instances, a range of methods was utilized for the scanner data, including the following:

1. *Implicit imputation.* Price comparisons were only used when January prices could be matched with the month in question. In our example the January to March comparisons were based on varieties 1 and 5 only, the price changes of varieties 2 and 6 being assumed to be the same as these remaining varieties. The weights for varieties 1 and 5 would be  $w_1/(w_1 + w_5)$  and  $w_5/(w_1 + w_5)$ , respectively.

2. *Targeted implicit imputation.* The price changes of missing varieties for a specific make within an outlet type were assumed to be the same as for the remaining active sample for that make within its outlet type. If we assumed varieties 1 and 2 and 5 and 6 are of the same make, then the weights for varieties 1 and 5 would be  $(w_1 + w_2)/W$  and  $(w_5 + w_6)/W$ , respectively, where  $W = (w_1 + w_2 + w_5 + w_6)$ .

3. *Direct comparison.* Within each outlet type a search was made for the best match first, by matching brand, then in turn by type (see earlier section "The Variables"), width, and spin speed. If more than one variety was found, the selection was according to the highest value of transactions (expenditure). In our example, varieties 3 or 4 and 7 would, respectively, replace varieties 2 and 6 in March.

4. *Explicit hedonic: predicted versus actual.* A hedonic regression of the

(log of the) price of model  $i$  in period  $t$  on its characteristics set  $z_{iki}$  was estimated for each month, given by

$$(8) \quad \ln p_{it} = \beta_{0t} + \sum_{k=1}^K \beta_{kt} z_{kit} + \varepsilon_{it}.$$

Say the price of variety  $m$  goes missing in March, period  $t + 2$ . The price of variety  $m$  can be predicted for March if we insert the characteristics of variety  $m$  into the estimated regression equation for March and similarly for successive months. The predicted price for this “old” unavailable variety  $m$  in March and its price comparison with January (period  $t$ ) are respectively given by

$$(9) \quad \hat{p}_{m,t+2} = \exp \left[ \hat{\beta}_{0,t+2} + \sum \hat{\beta}_{k,t+2} z_{k,m} \right] \quad \text{for} \quad \frac{\hat{p}_{m,t+2}}{p_{m,t}} - \text{old}$$

The “old” denotes that the comparison is based on a prediction of the price of the unavailable variety in the current period rather than the (new) replacement variety’s price in the base period. In our example we would estimate  $\hat{p}_{23}$ ,  $\hat{p}_{24}$ , and so on, and  $\hat{p}_{63}$ ,  $\hat{p}_{64}$ , and so on, and compare them with  $p_{21}$  and  $p_{61}$ , respectively. We would effectively fill in the blanks for varieties 2 and 6.

An alternative procedure is to select for each missing  $m$  variety a replacement  $n$  variety using the routine described in step (3) above. In this case the price of  $n$  in period  $t + 2$ , for example, is known, and we require a predicted price for  $n$  in period  $t$ . The predicted price for the “new” variety and required price comparison are

$$(10) \quad \hat{p}_{n,t} = \exp \left[ \hat{\beta}_{0,t} + \sum \hat{\beta}_{k,t} z_{k,n} \right] \quad \text{for} \quad \frac{p_{n,t+2}}{\hat{p}_{n,t}} - \text{new},$$

that is, the characteristics of variety  $n$  are inserted into the right-hand side of an estimated regression for period  $t$ . The price comparisons of equation (9) would be weighted by  $w_{m,t}$ , as would those of its replaced price comparison in equation (10).

A final alternative is to take the geometric mean of the formulations in equations (9) and (10) on grounds akin to those discussed by Diewert (1997) for similar index number issues.

5. *Explicit hedonic: predicted versus predicted.* A further approach was the use of predicted values for, say, variety  $n$  in *both* periods, for example,  $\hat{p}_{n,t+2}/\hat{p}_{n,t}$ . Consider a misspecification problem in the hedonic equation. For example, there may be an interaction effect between a brand dummy and a characteristic—say, a Sony television set and Nicam stereo sound. Possession of both characteristics may be worth 5 percent more on price (from a semilogarithmic form) than their separate individual components (for evidence of interaction effects, see Curry, Morgan, and Silver 2001). The use of

$p_{n,t+2}/\hat{p}_{n,t}$  would be misleading since the actual price in the numerator would incorporate the 5 percent premium, whereas the one predicted from a straightforward semilogarithmic form would not. A more realistic approach to this issue might be to use predicted values for *both* periods. We stress that in adopting this approach we are substituting for a recorded, actual price an imputation. This is not desirable, but neither would be the form of bias discussed above.

The comparisons using predicted values in both periods are given as

$$(11) \quad \frac{\hat{p}_{n,t+2}}{\hat{p}_{n,t}} \text{ for the "new" variety}$$

$$\frac{\hat{p}_{m,t+2}}{\hat{p}_{m,t}} \text{ for the disappearing or "old" variety, or}$$

$$\left[ \left( \frac{\hat{p}_{n,t+2}}{\hat{p}_{n,t}} \right) \left( \frac{\hat{p}_{m,t+2}}{\hat{p}_{m,t}} \right) \right]^{0.5}$$

as a (geometric) mean of the two.

6. *Explicit hedonic: adjustments using coefficients.* In this approach, a replacement variety was found using the routine in step (3) above and any differences between the characteristics of the replacement  $n$  in, for example,  $t + 2$  and  $m$  in period  $t$  ascertained. A predicted price for  $n$  in period  $t$ , that is,  $\hat{p}_{n,t}$ , was determined and compared with the actual  $p_{n,t+2}$ . However, unlike the formulation in equation (10),  $\hat{p}_{n,t}$  was estimated by applying the subset of the  $k$  characteristics that distinguished  $m$  from  $n$  to their respective marginal values in period  $t$  estimated from the hedonic regression and adjusting the price of  $p_{m,t}$ . For our illustration, if the nearest replacement for variety 2 was variety 3, then the characteristics that differentiate variety 3 from variety 2 were identified, and the price in the base period,  $p_{31}$  is estimated by adjusting  $p_{21}$  using the appropriate coefficients from the hedonic regression in that month. For example, if variety 2 had an 800 rpm spin speed and variety 3 a 1,100 rpm spin speed, other things being equal, the marginal value of the 300 rpm differential would be estimated from the hedonic regression, and  $p_{21}$  would be adjusted for comparison with  $p_{33}$ . Note that if the  $z$  variables in the characteristic set are orthogonal to each other the results from this approach will be identical to those from equation (10). A similar approach to equation (9) was also undertaken that only used the salient distinguishing characteristics and the geometric mean of the two calculated. These six methods provide in all twelve different measures of quality-adjusted price changes.

### 9.5.3 The Study

Table 9.6 provides a summary of the data used. In January 1998 there were 500 varieties (models in one of the four outlet types—multiples, mass merchandisers, catalogues, and independents) of washing machines accounting

**Table 9.6** Number of Observations and Expenditure Shares Used by Different Methods

	Number of Observations, Models in Outlets	Number of Transactions, Thousands	Expenditure, £1000s	Imputation—No Replacements				Direct Comparisons—Replacements			
				% of January Observations Used		% of January Expenditure Used		% of January Observations Used		% of January Expenditure Used	
January	500	126,171	45,727								
February	488	111,358	39,938	85.8	97.19	97.06	99.8	99.999			
March	605	134,049	48,877	83.0	99.13	91.06	99.2	99.991			
April	625	113,338	41,841	81.0	98.65	81.79	99.4	99.996			
May	647	112,549	41,623	77.6	98.32	76.62	98.2	99.899			
June	711	137,070	49,721	78.6	97.97	72.90	99.2	99.992			
July	744	116,273	42,924	76.0	97.05	64.21	99.0	99.990			
August	711	122,977	45,139	69.6	93.52	57.51	98.8	99.932			
September	717	150,422	54,665	66.2	86.41	54.24	99.0	99.935			
October	695	129,235	47,620	59.8	87.13	51.63	98.0	99.910			
November	643	124,845	45,198	54.0	83.51	49.35	97.0	99.788			
December	664	138,634	49,707	53.0	81.58	48.16	97.6	99.818			

for 126,171 transactions. The distribution was highly skewed, with the top 5 percent and 10 percent of varieties (in an outlet type) accounting for 49 percent and 66 percent of transactions, respectively, in January. The indexes in this section are fixed-base January 1998 = 100.0 indexes compiled over the period January to December 1998. Unlike the chained indexes considered in previous sections, they take no account of varieties introduced after January unless they replace a variety that is “missing”—that is, it no longer has any transactions. The imputation approach does not replace missing varieties. Table 9.6 shows that by December, only 53 percent of the January basket of varieties were used for the December/January index, although these accounted for 81.6 percent of January expenditure. Varieties with lower sales values dropped out more quickly. However, the remaining  $0.53(500) = 265$  varieties in December only accounted for 48.2 percent of the value of transactions in December. The active sample relating to the universe of transactions in December had deteriorated, and although they are not part of this current study, insights into the nature of the deterioration can be ascertained from the experiment. For example, this type of experiment could be used to consider the effects of renewing the item selection more frequently. The fall in the coverage of the active sample is mitigated by the introduction of replacements when using direct comparisons. The weights used remain, of course, those in January 1998. However, the sample of information used in successive months draws on the replacement varieties.

The search procedure outlined above for replacements for direct comparisons, and used in some of the hedonic methods, had a minimum condition—that the replacement must at least be of the same make in a given outlet type. This can be seen from table 9.6 to have been generally met, the fall-off being negligible. The hedonic approach benefited from use of information on the whole sample in each month for the estimation of the equations. The estimated regression equations were based on the whole sample in spite of the inclusion of varieties with limited transactions on the grounds that such varieties were more likely to be the ones going missing. The twelve regressions estimated in each month had a mean  $R^2$  of 0.80, and in each case the null hypothesis of the individual coefficients being jointly equal to zero was rejected by an  $F$ -test. The average monthly sample size was 558.

#### 9.5.4 The Results

Table 9.7 provides the results. The extent of any bias arising from the imputation approach is dictated by the ratio of missing price comparisons to the total number of comparisons and by the difference between *quality-adjusted* price changes of the missing varieties, had they continued to exist, and those of other varieties (see the appendix). The bias from the class mean imputation approach should be smaller than the imputation approach, and these methods show the choice between the results matters—an approximately 2 percentage point difference over the year, for a roughly 10 percent fall.

**Table 9.7 Laspeyres Indexes**

	Imputation	Targeted Imputation	Direct Comparison	Predicted versus Actual			Adjustment via Coefficients			Predicted versus Predicted		
				Old	New	Geometric Mean	Old	New	Geometric Mean	Old	New	Geometric Mean
January	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
February	99.49	99.50	99.30	99.75	99.16	99.45	99.41	99.38	99.40	99.49	99.48	99.49
March	99.29	99.29	99.26	99.35	99.17	99.25	99.23	99.20	99.22	99.27	99.29	99.28
April	98.30	98.29	98.45	98.23	98.29	98.25	98.21	98.38	98.30	98.09	98.31	98.20
May	97.06	98.89	97.30	96.95	97.10	97.02	96.94	97.18	97.06	96.75	97.09	96.92
June	94.98	94.97	95.09	95.00	94.95	94.98	94.92	95.05	94.98	94.91	95.07	94.99
July	93.14	92.97	93.13	93.44	93.08	93.26	93.28	93.10	93.19	93.40	93.22	93.31
August	93.64	92.62	94.49	94.64	93.83	94.24	94.28	94.07	94.18	94.09	93.87	93.98
September	93.55	92.01	94.03	94.91	92.26	93.58	93.37	93.29	93.33	94.06	93.96	94.01
October	91.84	90.74	91.75	92.99	90.48	91.73	91.49	91.30	91.40	92.13	92.14	92.13
November	91.52	90.20	91.47	93.26	89.94	91.59	90.85	90.83	90.84	91.75	92.17	91.96
December	90.87	88.94	91.81	92.74	90.09	91.41	91.26	91.19	91.23	91.63	91.66	91.55

The price fall measured using direct comparisons was found to be smaller than the class mean result—a fall of 8.2 percent compared with 11.1 percent. A priori there is no expectation of the direction of bias from the targeted imputation approach (see the appendix), although for direct comparisons the replacements in varieties are more likely to be priced higher than the missing ones, the smaller fall in price thus being expected. However, given that the selection of replacement is not governed by the judicious selection of price collectors but by a computational procedure, the extent of the difference is smaller than expected.

The use of hedonic adjustment methods provides an explicit basis for the quality adjustments. The results for the predicted versus actual have a smaller fall for the “old” (equation [9]) than “new” (equation [10]) comparison (table 9.7). The actual old, disappearing varieties could be argued to be more likely to be below the hedonic surface as they price to clear the market, whereas the actual new varieties are more likely to be above the hedonic surface as they price-skim segments of consumers with higher price elasticities. However, the effect of both is to increase the price change. The geometric mean of an 8.6 percent fall is slightly higher than the other results, as expected.

The above estimates were also undertaken using a subset of the regression coefficients applied only to the characteristic differences that took place. If the variables for the characteristics are orthogonal to each other, the results will be the same as those for equations (9) and (10) discussed above. The adjustments are affected by multicollinearity as imprecise estimates of individual marginal values are utilized for the adjustment. Multicollinearity occurs when characteristics are bundled together, and it is not clear whether the differences between the results of the two formulations are some measure of this. Nonetheless, the results for the two geometric means are quite similar—a fall of 8.8 percent for the adjustment via coefficients.

The final set of estimates is the predicted versus predicted. The old and new estimates are very close, with a geometric mean and a fall of 8.4 percent. This compares with the geometric mean of the predicted versus actual of 8.6 percent, implying that the results, when averaged using the geometric mean, are not subject to serious misspecification error.

Thus, in summary, different hedonic adjustment techniques provide similar results, although the old and new predicted to actual appear to work best as a geometric mean. This similarity is encouraging, given the plethora of such approaches. These results are also similar to those from a computational, direct replacement method, but both imputation approaches lead to larger falls in prices, a result with no immediate explanation, given the analysis in the appendix.

### 9.5.5 Further Work

The results are exploratory in the sense that they arise from an experimental formulation that is subject to some limitations that cannot be easily

remedied as well as some restrictions that can. A major limitation is that the observations are for a product variety in a specific outlet type, as opposed to in a specific outlet (in a geographical place). That some models are specific to some outlet chains helps, but we cannot distinguish here between the locations of the outlets, although in principle this is possible with scanner data. This in itself may not be problematic for price comparisons, since there is still some debate over the validity of using the aggregated unit values over outlets for price comparisons (Balk 1999; Diewert 1990; de Haan and Opperdoes 1998). However, the concept of “missing” prices used here is not appropriate since a price collector may, for example, find a price missing for a variety in an outlet in Cardiff while other price collectors may find price quotes for the same variety in different stores or locations. The experiment would only treat prices as missing if there were no transactions anywhere for the product variety. Scanner data provide a proxy variable on the extent to which each variety is sold in different outlets, and use of this is being considered to develop the experiment.

Further work might also include consideration of the effects of the following:

- different sampling schemes for the January selection as opposed to the use of the universe of transactions
- more frequent item selection (rebasing) on the need for quality adjustments
- more frequent item selection (rebasing) on the coverage of the universe of transactions
- variations in the specification and sample used for the hedonic regressions
- different selection criteria for replacements
- use of different formulas
- comparison with hedonic indexes
- different rules for deciding when a variety is “missing”
- more refined class imputation procedures
- missing market innovations
- extension to other products

## Appendix

Triplett (2001) in a draft OECD manual has been responsible for a detailed analysis of the implicit bias from imputations, and although the formulation here is quite different, there is much that has been usefully applied from his analysis. For  $i = 1 \dots m$  varieties where  $P_{m,i}$  is the price of variety  $m$  in

period  $t$ ,  $P_{n,t+1}$  is the price of a replacement variety  $n$  in period  $t + 1$ ,  $A(h)$  is a quality adjustment to  $P_{n,t+1}$  that equates its quality services to  $P_{m,t+1}$  such that the quality-adjusted price  $P_{m,t+1}^* = A(h)P_{n,t+1}$ , and  $Q$  is the implicit adjustment that allows the method to work, the arithmetic formulation for one missing variety is given by

$$(A1) \quad \frac{1}{m} \left( \frac{P_{m,t+1}^*}{P_{m,t}} + \sum_{i=1}^{m-1} \frac{P_{i,t+1}}{P_{i,t}} \right) = \left[ \frac{1}{(m-1)} \sum_{i=1}^{m-1} \frac{P_{i,t+1}}{P_{i,t}} \right] + Q$$

$$(A2) \quad Q = \frac{1}{m} \frac{P_{m,t+1}^*}{P_{m,t}} - \frac{1}{m(m-1)} \sum_{i=1}^{m-1} \frac{P_{i,t+1}}{P_{i,t}}$$

and for  $x$  missing varieties by

$$(A3) \quad Q = \frac{1}{m} \sum_{i=1}^x \frac{P_{m,t+1}^*}{P_{m,t}} - \frac{x}{m(m-x)} \sum_{i=1}^{m-x} \frac{P_{i,t+1}}{P_{i,t}}$$

The relationships are readily visualized if  $r_1$  is defined as the respective geometric or arithmetic mean of price changes of varieties that continue to be recorded and  $r_2$  that of quality-adjusted missing varieties, that is, for the arithmetic case where

$$(A4) \quad r_1 = \left[ \sum_{i=1}^{m-x} \frac{P_{i,t+1}}{P_{i,t}} \right] \div (m-x)$$

$$r_2 = \left[ \sum_{i=1}^x \frac{P_{i,t+1}^*}{P_{i,t}} \right] \div x$$

then the ratio of arithmetic mean biases from substituting equations (A4) in (A3) is

$$(A6) \quad Q_g = \frac{x}{m} (r_2 - r_1),$$

which equal zero when  $r_1 = r_2$ . The bias depends on the ratio of missing values and the difference between the mean of price changes for existing varieties and the mean of quality-adjusted replacement to missing price changes. Note that the bias is small if *either*  $(x/m)$  or the difference between  $r_1$  and  $r_2$  is small. Furthermore, note that the method is reliant on a comparison between price changes for existing varieties and *quality-adjusted* price changes for the replacement/missing comparison. This is more likely to be justified than a comparison without the quality adjustment. For example, suppose we had  $m = 3$  varieties, each with a price of 100 in period  $t$ . Let the  $t + 1$  prices be 120 for two varieties, but assume the third is missing—that is,  $x = 1$ —and is replaced by a variety with a price of 140, 20 of which is due to quality differences. Then the arithmetic bias as given in equation (A6), where  $x = 1$ ,  $m = 3$  and  $r_2 = [(A(h) + P_{n,t+1})/P_{m,t}]$ , is

$$\frac{1}{3} \left[ \frac{140 - 20}{100} - \frac{240/2}{100} \right] = 0.$$

Had the bias depended on the *unadjusted price* of 140 compared with 100, the method would be prone to serious error. In this calculation the direction of the bias is given by  $(r_2 - r_1)$  and does not depend on whether quality is improving or deteriorating, that is, whether  $A(h) > P_{n,t+1}$  or  $A(h) < P_{n,t+1}$ . If  $A(h) > P_{n,t+1}$ , a quality improvement, it is still possible that  $r_2 < r_1$  and that the bias may be negative, a point stressed by Jack Triplett.

## References

- Abraham, K., J. Greenless, and B. Moulton. 1998. Working to improve the Consumer Price Index. *Journal of Economic Perspectives* 12 (1): 23–36.
- Arguea, N. M., C. Haseo, and G. A. Taylor. 1994. Estimating consumer preferences using market data: An application to U.S. automobile demand. *Journal of Applied Econometrics* 9:1–18.
- Armknecht, P. A., W. F. Lane, and K. J. Stewart. 1997. New products and the U.S. Consumer Price Index. In *The economic of new goods*. Studies in Income and Wealth, vol. 58, ed. T. F. Bresnahan and R. J. Gordon, 375–96. Chicago: University of Chicago Press.
- Armknecht, P. A., and F. Maitland-Smith. 1999. Price imputation and other techniques for dealing with missing observations, seasonality, and quality changes in price indices. IMF Working Paper no. WP/99/78. Washington, D.C.: International Monetary Fund, Statistics Department.
- Astin, J. A., and D. J. Sellwood. 1998. Harmonization in the European Union: A review of some technical issues. *Proceedings of the third meeting of the International Working Group on Price Indices*, ed. B. Balk, 291–308. Research Paper no. 9806. Voorburg, the Netherlands: Statistics Netherlands.
- Balk, B. 1999. On the use of unit value indices as consumer price sub-indices. In *Proceedings of the fourth International Working Group on Price Indices*, ed. W. Lane, 112–20. Washington, D.C.: Bureau of Labor Statistics.
- Berndt, E. R., Z. Griliches, and N. J. Rappaport. 1995. Econometric estimates of price indices for personal computers in the 1990s. *Journal of Econometrics* 68:243–68.
- Blow, L., and I. Crawford. 1999. *Cost of living indices and revealed preference*. London: The Institute of Fiscal Studies.
- Boskin, M. S., chairman. Advisory Commission to Study the Consumer Price Index. 1996. Towards a more accurate measure of the cost of living. Interim Report to the Senate Finance Committee. Washington, D.C.
- Boskin, M. S., E. R. Delberger, R. J. Gordon, Z. Griliches, and D. W. Jorgenson. 1998. Consumer prices in the Consumer Price Index and the cost of living. *Journal of Economic Perspectives* 12 (1): 3–26.
- Bradley, R., B. Cook, S. G. Leaver, and B. R. Moulton. 1998. An overview of research on potential uses of scanner data in the U.S. CPI. *Proceedings of the third meeting of the International Working Group on Price Indices*, ed. B. Balk, 169–182. Research Paper no. 9806. Statistics Netherlands.

- Cameron, J., and A. Collins. 1997. Estimates of hedonic ageing for partner search. *Kyklos* 50 (3): 409–18.
- Cole, R., Y. C. Chen, J. A. Barquin-Stolleman, E. Dulberger, N. Helvacian, and J. H. Hodge. 1986. Quality-adjusted price indexes for computer processors and selected peripheral equipment. *Survey of Current Business* 65 (1): 41–50.
- Cunningham, A. 1996. Measurement bias in price indices: An application to the UK's RPI. Bank of England Working Paper Series, no. 47. London: Bank of England.
- Curry, B., P. Morgan, and M. Silver. 2001. Hedonic regressions, misspecifications, and neural networks. *Applied Economics* 33:659–71.
- Dalen, J. 1998. Experiments with Swedish scanner data. *Proceedings of the third meeting of the International Working Group on Price Indices*, ed. B. Balk, 163–68. Research Paper no. 9806. Voorburg, the Netherlands: Statistics Netherlands.
- Diewert, W. E. 1976. Exact and superlative index numbers, *Journal of Econometrics* 4:115–45.
- . 1978. Superlative index numbers and consistency in aggregation. *Econometrica* 46:883–900.
- . 1980. Aggregation problems in the measurement of capital. In *The measurement of capital*, Studies in Income and Wealth, vol. 45, ed. Dan Usher, 433–528. Chicago: University of Chicago Press.
- . 1993. The early history of price index number research. In *Essays in index number theory*, vol. 1, ed. W. E. Diewert and A. O. Nakamura, 33–65. Amsterdam: North Holland.
- . 1995. Axiomatic and economic approaches to elementary price indices. Discussion Paper no. 95-01. Vancouver, Canada: University of British Columbia, Department of Economics.
- . 1996. Sources of bias in consumer price indices. Discussion Paper no. DP-96-04. University of New South Wales, School of Economics.
- . 1997. Alternative strategies for aggregating prices in the CPI: Comment. *Federal Reserve Bank of St Louis Review* 79 (3): 113–25.
- Feenstra, R. C. 1995. Exact hedonic price indices. *Review of Economics and Statistics* 77:634–54.
- Feenstra, R. C., and E. W. Diewert. 2000. Imputation and price indices: Theory and evidence from the International Price Program. BLS Working Paper no. 335. Washington, D.C.: Bureau of Labor Statistics, Office of Prices and Living Conditions.
- Fixler, D., and K. Zieschang. 1992. Incorporating ancillary measures of processes and quality change into a superlative productivity index. *Journal of Productivity Analysis* 2:245–67.
- Forsyth, F. G., and R. F. Fowler. 1981. The theory and practice of chain price index numbers. *Journal of the Royal Statistical Society*, series A, 144 (2): 224–47.
- Gandal, N. 1994. Hedonic price indices for spreadsheets and empirical test for network externalities. *RAND Journal of Economics* 25 (1): 160–70.
- . 1995. Competing compatibility standards and network externalities in the PC market. *Review of Economics and Statistics* 77:599–608.
- Goldberger, A. A. 1968. The interpretation and estimation of Cobb-Douglas functions. *Econometrica* 35 (3–4): 464–72.
- Gordon, R. L. 1990. *The measurement of durable goods prices*. Chicago: University of Chicago Press.
- Griliches, Z. 1990. Hedonic price indices and the measurement of capital and productivity: Some historical reflections. In *Fifty years of economic measurement: The Jubilee Conference of Research in Income and Wealth*, Studies in Income and

- Wealth, vol. 54, ed. E. R. Berndt and J. E. Triplett, 185–206. Chicago: University of Chicago Press.
- Haan, J. de, and E. Opperdoes. 1998. Estimation of the coffee price index using scanner data: The choice of the micro index. *Proceedings of the third meeting of the International Working Group on Price Indices*, ed. B. Balk, 191–202. Research Paper no. 9806. Voorburg, the Netherlands: Statistics Netherlands.
- Hausman, J. R. 1997. Valuation of new goods under perfect and imperfect conditions. In *The economics of new goods*, Studies in Income and Wealth vol. 58, ed. T. Bresnahan and R. J. Gordon, 209–48. Chicago: University of Chicago Press.
- Hawkes, W., and R. Smith. 1999. Improving consumer price measurements through the use of scanner data and market segmentation. In *Proceedings of the Measurement of Inflation Conference*, ed. M. Silver and D. Fenwick, 283–99. Cardiff, Wales: Cardiff Business School.
- Hicks, J. R. 1940. The valuation of the social income. *Economica* 7:105–24.
- Hoffmann, J. 1998. Problems of inflation measurement in Germany. Economic Research Group of Deutsche Bundesbank Discussion Paper no. 1/98. Frankfurt: Deutsche Bundesbank.
- Ioannidis, C., and M. Silver. 1999. Estimating hedonic indices: An application to U.K. television sets. *Zeitschrift für Nationalökonomie* (Journal of Economics) 69 (1): 71–94.
- Kmietowicz, Z. W., and M. S. Silver. 1980. New products and the Index of Industrial Production. *Journal of Development Studies* 16 (4): 463–67.
- Kokoski, M., K. Waehrer, and P. Rozaklis. 1999. Using hedonic methods for quality adjustment in the CPI: The consumer audio products component. BLS Working Paper. Washington, D.C.: Bureau of Labor Statistics, Division of Price and Index Number Research. Available at [<http://www.bls.gov/cpi/cpiaudio.htm>].
- Kotüs, A. A. 1939 (1924). The problem of the true index of the cost-of-living. *Econometrica* 7 (1): 10–29. First published in Russian.
- Lerner, J. 1995. Pricing and financial resources: An analysis of the disk drive industry, 1980–88. *The Review of Economics and Statistics* 77:585–98.
- Lowe, R. 1999. Televisions: Quality changes and scanner data. In *Proceedings of the fourth meeting of the International Working Group on Price Indices*, ed. W. Lane, 5–20. Washington, D. C.: Bureau of Labor Statistics.
- McFadden, D. 1983. Econometric models of probabilistic choice. In *Structural analysis of discrete data with econometric applications*, ed. Charles F. Manski and Daniel McFadden, 198–272. Cambridge: MIT Press.
- Mills, E. S., and R. Simenauer. 1996. New hedonic estimates of regional constant quality house prices. *Journal of Urban Economics* 39 (2): 209–15.
- Moulton, B., T. J. LaFleur, and K. E. Moses. 1999. Research on improved quality adjustments in the CPI: The case of televisions. In *Proceedings of the fourth meeting of the International Working Group on Price Indices*, ed. W. Lane, 77–99. Washington, D.C.: Bureau of Labor Statistics.
- Moulton, B., and K. E. Moses. 1997. Addressing the quality change issue in the Consumer Price Index. *Brookings Papers on Economic Activity*, issue no. 1:305–49.
- Murray, J., and N. Sarantis. 1999. User cost, forward-looking behaviour, and the demand for cars in the U.K. *Journal of Economics and Business* 51 (3): 237–58.
- Nelson, R. A., T. L. Tanguay, and C. D. Patterson. 1994. A quality-adjusted price index for personal computers. *Journal of Business and Economics Statistics* 12 (1): 23–31.
- Rasmussen, D. W., and T. W. Zuehlke. 1990. On the choice of functional form for hedonic price functions. *Applied Economics* 22 (4): 431–38.

- Reinsdorf, M. 1996. Constructing basic components for the U.S. CPI from scanner data: A test using data on coffee. Working Paper no. 277. Washington, D.C.: Bureau of Labor Statistics.
- . 1998. Divisia indices and the representative consumer problem. In *Proceedings of the fourth meeting of the International Working Group on Price Indices*, ed. W. Lane, 210–235. Washington, D.C.: Bureau of Labor Statistics.
- Reinsdorf, M., P. Liegey, and K. Stewart. 1995. New ways of handling quality change in the U.S. Consumer Price Index. Working Paper no. 276. Washington, D.C.: Bureau of Labor Statistics.
- Rosen, S. 1974. Hedonic prices and implicit markets: Product differentiation and pure competition. *Journal of Political Economy* 82:34–49.
- Saglio, A. 1995. Comparative changes in average price and a price index: Two case studies. In *Proceedings of the first meeting of the International Working Group on Price Indices*, ed. L. M. Duchare. Ottawa: Statistics Canada. Available at [<http://www4.statcan.ca/secure/english/ottawagroup/toc1.htm>].
- Shapiro, M. D., and D. W. Wilcox. 1997. Alternative strategies for aggregating prices in the CPI. *Federal Reserve Bank Review* 79 (3): 113–26.
- Silver, M. 1995. Elementary aggregates, micro-indices, and scanner data: Some issues in the compilation of consumer prices. *Review of Income and Wealth* 41:427–38.
- . 1999. An evaluation of the use of hedonic regressions for basic components of consumer price indices. *Review of Income and Wealth* 45 (1): 41–56.
- Stewart, K. G., and J. C. H. Jones. 1998. Hedonic and demand analysis: The implicit demand for player attributes. *Economic Inquiry* 36 (2): 192–202.
- Szulc, B. J. 1983. Linking price index numbers. In *Price level measurement*, ed. W. E. Diewert and C. Montmarquette, 537–66. Ottawa: Statistics Canada.
- Triplett, J. E. 1987. Hedonic functions and hedonic indices. In *The new Palgrave dictionary of economics*. 630–34. Basingstoke, England: Palgrave Macmillan.
- . 1990. Hedonic methods in statistical agency environments: An intellectual biopsy. In *Fifty years of economic measurement: The Jubilee Conference on Research in Income and Wealth*, Studies in Income and Wealth, vol. 56, ed. E. R. Berndt and J. E. Triplett, 207–38. Chicago: University of Chicago Press.
- . 2001. Draft of handbook on quality adjustment of price indexes for information and communication technology products. OECD Directorate for Science technology and Industry. Paris: Organization for Economic Cooperation and Development.
- Turvey, R. 1989. *Consumer price indices*. Geneva: International Labour Organisation.
- . 1999a. Draft of elementary aggregate indices manual. Available at [<http://www.turvey.demon.co.uk>]. Revised April.
- . 1999b. Incorporating new models into a CPI: PCs as an example. In *Proceedings of the Measurement of Inflation Conference*, ed. M. Silver and D. Fenwick, 505–10. Cardiff, Wales: Cardiff Business School.
- White, H. 1980. A heteroskedasticity-consistent covariance matrix and a direct test for heteroskedasticity. *Econometrica* 48:721–46.