INVITED PAPER No 7 (9)*
19 October 2001

ENGLISH ONLY

STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

INTERNATIONAL LABOUR ORGANISATION (ILO)

CONFERENCE OF EUROPEAN STATISTICIANS

Joint ECE/ILO Meeting
on Consumer Price Indices
(Geneva, 1-2 November 2001)

WHY THE CPI MATCHED MODELS METHOD MAY FAIL US: RESULTS FROM AN HEDONIC AND MATCHED EXPERIMENT USING SCANNER DATA

Invited paper submitted by Bureau of Labor Statistics of the United States of America**

---

* Paper posted on Internet as submitted by the country.

** Prepared by Mick Silver*** and Saeed Heravi, Cardiff Business School, Cardiff University, Colum Drive, Cardiff CF10 3EU, United Kingdom.
Earlier versions of this paper benefited from useful comments from Ernst Berndt (MIT), Erwin Diewert (University of British Columbia), Rob Feenstra (University of California, Davis) and John Greenlees (BLS) when presented at *National Bureau of Economic Research, Conference on Research in Income and Wealth*; Washington DC, September 2000, from participants at the Ottawa Working Group on Price Indices, Sydney, March 2001 and from Johannes Hoffmann (Deutches Bundesbank) at a Symposium on Hedonic Methods, Wiesbaden, June 2001. In addition the authors had comments from, and sight of a draft OECD Manual by, Jack Triplett (Brookings Institution) which proved most useful for this work. The authors acknowledge support from an American Statistical Association/ National Science Foundation/ Bureau of Labor Statistics Fellowship.The usual disclaimers apply.

*** Visiting Senior Research Fellow, Bureau of Labor Statistics, 2 Massachusetts Ave., NE, Washington DC 20212, and Visiting Scholar, International Monetary Fund, Washington DC 20431.

Abstract

Statistical offices use the matched models method to compile consumer price indices (CPIs) to measure inflation. Price collectors record the prices of a sample of models or varieties of selected products existing in period *t,* and then continue to record the prices of these same matched models in subsequent periods *t+n.* The matched models method is designed to control for quality changes by ensuring that like is always compared with like. But new, unmatched models launched in periods *t+1,.. t+n* have their prices ignored unless they are used as replacements. The matched models method also fails when an old model is no longer available and a matched comparison of like with like cannot take place. In markets where models change rapidly due to technological innovations the matched sample may, by the end of a year, be quite unrepresentative of what is, and what was, bought. We show that serious sample degradation is a necessary condition for the matched approach to fail and then consider how the prices of unmatched models differ from matched ones and estimate their impact on the index. The paper thus argues that a form of sampling bias may exist as a by-product of the method used to control for quality changes. The bias is one not recognised in major studies of CPIs including Boskin (1996). An alternative approach to measuring quality adjusted price changes is the use of hedonic indices which can use data on matched and unmatched models. This study is in part motivated by evidence of differences between the results of the two approaches. A replication of CPI procedures is attempted on scanner data. Bar-code scanner data is particularly suited to this purpose covering the universe of transactions. The data allows us to identify the extent of sample degradation, compare unmatched (new and old) prices and matched prices and their effects on price measurement and demonstrate the effects of more frequent sample rotation, such as chaining. Hedonic indices can use the whole sample and chained indices refresh the sample on a regular basis. They are argued to be more useful to price measurement in markets with a rapid turnover of models.

JEL classification: C43, C81, D12, E31, L15, L68, O47

Keywords: Superlative index numbers; Cost of living indices; COLI; Consumer price indices; Hedonic regressions; Quality adjustment; Inflation.

1. INTRODUCTION

The matched models method is the universally accepted procedure for the compilation of consumer price indices (CPI) and the measurement of inflation. The details and prices of a representative selection of items are collected in some base/reference period, and their matched prices are collected in subsequent periods, so that 'like' is compared with 'like'. It is no mean affair. The local (as opposed to central) collection of prices for the U.S. CPI includes about 835,000 price quotations (Armknecht, 1996) and for the U.K. Retail Prices Index (RPI) some 110,000 quotations each month (Baxter, 1998). It has nonetheless been criticised for its inability to properly incorporate quality changes, substitution bias and the effects of new goods (Boskin *et al*., 1996 and 1998; Diewert, 1996; Cunningham, 1996; Hoffmann, 1998; Abraham *et al*., 1998). Results from hedonic regression studies have shown different, usually lower rates of price changes (recent examples including, Hoffmann (1998), Silver and Heravi (2000 and 2001), Koskimäki and Vartia (2001) and Pakes (2001)). Yet the matched method has its supporters. Alan Greenspan, in commenting on the need for better micro data for price measurement, reported on how the conceptually simpler matched model method can give comparable results to the hedonic approach when detailed micro data are used (Greenspan, 2001 citing Aizcorbe *et al.,* 2000), though we comment on this later.

Hedonic approaches use regression techniques whereby, in their simplest form, the price of an item is regressed on its quality characteristics and dummy variables for the time period to which the observations relate. The coefficients on these time dummies are estimates of the change in price over the period concerned, controlling for changes in the quality mix of what was bought. Studies have found substantial differences in the results from the matched and hedonic approaches. When comparing results there is often a preference for hedonic approach as the benchmark, though the basis for this is not always apparent Boskin *et al*., (1996 and 1998) and Hoffmann (1998).

The matter is of some importance. Boskin *et al*., (1996 and 1998) estimated the cumulative additional national debt from over-indexing the budget over a dozen years would be more than $1 trillion. For the U.K. a 0.1 percentage point overstatement of the RPI is estimated by the ONS to affect Government expenditure and receipts by about £100 million a year (Fenwick et al., 2000). Such bias affects the targeting of inflation; taxation liability and bank payment adjustments; indexation in legal contracts, wages and benefits; current cost accounting; deflation of national accounts, wages and retail sales.

So what is going wrong? Is there some problem with the matched models method from which the hedonic approach does not suffer?

Four reasons are considered for why the matched models approach may fail us:

Its static sampling universe*:* The matching procedure has at its roots a conundrum. Matching is designed to avoid price changes being contaminated by quality changes. Yet its adoption constrains us to a static universe of models which exist in both the reference and base period. Outside of this there is of course something more: models which exist in the reference period, but not in the current period, and are therefore not matched, and similarly those new models existing in the current period but not the reference one – the dynamic universe (Dalén, 1998 and Sellwood, 2001). Of course the new models not in the reference period may be the ones undergoing more rapid technological change and the old ones may incorporate an obsolete technology, both experiencing unusual (quality-adjusted) price changes. The conundrum is that the very device used to maintain a constant-quality sample not marred by technological change, may itself give rise to a biased sample which excludes new, unmatched technological developments and old, unmatched obsolete ones.

Missing items and quality adjustment: when items in subsequent periods are missing – the items are simply not available for any record to be made of their prices. They may be obsolete or no longer stocked by the establishment. A number of procedures are available to statistical offices in such circumstances to impute the missing prices, replace the missing prices with comparable models or a non-comparable ones with associated adjustments for quality differences. If this is being undertaken on a substantial basis and there is a bias in these procedures the matched models method fails us.

Quality adjustment and the sample space: a third point is the inter-relationship between sampling and quality adjustment. As noted above quality adjustment techniques are used when an item is permanently missing. The practices used when items are missing have implications for the active sample. For example, the selection of comparable items may be directed towards items with high sales to ameliorate the sample depletion, or conversely towards ‘similar’ items which by their nature will have low sales. Or the use of imputations based on the price changes of the existing sample may suffer from allowing the sample to further deteriorate***.*** And when this degraded sample is used to make imputations as to the price changes of replacement items, it of course reflects the changing technology of a sample not representative of current technological changes.

New products: A final potential source of error lies when something ‘new’ is introduced into the market place. There is a difficulty in distinguishing between new items and quality changes. When a quite new item is introduced there is a gain to consumers’ utility. For example the introduction of the video cassette recorder (VCR) was a completely new good that led to an initial gain to consumers’ utility. This welfare gain from its introduction could not be properly brought into the index by waiting until the index was re-based and a new basket of goods that included the VCR formed. We would include the subsequent price changes, but not the initial gain in welfare accompanying its introduction. A measure of the welfare gain requires an estimate of the reservation price of the VCR in the period before its introduction, that is, when demand would be zero (Hicks, 1940 and Hausman, 1997).

The subject of bias in CPI measurement has a long history more recent interest following Boskin (1996). Here the focus was on quality adjustment bias when models are missing, substitution bias at a higher and lower level for products and items and also between outlets, and new product bias. The study of bias from different quality adjustment procedures when items go missing, has benefited from a number of useful studies including Reinsdorf *et al*., 1995, Moulton and Moses, 1997, Armknecht *et al*., 1997 and Moulton *et al*., 1998, Lowe (1998), Kokoski *et al*., 1999a, Silver and Heravi, 2000 and Triplett, 2001). Here our concern is with the more neglected area of the use of a static universe in a dynamic market.

In section 2 of this paper we provide some background to the matched and hedonic approaches. Section 3 continues by considering why the matched model method may fail and lays down the study objectives for the empirical section. These include sample degradation when matching, the results from some studies on, and an analytical framework for, the effects of dissimilarities between matched and unmatched models, the effect of weighting and leverage. Section 4 provides an outline of an experimental formulation and its use of scanner data on washing machines to replicate the matched model method. The scanner data are taken to be the active universe of transactions. In section 5 we start the empirical results: the effect of the matched models method on sample coverage is demonstrated for long run and short run Laspeyres and for more frequent intervals of sample rotation including chained indices. One reason why hedonic and matched indices (as used by statistical offices) may differ is that the latter, in matching, uses a depleted sample. A concern here is with whether the coverage is substantially depleted and the extent to which more regular sample rotation or chaining militates against this. Yet sample depletion is only a necessary condition for bias. If matched and unmatched models were priced in a similar manner, the depletion would not lead to discrepancies in the two approaches. Section 6 continues by considering the nature and impact on the index of the pricing of unmatched new and old models compared with matched models. Their different prices, characteristics, hedonic (quality-adjusted) prices and price changes, residuals from a common hedonic surface and leverage are all examined. Section 7 concludes the empirical analysis with summary results from the matched models method which uses only matched models, hedonic indices using the full sample, chained monthly indices and more frequent sample rotation which refreshes the sample more regularly, and matching using explicit hedonic quality adjustment when models are missing. Overall conclusions are drawn in section 8. In particular it is proposed that the matched models method be no longer used for products areas where rapid changes in models take place. A sample of prices of all or major selling models can be taken in each period and hedonic or chained indices used. The matched models approach was developed for, and is suitable for use in, product areas where models that remain responsible for sizeable market shares continue over relatively long periods of time. Product areas such as personal computers, mobile phones, televisions, washing machines and many consumer durables are no longer like this and the matched models method can lead to bias in such circumstances, the nature of which is outlined in this study.

A novel feature of the paper is its use of scanner data on the universe of transactions, which also allows us to follow a fixed base matched sample and compare matched and unmatched observations, something not available to studies using CPI matched data.

2. BACKGROUND

a) Background to the matched models approach

Consumer price indices serve as measures of the cost of living. The conceptual basis for such indices is well established in economic theory. Konüs (1939) and Diewert (1976) define a theoretical cost-of-living index (COLI), P_c as the ratio of the minimum expenditures of achieving 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}; i.e.

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

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

$$\mathrm{P_c}\left(P_{t,} P_{t-1,} z_{t,} z_{t-1,} U\right) = E\left(p_{t,} z_{t,} U\right) / E\left(p_{t-1,} z_{t-1,} U\right) \quad (2)$$

i.e. the ratio of the minimum expenditures 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}.

Against these frameworks is the practical need to compile a consumer price index. Representative samples of prices are required in each period as are data on expenditure patterns to act as weights. There is also a need for an aggregator, which calculates some weighted average of price changes. Economic theory, along with axioms as to good properties of index numbers, has proved most useful to the selection of such formulae. Laspeyres is now well known to have a substitution bias as consumers substitute away from items, products and outlets with relatively high price changes. The Laspeyres fixed weights index corresponds to a Leontief utility function, which cannot adjust consumption patterns for such changes. It thus overstates its theoretical cost-of-living counterpart. A class of index numbers referred to a *superlative* has also been developed. These correspond to flexible functional forms that allow for substitution effects and include Fisher and Törnqvist indices (Diewert, 1976 and 1978).

All of this is well and good, except for the fact that in many product areas the items that are purchased change over time, as new models/varieties are introduced to markets which are often highly differentiated with several brands each having different offerings to different segments. Also new models, which may be radically different to the old models, may appear on the market. Taking a random sample of prices and comparing some average to an average from a previous month will not provide a measure of price changes untainted by the quality changes. Prices of, for example personal computers may fall over time, but when the increased processing power of the PC is taken into account, the fall in terms of services rendered or utility derived, will be far greater than that of the measured price comparison.

So statistical offices use the matched models method to help get round this problem. In a reference month models are selected and their prices and details recorded so that matched prices can be collected and recorded in subsequent months and like is compared with like.

b) Background to the hedonic approach

We distinguish here between four types of hedonic studies (see also Triplett, 1988 and 2001).

(I) Filling in missing unmatched prices - patching

The first is where the matched models method is being used and statistical agencies have missing unmatched models. The price collector can only find a replacement model which is not directly comparable and the coefficients from a hedonic regression are used to make a quality adjustment, so that the old and new price can be compared. This is a 'patched' solution in the sense that adjustments for quality differences are made to non-comparable models and the adjusted 'patched' price used for price comparisons.

(ii) Hedonic price functions for given quality points

The second method is an extension of the 'patching' approach to the whole data set. Hedonic regressions are estimated for period *t* and/or *t+n*. The features of models in period *t* are inserted into the hedonic regression equation for period *t+n* and price imputations made for each observation: the estimates are the prices in period *t+n of* the period *t* sample. Some average of the actual period *t* prices can then be compared with the imputed period *t+n* one. Alternatively the imputations could be made for period *t* using the period *t* regression and period *t+n* features: the imputations would be the prices in period *t+n* using period *t* valuations. Some average of the imputed period *t* prices would be compared with the actual period *t+n* ones.

(iii) Dummy variable hedonic

A third approach is the dummy variable method. This is again separate from the matched models method. The sample required does not have to be matched. A set of (z_k = 1,….*K*) characteristics of a product are identified and data over *i*=1,…*N* product varieties (or models) over *t*=1,…,*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:

$$1np_{ti} = \boldsymbol{b}_0 + \sum_{t=2}^{T} \boldsymbol{b}_t D_t + \sum_{k=1}^{K} \boldsymbol{b}_k z_{tki} + \boldsymbol{e}_{ti} \qquad (3)$$

where D_t are dummy variables for the time periods, D_2 being 1 in period *t*=2, zero otherwise; D_3 being 1in period *t*=3, zero otherwise etc.

The coefficients $\boldsymbol{b}_t$ are estimates of quality-adjusted price changes, that is estimates of the change in the (the logarithm of) price between period *t* and period *t+n*, having controlled for the effects of variation in quality (via $\sum_{k=1}^{K} \boldsymbol{b}_k z_{tkj}$).

(iv) Superlative/exact hedonic framework

The final approach arises out of the economic theory of price indices and involves the compilation of superlative/exact hedonic indices (Fixler and Zieschang, 1992, Feenstra, 1995 and Diewert, 2001). Superlative indices can, and will, be calculated using matched scanner data. However, the superlative/exact *hedonic* framework attempts to minimise loss of data through failures to match. This is not the subject of this study but is considered in Silver and Heravi (2000 and 2001).

3. WHY THE MATCHED MODELS METHOD MAY FAIL

(a) Its static sampling universe:

(i) Coverage of dynamic universe

The matching of prices of identical models over time, by its nature, is likely to lead to a serious depleted sample. Sellwood (2001:3-4) notes:

> "Matching is a procedure that is undertaken as part and parcel of the Laspeyres concept but it serves to prescribe a particular universe. In simple terms this is a 'static sub-universe' of transactions…in the current and reference periods for which there is a correspondence given by the matching criteria actually applied. That is where, for the purpose of the CPI there is deemed to be no product change and hence no quality change……….Matching of identical and 'closely similar' sample products in pairs implies a universe that is only part of the complete universe of all transactions. It is those transactions in products available at both t=1 (the reference period) and t=2 (the current period but not transactions in those products only available at either t=1 or t=2 (but not both)….These transactions excluded from consideration by the matching are a "dynamic sub-universe' the complement of the static sub-universe."

Instructions to price collectors to hold on to models until forced replacements are required has the effect of including unusual price changes of models as they become obsolete, and ignoring the unusual price changes of models on their launch. It is of course difficult to ascertain the extent to which matching constrains our penetration into the dynamic universe. Sellwood (2001) advocates simulations using the universe of scanner data – a subject of the empirical section of the paper. We consider some recent studies which use hedonic approaches on the whole sample and matched approaches.

(ii) Similarity of matched and unmatched observations and their effects on measured price changes: some studies

The matching of prices of identical models over time may also lead to the monitoring of a sample of models increasingly unrepresentative of the population of transactions. There are old models that existed when the sample was drawn, not being available in the current period and new ones that enter it, not being available in the base period. It may be that the exits have relatively low prices and the entrants relatively high ones and by ignoring these prices we introduce a bias.

Koskimäki and Vartia (2001) attempted to match prices of models of personal computers (PCs) over three two-month periods (spring, summer and fall) using a sample of prices collected as part of standard price collection for the Finish CPI. Of the 83 spring prices only 55 matched pairs could be made with the summer, and then only 16 continued through to the fall. They noted the sample of matched pairs to get rapidly biased: of the 79 models in the fall, the 16 matched ones had a mean processor speed of 518MHz compared with 628MHz for the remaining 63 unmatched ones; the respective hard disk sizes were 10.2 and 15.0 Gb., and percentages of high-end processors (Pentium III and AMD Atl.) 25 and 49.2 respectively. Hardly any change in *matched* prices was found over this

6 month period, while a hedonic regression analysis using all of the data found quality-adjusted price falls of around 10%. Instructions to price collectors to hold on to models until forced replacements are required may thus lead to an increasingly unrepresentative of the population and be biased towards technically inferior variants. In this instance the hedonic price changes fell faster, the newer models being cheaper for the services supplied.

Kokoski et al. (1999b) provides some interesting insights in an empirical study of inter-area price comparisons across U.S. urban areas using U.S. Consumer Price Index data. They provide the detailed regression results from data on prices, geographical area in which the prices were collected, and an extensive list of quality characteristics for over 6,700 separately priced potatoes, lettuces and tomatoes and bananas, 9,400 observations on apples, 12,000 on oranges, 14,200 on other fresh vegetables, and 26,009 on other fresh fruit. They regressed the price of, for example, potatoes each of the 43 geographical areas (42 dummy variables) in which the prices were collected, the 12 months, 7 outlet types, 4 package types, 3 sizes and 9 varieties. They also included a variable if the price was collected from a sample, which was recently rotated, as opposed to being in the same sample as the previous month. This is of particular interest since the coefficient on this dummy variable for the 'new sample' is an estimate of the effect on prices of more recent sample selection. For all eight products the coefficients on the new sample were negative, though in only 3 cases were they statistically significant, with magnitudes of about minus 3 and 4 percent. The negative sign indicated that quality adjusted prices were lower for the newly included model than the old one. And this is for a product group in which technological quality change is not immediately apparent.

Pakes (2001) estimated hedonic and matched indices for personal computers (PCs) sold between 1995 and 1999. Using annual data he found that about 85 percent of base-period 1995 observations could not be matched in the comparison year and were dropped in the calculation of matched model indices. The results were compared with those from a hedonic equation:

> "The estimated matched model index is the *opposite sign* of the proper hedonic in every year and averages +.27 per cent over the years (compared to about -.16 per cent for the proper hedonics) indicating that the positive bias generated by the selection on survival dominates the fall in prices generated by technical change. Also there is a negative correlation between the matched model and proper hedonic over the years. The selection argument can also explain this finding. Selection effects should be most positive in the years with the largest rate of technical change as those are the years in which all but the mot superior of base period goods are obsoleted by entering products. Years with a lot of technical change are the years that we expect to also see the largest fall in prices." [author's emphasis](Pakes, 2001, p.38).

Aizcorbe et al. (2000) undertook an extensive and meticulous study of high technology goods (personal computers and semiconductors) using quarterly data for the period 1993 to 1999. They used Fisher indices for matched models as a benchmark and compared the results with dummy variable hedonic indices. The difference was recognised as being composed of two terms: the difference between the hedonic and an (unweighted) geometric mean matched model and the difference between the geometric mean and the (superlative) Fisher index. The latter difference was due to the superior weighting system which Fisher benefits from. The former is the fairer comparison: between an unweighted matched geometric mean and an unweighted OLS semi-logarithmic dummy variable hedonic estimator. Feenstra (1995) has shown that a semi-logarithmic hedonic function has a correspondence to a geometric mean aggregator. The results from all three approaches were remarkably similar over the seven years of the study. For, example, for desktop CPUs the index between the seven years of 1993:Q1 and 1999:Q4 fell by 60.0 % (dummy variable hedonic), 59.9% (Fisher) and 57.8% (geometric mean). So why are these results similar when matched indices use a different sample space to hedonic studies which include unmatched ones? In this study *chained* Fisher and *chained* geometric means were used [correspondence with authors] so the sample was refreshed in each period. The number and proportion of models in a say, direct matched comparison between period 6 and period 0, would be very much smaller than in an index with chained comparisons. A chained index compares prices of models in period 0 with period 1 ($Index_{0,1}$) and then as a new exercise, studies the universe of models in period 1 and identifies matches in period 2 and links the result ($Index_{1,2}$) to the index $Index_{0,1}$ by successive multiplication continuing to $Index_{5,6}$ to form $Index_{0,6.}$ Only computers available in both period 0 and period 6 would be used in normal CPI compilation. The differences in sample space were therefore less marked and the results therefore closer. Chained matching still suffers from excluding some new models: for example *new* models in period 2 would be excluded from $Index_{1,2}$, but included in $Index_{2,3}$; the sample depletion is much less severe. We investigate the effect of chaining on the sample space later in this study.

Yet the authors stressed that substantial differences for a few annual comparisons existed. For example for desktop CPUs in 1996:Q4 the 38.2 percent annual fall measured by the dummy variable hedonic method differed from the geometric mean index by 17 percentage points and from the Fisher index by 22.2 percentage points. So why was there a difference between the (chained) geometric mean and the hedonic indices in this case? This should only happen if there was an unusually high proportion of models being turned over for the comparisons in question. In 1996:Q4 they reported 2 new observations (15.4% of all observations) entering the market which could not be matched with previous observations, and were therefore excluded from the chained matched analysis. And 2 exiting models which again were lost from the matched comparisons, though these 2 exiting models accounted for 15.4 percent of observations, but only 0.8 percent of revenue share. Other periods in which the indices diverged were marked by high model turnover. The study thus supports our contention regarding sample space. For chained indices with little model turnover there is little discrepancy between hedonic and matched models methods. It is only when binary comparisons or links have a high model turnover that differences arise.

The conclusions from the study are worth considering in this light. First, for high frequency, very disaggregated data matched models methods will generally capture the rapid quality change. This is because for such data the "..market share of turnover varieties tends to be small." Aizcorbe et al. (2000,p19). This applies only for chaining and is a conclusion that cannot be carried on to normal CPI practice. This is an important caveat since the chaining matches models in period *t* with prices in the succeeding period. It thus misses data not available in *both* periods. It works well if these missing models are unimportant - a finding supported by Silver and Heravi, 2000 and 2001. Aizcorbe et al. (2000) note the need for alternative approaches to data collection to allow such chained matching. For now we want to know if fixed base matching as practically practised in CPI compilation works –

and the sample turnover is likely to be much larger in this case as we will seek to show, than in the careful matching of chained indices used in Aizcorbe et al. (2000).

The above is not to say that chained indices will always be close to hedonic indices. In studies of price changes of six consumer products Silver and Heravi (2000 and 2001) found OLS hedonic estimates to fall slower than matched estimates for five of the six products: vacuum cleaners, washing machines, personal computers, televisions sets and cameras, the results for dishwashers being borderline. This supports the view that the matching is ignoring models benefiting from higher levels of technological developments, thus understating quality-adjusted price changes.

(iii) <u>Similarity of matched and unmatched observations and their effects on measured price changes</u>
We borrow now on the valuable analytical framework used for analysis by Aizcorbe et al. (2000) and generalise it from one new unmatched model to include more than one new and more than one old unmatched model. The analysis is for period t and t-1 comparisons, but holds for period 0 to $t+n$ comparisons - so the implications for longer term fixed base comparisons remain. An index from a semi-logarithmic (prices in logs) dummy variable hedonic regression for matched models can be shown to equal to:

$$p_t/p_{t-1} = \sum_{m\in Mt} (\ln p_{mt} - Z_m)/M_t \; - \sum_{m\in Mt-1} (\ln p_{mt-1} - Z_m)/M_{t-1} \qquad (4)$$

where Z_t and Z_{t-1} are in principle the quality adjustments to the dummy variables for time in equation (3), that is, $\sum_{k=1}^{K} \boldsymbol{b}_k z_{tk}$ - in this context, since matched models are being used, they are fixed effect dummy variables for each model. Equation (4) is simply the difference between two geometric means of quality-adjusted prices. The sample space $m = M_t = M_{t-1}$ is the same model in each period. Consider the introduction of a new model n introduced in period t with no counterpart in t-1 and the demise of an old model o so it has no counterpart in t. So M_t is composed of m and n, and M_{t-1} is composed of m and o and M are only the matched models m. The dummy variable hedonic comparison is now:

$$\begin{aligned}
\ln p_t/p_{t-1} \;=\; & \left[m/(m+n)\sum_m (\ln p_{mt} - Z_m)/m \;+\; n/(m+n)\sum_n (\ln p_{nt} - Z_n)/n\right] \\
& - \left[m/(m+o)\sum_m (\ln p_{mt-1} - Z_m)/m \;+\; o/(m+o)\sum_o (\ln p_{ot-1} - Z_o)/o\right] \\
=\; & \left[m/(m+n)\sum_m (\ln p_{mt} - Z_m)/m \;-\; m/(m+o)\sum_m (\ln p_{mt-1} - Z_m)/m\right] \\
& + \left[n/(m+n)\sum_n (\ln p_{nt} - Z_n)/n \;-\; o/(m+o)\sum_o (\ln p_{ot-1} - Z_o)/o\right] \qquad (5)
\end{aligned}$$

Consider the second expression in equation (5). We first have the change for *m* matched observations. It is the change in mean prices of matched models *m* in period *t* and *t*-1 adjusted for quality. Note that the weight in period *t* for this matched component is the proportion of matched to all observations in period *t.* And similarly for period *t*-1 the matched weight depends on how many unmatched old observations are in the sample. In the last line of equation (5) the change is between the unmatched new and the unmatched old mean (quality adjusted) prices in periods *t* and *t*-1. Thus matched methods can be seen to ignore the last line in equation (5) and will thus differ from the hedonic dummy variable approach. The hedonic dummy variable approach in its inclusion of unmatched old and new observations can be seen from equation (5) to change the emphasis given to matched observations.

(iv) The weighting of matched and unmatched price changes

The weights again refer to the relative number of observations in each period. Ideally imputed prices for unmatched new models in period *t*-1 should have been calculated and used. The imputed price would be the price of model n in period *t*-1 had its demand been reduced to zero. This Hicksian reservation price would have allowed the inclusion of welfare effects from the introduction of new models. However, our analysis here is concerned with sampling issues. Ideally the weights should relate to sales and not the number of observations, and the sales weights should correspond to a superlative formulation as outlined in Fixler and Zeischang (1992), Kokoski et al. (1999a and Silver (1999 and 2001).

(v) Differences between estimates from matched data and matched and unmatched data: leverage effects

We develop here the analysis of the effects of excluding unmatched data on the measurement of price changes based on leverage effects. Consider the case where for simplicity the coefficients are constant between the two periods *t* and *t+n*. The only difference between their hedonic surfaces is an intercept shift reflecting an overall price change. The dummy variable method works well in such circumstances and we assume again that the quite simplistic functional form given by equation (3) is appropriate – though there is a case for more flexible forms (Curry, 1999 and Diewert, 2001). Assume a full data set of observations are available which includes (I) those observations that might be matched in periods *t* and *t+n*, (ii) those available in period *t* but not *t+n* (unmatched old *t*) and (iii) those available in period *t+n* but not *t* (unmatched new *t+n*).

The concern here is that the unmatched observations may not lie on the matched hedonic surface given by equation (3). Bear in mind that equation (3) has two effective surfaces, one for period *t* and one for period *t+n* with an intercept shift. The concern is therefore that the unmatched period *t* observations do not lie on the former and the unmatched period *t+n* do not lie on the latter. Berndt (1995), Silver (1999), Triplett (2001) and Pakes (2001) have amongst others provided reasons why this might be the case. Unmatched period *t* models are near the end of their life cycle and unmatched period *t+n* near the start of theirs, and economic and marketing theory and practice has shown how pricing varies over the life cycle of a products and a brand's differentiated varieties. So what is the effect on our estimates of including such additional data?

This is not a difficult matter to investigate empirically. For an unmatched observation to have an influence on the hedonic results it needs to have unusual characteristics, say due to its technology have a relatively slow spin speed for a washing machine combined with being energy inefficient. However this may still have a price that can be explained by the regression – it can lie on the hedonic surface. To lie off the hedonic surface its price must be unusual given its characteristics, say due to high price skimming for a new model or low price dumping of an old one, i.e. the residuals should be high. To examine such issues we borrow from the econometrics of leverage effects using Davidson and McKinnon (1993). They consider the estimation of an equation with influential observations.

It is first noted that an OLS vector of $\mathbf{b}$ estimates is a weighted average of the individual $\mathbf{y}$ elements – the prices of individual models where:

$$\hat{\hat{\mathbf{a}}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \tag{6}$$

and $(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ the weights given to the prices with some observations likely to have more influence than others. If there are a number of such unusual observations belonging to a different data generating process, the larger the number of such observations the more influence they will have. Our concern is with the effect of adding a, for simplicity, single new unmatched observation to the regression estimate in period *t,* though via equation (6), the extension to several new observations is straightforward and considered later.

Following Davidson and McKinnon (1993) we compare $\hat{\hat{\mathbf{a}}}$ with $\hat{\hat{\mathbf{a}}}^{(t)}$ where the latter is an estimate of $\mathbf{b}$ if OLS was used on a sample omitting the period t unmatched observation, hereafter - the t^{th} observation. We distinguish between the leverage of the t^{th} observation, h_t and its residual, $\hat{u}_t$. An influential observation may, for example, have high leverage, that is influence on at least one element of $\hat{\hat{\mathbf{a}}}$, but a smaller impact on $\hat{u}_t$, or it may have little leverage but have a high residual. The leverage for observation t is given by:

$$h_t = \mathbf{X_t}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X_t}^T \quad \text{where } 0 \le h_t \le 1 \tag{7}$$

and the difference between the hedonic coefficients with the t^{th} observation omitted and included given by:

$$\hat{\hat{\mathbf{a}}}^{(t)} - \hat{\hat{\mathbf{a}}} = -\left(\frac{1}{1-h_t}\right)(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X_t}^T\hat{u}_t \tag{8}$$

When $\hat{u}_t$ is large and/or h_t is small the effect of the t^{th} observation on at least some of $\hat{\hat{\mathbf{a}}}$ is likely to be substantial. It follows that including the t^{th} observation in the regression affects the fitted value for that observation by:

$$\mathbf{X_t\hat{\hat{a}}} = \mathbf{X_t\hat{\hat{a}}^{(t)}} + \left(\frac{h_t}{1-h_t}\right)\hat{u}_t \qquad (9)$$

and therefore the change in the t^{th} residual by including the t^{th} observation by:

$$-\left(\frac{h_t}{1-h_t}\right)\hat{u}_t \qquad (10)$$

Naturally if a dummy variable was added to the model, which took the value of 1 for the unusual observations and zero otherwise, then $\hat{\hat{\mathbf{a}}} = \hat{\hat{\mathbf{a}}}^{(t)}$. It can be shown that h_t must on average equal k/n where there are k explanatory variables and n observations. If all h_t were equal to k/n then every observation would have the same leverage. The matched sample is not equivalent to this *balanced design*, in which the only regressor would be a constant, but it is an attempt to be close to it. The quantity h_t may be used to measure potential leverage of an observation, a value greater than say, $2*k/n$ denoting high leverage. However it can be seen from the above equations that when h_t is large, dropping observation t will only have a relatively large effect on $\hat{\hat{\mathbf{a}}}$ if $\hat{u}_t$ is very close to zero. The leverage h_t only has a *potentially* large effect and whether this is realised depends on $\hat{u}_t$ and thus y_t.

We can thus explore on an empirical basis the values of $\left(\frac{h_t}{1-h_t}\right)\hat{u}_t$ and h_t for both matched and unmatched observations and identify whether or not the unmatched observations are unduly influential and whether such influence is realised. In the context of CPI construction the analysis is of whether the characteristics of the unmatched observations are sufficiently different to give them the potential to allow them to have an undue effect on the hedonic estimates. And, if so, whether the associated prices are not aligned with what would be expected from the hedonic relationship, say due to price skimming or dumping, to allow large residuals to be created which will realise the potential of the leverage. The above statistics will provide measures of these effects.

(b) Missing models

A second reason why the matched models method may fail is that when models in subsequent periods are missing, no record can be made of their prices. They may be obsolete or no longer stocked by the establishment. Statistical offices employ the matched models method to ensure quality changes do not affect measured price changes. It is only when there are missing prices that the quality adjustment problem potentially arises. Any bias in the measurement of inflation through its inability to properly incorporate quality changes results from such missing prices (Boskin *et al*., 1996 and 1998; Diewert, 1996; Cunningham, 1996; Hoffmann, 1998; Abraham *et al*., 1998). A number of procedures are available in such circumstances:

- the models may be dropped on the assumption that the aggregate price changes of other models reflect those of the missing models – an implicit quality adjustment;

- prices on both the missing and replacement models may both be available in an overlap period before the model was missing; the price difference in this overlap period is used as a measure of the quality difference;

- the price collector may select replacement models whose quality is deemed to be similar and continue to use their prices instead;

- replacement models whose characteristics differ from the unavailable ones may only be available and explicit adjustments of quality differences may be used to extricate the 'pure' price and quality changes.

The quality adjustment problem arises because many of the more readily available and practically used methods are not always satisfactory. This is not the subject of this study, though see Turvey, 1989, Reinsdorf *et al*. (1995), Moulton and Moses (1997), Armknecht *et al*. (1997), Moulton *et al*.,(1998), Lowe (1999), Kokoski *et al*. (1999), Silver and Heravi (2000) and Triplett (1990 and 2001).

(c) Quality adjustment methods and their effect on the sample space

There are a number of interrelationships between quality adjustment procedures and the active sample space. The first concerns the use of imputations of similar price changes for existing and missing varieties. The second concerns selection of models for forced replacements. Such models may be 'comparable' requiring no quality adjustment the price difference between the old and new model being taken as a pure price change. Or they may be 'non-comparable' with an explicit adjustment being required to separate out the pure price change by either an option cost, subjective judgement, quantity or hedonic adjustment.

(i) <u>imputation</u>

Imputations involve effectively dropping the missing models from the sample; they rely on assumptions of similar price changes for quality-adjusted missing models, were they available, and the active sample. Class means or targeted imputation relates this assumption to a more specific class of items (see Armknecht and Maitland-Smith, 1999, Feenstra and Diewert, 2001 and Triplett, 2001). Yet, of course this implicitly equates the missing (quality adjusted) price change to the price change in the static universe. Yet what we want is the price change in the dynamic universe. The more representative the static universe, the better the imputations. In practice this will arise when the price change is closer to the base period in which the item sample was last rotated. One implication is that chaining and sample rotation has some merit as devices to refresh the sample. Another is that imputations should be short run as opposed to long-run comparisons. In the former case matched prices in period t are compared with $t+n$. In the latter case matched prices in period t are compared with those in period $t+n-1$, and multiplied by a comparison between matched prices in period $t+n-1$ and period $t+n$. There are two effects. The first is that assumptions of similar price changes are more likely to hold over the short-run and the second that the sample space is larger. This is particularly true for the rotated sample and chaining, but also holds for short-run price changes. The comparison of say prices for 6 models in January may be depleted by one a month to 3 in April. Yet instead of having a sample space of three comparisons for the long-run January:April, we have 4 for January:March and 3 for March:April.

(ii) Comparable replacements

The selection of comparable replacements by price collectors when a model is missing puts the coverage of the sample to some extent under the control of the price collectors. Replacement models should intrude into the universe of transactions in a substantial and representative sense, to be more representative of the dynamic universe. It is of little merit to substitute a new model with limited sales for a missing model, just because they have similar features, both being 'old'. It is bad enough that the matched models method encourages price collectors to hang on to the model until it dies. It compounds the misdemeanour when guidelines encourage its replacement with a similar old variety with low sales - as Lane (2001:21) notes for the U.S. CPI: "..when the outlet finally discontinued selling the item the regular replacement rules require the CPI field agent to replace it in the sample with the most similar thing that the store still sells. This could be something almost as obsolete." The motivation behind this is of course to ease quality adjustment between the old and new item, so that the two models are similar. The institutional mechanism devised to help quality adjustment can lead to bias due to their adherence to a sample of models which do not enjoy the benefits of recent technological innovations and are unrepresentative of what we buy. Quality adjustment and representativity are interrelated since the former affects the sample space of the index.

(d) New products

As noted above welfare gains from new product cannot be readily measured by hedonic or the matched model approach. There are also many sources of quality change such as reliability and 'ease of use' which may be difficult to quantify, especially when producing CPIs in real time. We do not consider these aspects of bias in this study.

4. DATA AND IMPLEMENTATION

(a) Data: Scope and coverage

The study is for monthly price indices 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 includes 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 analysing market share and pass on such data to market research agencies for analysis. By cumulating these records for all outlets (supplemented by visits to independent outlets without scanners) the agencies can provide, on a monthly basis, comprehensive data, for each model for which there is a transaction, on: price (unit value), volume of sales, quality characteristics, make, and outlet type. There is a reluctance for them 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, governmental and other bodies to check on anticompetitive pricing. Data are however identifiable by broad types of outlets and models codes often apply to specific outlets, though they are not identifiable.

It should be stressed that the data, unlike that collected by price collectors:

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

The data are not without problems in that the treatment of multi-buys and discounts varies between outlets and the coverage varies between product groups. For example, items such as cigarettes sold in a variety of small kiosks are problematic. Nonetheless, they provide a recognised alternative, first proposed by Diewert (1993) and used by Silver (1995) and Saglio (1995), though see also, for example, Lowe (1998), Moulton, LaFleur and Moses (1998); as Astin and Sellwood (1998 p297-298) 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 [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; 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 catalogue. We stress that we differentiate models as being sold in different types of outlets. 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 both in terms of transactions and features. For the in 1998, there were 1.517 million transactions of washing machines involving 7,750 observations (models/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.

(b) Data: variables

The variable set includes:

<u>Price</u> **-** the unit value = value of sales/quantity sold of all transactions for a model in an outlet type in a month.

<u>Volume</u> is 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/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 – each only sold 1 machine in 1998. There were 1,684 observations (models in outlet types) selling between 2 and 10 machines in a month (on average) selling about 8 thousand machines: so far a total of 2,407 observations managing a sales volume of about 8,800. Yet the 12 models achieving a sales volume of 5,000 or more in any outlet/month accounted for 71,600 transactions.

Vintage is the year in which the first transaction of the model took place. With durable goods models are launched (usually) 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 which are part of the new model. New models can coexist with old models; 1.1787 million of the about 1.517million washing machines sold in 1998 were first sold in 1997 or 1998 – about 77.7% leaving 22.3% of an earlier vintage coexisting in the market.

Makes: transactions occurred in 1998 for machines of 24 different makes. The market was, however, relatively concentrated with the three largest selling (by volume) makes accounting for between about 60% of the market. Hotpoint had a substantial 40% of sales volume in 1998. This was achieved with 15% of models (observations). Zannusi, Hoover and Bosch followed with not unsubstantial sales of around 10% each by volume.

The characteristics set includes:

Type of machine: 5 types – top-loader; twin tub; washing machine (WM) (about 90% of transactions); washer dryer (WD) with and without computer;

WD with /without condensors (about 10% with);

Drying capacity of WD – a mean 3.15kg and standard deviation of 8.2 KGs for a standard cotton load;

Height of machines in cms - about 90% of observations being 85cms tall;

Width - 94% being about 60cms. Depth - most observations taking values between 50 and 60 cms inclusive;

Spin speeds: 5 main - 800rpm, 1000rpm, 1100rpm, 1200rpm and 1400rpm accounting for 10%, 32%, 11%, 24%, and 7%v of the volume of sales respectively.

Water consumption which is advertised on the displays as "..not a measure of efficiency since it will vary according to the programme, washload and how the machine is used." It is highly variable with a mean of about 70 litres and standard deviation of 23 litres;

Load capacity is another such measure for"…a maximum load when loaded with cotton" - a mean about 50Kgs with a standard deviation of about 13 Kgs;

Energy consumption (kWh per cycle) is"…based on a standard load for a 60 degree cotton cycle - a mean of about 12kWh with again, a relatively large standard deviation of about 6kWh.;

Free standing, built-under and integrated; built-under not integrated; built-in and integrated.

Outlet-types include multiples, mass merchandisers, independents, multiples.

(c) 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. However, it is hoped it will be useful for experimental purposes. We start by taking a January fixed basket of washing machines comprising all varieties for which there was a transaction in January. It is assumed that statistical offices have unlimited resources to examine all transactions. Our varieties are for a model in one of four outlet types; multiples, mass-merchandisers, catalogue 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 for illustration Table 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 were collected from different outlet types, multiples and mass merchandisers in this example.

[Table 1 about here]

It should be borne in mind that the price quotes may be unavailable because they are missing on a temporary – seasonal or out-of-stock- or permanent basis. In the former case our concern is with imputations for the missing prices (Armknecht and Maitland-Smith, 1999 and Feenstra and Diewert, 2000). For permanently 'missing' prices comparable substitutes may be available. If not, non-comparable substitutes may be used with explicit, direct quality adjustments. Such explict, direct methods are preferred, though often imputation techniques similar to those used for temporary missing price quotes are used. In this study, using scanner data, all missing prices will be treated in the same way irrespective of whether they are permanently or temporarily missing. At first it will be assumed that implicit indirect overall mean imputations are used – a widely used procedure that simply assumes prices of missing unmatched models follow those of the rest of the sample. The sample thus becomes degraded. The use of explicit methods will be considered later to counter any claim that we are unfairly comparing matching by using an 'inadequate', though widely used, form of imputation for our comparisons with hedonic methods.

Missing varieties are not a trivial matter. Moulton *et al*. (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%) were replacements of which, in turn, 680 (42%) were judged to be not directly comparable. Canadian experience for TVs over an almost identical period found 750 of the 10,050 (7.5%), to be replacements of which 572 (76%) were judged to be not directly comparable (Lowe, 1998). For international price comparisons the problem is much more severe (Feenstra and Diewert, 2000).

The experimental framework and initial results were outlined in Silver and Heravi (2000). The study demonstrated how scanner data might be used to simulate CPI practices to help judge the veracity of alternative quality adjustment procedures. The concern of this study with the static sample is quite different.

5. SAMPLE COVERAGE

The active sample is first considered for the long-run matched Laspeyres. For a matched comparison between January and December, the coverage is the sample in January which can still be matched in

December. Table 2 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, catalogue and independents) of washing machines accounting for 126,171 transactions. The distribution was highly skewed with the top 5% and 10% of varieties (in an outlet type) accounting for 49% and 66% of transactions respectively in January. Table 2 shows that by December, only 53% of the January basket of varieties were used for the December/January index, though these accounted for 81.6% of January expenditure. Varieties with lower sales values dropped out quicker. However, the remaining 0.53 (500) = 265 varieties in December only accounted for 48.2% of the value of transactions *in December*. The active sample relating to the universe of transactions in December had substantially deteriorated.

[Table 2 about here]

The short run modified Laspeyres is an alternative, though not widely used formula (Armknecht and Maitland-Smith, 1999). It combines the long run comparison between the base period and the preceding month and the short run comparison between the current month and its preceding one. For a comparison between January and December, for example, the short run comparison is first based on the long run January to November comparison using 54% of January observations accounting for 83.5% of January expenditure, though only 49.4% of November expenditure. And second on the November to December comparison, which uses 47.9% of December expenditure. The short run comparisons always make slightly less intrusion into current expenditure than the long run ones. This is because the long run comparison requires price data to exist for January and December, while the short run one comparison is still based on the active January sample, but also requires November and December information on prices.

Imputations based on short-run Laspeyres or with item rotation on a biannual, quarterly or monthly basis benefit not only from the more likely veracity of assumptions of similar price changes, but also because of the larger sample space upon which they rely. There is thus an increased likelihood that they replicate the price movements of the dynamic sample. The focus here is on the effects of sample coverage, returning in section 8 to the results for the indices. Table 2 presents results on coverage for Laspeyres long run comparisons using overall mean imputation with re-weighted sample rotation conducted on a biannual, quarterly and then monthly basis. The coverage relates to the percentage of the current month's expenditure captured in the matching of prices between the base and current period. For example in the biannual comparison in October (compared with June 1998) 75.25% of expenditure in October was covered by the matching procedure, the remaining being implicitly imputed using the overall mean. The use of biannual sample rotation improves the coverage of the matching to at worst, a little over 70%, compared with 48.2% in Table 2 when rotating annually: a substantial improvement. Table 2 shows the upgrading of the rotation to a quarterly basis to further improve coverage to an at worst, 76.71% (September) and annual chaining to 83.33% (July). The average coverage over the 12 months for the biannual, quarterly and monthly chained procedures were 73.5%, 79.8% and 86.8 % respectively compared with 48.2% for the annual sample rotation.

6. MATCHED AND UNMATCHED COMPARISONS

(a) Characteristics of matched and unmatched models for long run comparisons

Having established that long run matched and hedonic indices can in principle diverge, because of the limited coverage of the former (section 3). And having further established that the differences in coverage in practice can be substantial (section 5). And bearing in mind that this was undertaken by recourse to the special nature of scanner data that allows us to examine the universe of transactions. We now ask whether this depletion in coverage matters? Is it the case that the prices and characteristics of these matched and unmatched models differ and, if so, do they have an effect on the indices? It may be that the prices of new models entering the market after the base period – hereafter unmatched 'new' - command a higher price, being at the initial stage of their life cycle. But incorporate new technology and have a lower quality adjusted price. It may also be the case that prices of models existing in the base period and then exiting the market – hereafter unmatched 'old' – have relatively low quality adjusted prices, say because they are being dumped to make way for new entrants. Alternatively they may coexist with the new model and continue to serve a special segment and even have a higher price than their quality merits (Berndt, 2001 and Pakes, 2001). Table 3 provides some summary statistics on the price, number, value, vintage and spin-speed of matched and unmatched new and old models.

[Table 3 about here]

Table 3 clearly shows how prices differ between matched and unmatched observations. The mean January and current price of matched models were 440.57 and 424.35 pounds sterling respectively, the current price of a matched model falling over time. However, the mean price for old models, those available in January, but exiting in the current month, was lower than the January price, but higher, at 435.60, than the matched current price. The former is as expected but the latter less so and we will revisit this issue in a more analytical framework. The mean price of new entrants is of course much higher, at 484.94, than existing matched prices. They are also newer, their mean launch date being half-way through 1996, as opposed to half-way through 1995 for matched models and nearly a year earlier for old exits. New entrants are also more advanced than matched ones in terms of their spin speeds (1118rpm compared with 1097rpm) and old exits are technologically inferior in this sense (1083 rpm). The different make-up or characteristics of the matched and unmatched models of course influence the above descriptive statistics on price. The growth in the number of new models again illustrates how the matched models method can lose out – their being 399 models in December not on the market in January, and thus not in the matched sample selection, compared with only 39 new models launched in February. The potential for error increases as we move away from the base month.

(b) hedonic, quality-adjusted price differences between matched and unmatched models.

The comparison between matched and unmatched prices in Table 3 takes no account of the differential qualities of the models being compared. Hedonic regressions were estimated on the whole sample in each month. For each month a model's price (or its logarithm) was regressed on the characteristic set, outlet types and brand dummies described in section 4(b) - over 40 variables in all. The variable list was given in section 4(b) above and in the illustrative regression in Table 8, though this regression serves a different purpose. Also included were two dummy variables. The first took the value of 1 if the observation in that month was an unmatched *new* model, and zero otherwise, while the second took the value of 1 if the observation in that month was an unmatched *old* model, zero otherwise. Separate regressions were estimated for each month using linear and semi-logarithmic formulations and OLS and sales volume WLS. The *t*-statistics on these 'old' and 'new' unmatched variables provide test statistics on the null hypothesis of no difference in these matched and unmatched prices in the month in question. The sign and magnitude of the coefficients provide information, if statistically significant, on the nature and magnitude of their difference from matched models.

[Table 4 about here]

In Table 4 the results are only presented for the dummy variables on the new and old variables – though the regressions included all variables (full regression results are available from authors). Tables 3 and 4 show the mix of the sample: for example for the 1,047 models in August, 363 were unmatched new and 152 unmatched old. The $\overline{R}^2$ in August for the regression equation with the full 40 plus variable set was 0.81; the coefficient on the dummy variable for the old unmatched observations in this hedonic regression was an estimated –0.074, an about 7.4 per cent difference from their matched (omitted benchmark) counterparts, the difference being statistically significant at a 5 per cent level (Table 4). It is apparent from Table 4 that there is a universal negative impact for old unmatched observations in each month, generally statistically significant at the 5 per cent level. Hedonic prices of old unmatched models are on average lower than matched existing ones. The average difference (OLS semi-logarithmic) is 6.1 per cent lower than matched ones. We follow here *Teekens, R. and Koerts, J. (1972) in adding* one half the squared standard errors to the coefficients of these semi-logarithmic regression coefficients to correct for bias – though subsequently omit this to allow comparability between the text and the tables. The standard errors are in any event small and can be derived by the reader via the coefficients and *t*-statistics.

There is also evidence of a *positive* difference for unmatched new models, though it is less clear. The OLS semi-logarithmic results found only two significant coefficients at the 5 per cent level and they were both positive. Yet the WLS counterpart had positive results for all but one month, with positive differences from zero statistically significant for *six* months. New entrants seem to have higher prices than matched ones. These new unmatched ones are the dynamic sample CPI compilers never see because they are outside of the coverage of the initial selection. Note that the differences are after the influences of quality, outlet and brands have been controlled for in the regressions.

Having established that the quality-adjusted prices of new and old unmatched and matched observations differ for individual months, the next question is whether these differences matter with regard to the measurement of quality adjusted price changes?

(c) differences in hedonic, quality-adjusted price changes between matched and unmatched models.

Hedonic regressions akin to equation (3) were estimated. These were undertaken for data for January and February, then again for January and March, and continuing for January with September – 11 regressions in all. Each regression had the full variable set as illustrated in Table 9 and discussed in section and were estimated using four different data sets:

- matched data only;
- matched plus unmatched old and unmatched new;
- matched plus unmatched old;
- matched plus unmatched new.

Included in each regression was a dummy variable which was 1 for observations in the current period, for example for the January and December comparison/data, it was 1 for December, and zero otherwise. The coefficient is an estimate of the quality adjusted price change. There were three estimates in each month to identify whether the estimate of quality adjusted price changes differ according to the inclusion of unmatched new and old models. Any such differences would depend on the changes between the means for the respective matched and unmatched quality adjusted price changes and the weight given to such changes. For OLS estimates the weights would be the number of observations, though weighting systems based on sales would follow from a WLS estimator. Reminding ourselves of equation (5) for matched *m* and unmatched old *o* and new *n* models:

$$\ln p_t/p_{t-1} = \Big[m/(m+n)\sum_m (\ln p_{mt} - Z_m)/m - m/(m+o)\sum_m (\ln p_{mt-1} - Z_m)/m\Big]$$

$$+ \Big[n/(m+n)\sum_n (\ln p_{nt} - Z_n)/n - o/(m+o)\sum_o (\ln p_{ot-1} - Z_o)/o\Big] \qquad (11)$$

If the data are matched only, *n* and *o* are zero, equation (11) collapse to the ratio of the geometric means of the matched sample in the two periods. If there are only matched *m* and unmatched new *n*, *o* = 0, with some simple algebra:

$$\ln p_t/p_{t-1} = m/(m+n)\Big[\sum_m (\ln p_{mt} - \ln p_{mt-1})/m\Big]$$

$$+ n/(m+n)\Big[\sum_n (\ln p_{nt} - Z_n)/n - \sum_m (\ln p_{mt-1} - Z_m)/m\Big] \qquad (12)$$

and for only matched *m* and unmatched old *o, n*=0:

$$\ln p_t/p_{t-1} = m/(m+o)\Big[\sum_m (\ln p_{mt} - \ln p_{mt-1})/m\Big]$$

$$+ o/(m+o)\Big[\sum_m (\ln p_{mt} - Z_m)/m - \sum_o (\ln p_{ot-1} - Z_o)/o\Big] \qquad (13)$$

[Table 5 about here]

Table 5 provides the results. Both linear and semi-logarithmic functional forms were used given their prevalence in the literature (discussed in Triplett, 1988, Griliches (1990), Triplett (1990), Gordon (1990), Arguea *et al*., 1994, and Berndt *et al,* 1995). OLS and weighted least squares (WLS) estimators were used on both forms, the WLS estimator being weighted by sales volume. Sales value weights were also used and similar results resulted not reported here but available from the authors.

Table 5 contains the results for the single dummy variable on the current month from the 11 months x 4 data sets x 2 functional forms x 2 estimators (OLS and WLS) = 176 estimated regression equations.

The regression equations fitted the hedonic models well by the usual criteria, including a mean $\overline{R}^2$ of 0.85 with a standard deviation of a mere 0.0022, a minimum of 0.75 and maximum of 0.91. It is reiterated that the results are from fully specified regression equations each using about 40 variables as discussed in section 4(b), though only the coefficients on the time dummy are presented here. The detailed results are available on request. Equations (12) and (13) show that if new models in period *t* are higher priced than matched ones in period *t*-1, and old ones in period *t*-1 lower priced than matched ones in *t*, then the matched results will fall faster. Equation (11) however, shows that for *both new and old models*, the joint effect is more than the individual effect; the combination of higher new prices and lower old ones will lead to a larger fall in the matched results compared with the unmatched ones. For example, if the average quality adjusted prices of new models in period *t* was 120 for old models in *t*-1 80, and matched models in both periods 100, then the price change resulting from including both new and old unmatched models would be higher than if just new and just old were included. Equation (11) shows how the effect of including new and old unmatched models depends on their respective prevalence in periods *t* and *t*-1. The method can seriously understate (overstate) inflation (deflation) in a manner not previously recognised. The nature and extent of the bias can be seen to relate to the pricing policies of firms regarding old and new models and the patterns estableshed here need not be indicative of all markets. Berndt (2001) for example has shown how the prices of old, branded, pharmaceutical drugs can increase after the expiration of a patent and launch of new generic models.

The results in Table 5 show first, that the quality adjusted prices fell faster for the matched sample, than the matched and unmatched samples, this result holding for both functional forms and all estimators (though there was a single exception – August semi-log by volume). Second, the differences were marked. By December 1998 prices fell by 9.9 percent for matched data compared with 6.7 per cent for all data. The exclusion of unmatched data seriously overstates these price falls.

Third, the matched plus unmatched new *and* old does not fall as fast as the unmatched plus just new or unmatched plus just old. The combined effect of lower old prices and higher new prices, when weighted in various ways by the estimators, leads to lower overall falls. The differences can be substantial. For example, the weighted by volume semi-logarithmic OlS estimator in Table 5 finds falls of 6.7 percent for all data, compared with falls of 9.9 percent for matched, 7.0 percent for matched old and 8.9 percent for matched and new.

Fourth, when sales weighted estimators are used, the falls in prices generally becomes less marked overall. This pattern hold for matched models and matched plus new models. More popular matched models and unmatched new models in period *t* have higher than average prices while more popular matched periods in *t*-1 have lower relative prices.

The results confirm the differential quality adjusted prices and price changes of unmatched and matched models. They bring in the implicit weighting given to the old and new models apparent from equations (11), (12) and (13) and given via the number of observations in Table 3. For example, Table 3 shows that for the January to December comparison there were 1,013 prices in each month for matched models, 399 prices of new models in December and 235 for old models in January. The weighting for new and old models in December and January respectively using equation (11) are 399/(399+1,013) and 235/(235+1,013). The use of the hedonic framework thus allows a comparison of how unmatched models influence overall hedonic price changes.

(d) residuals and leverage

It was noted in section 3, that the unusual unmatched observations in a regression can take two forms. First they can have high residuals from the hedonic surface. That is, given their characteristics, the actual prices lies away from what would be predicted by the hedonic model. The second is leverage, the diagonal of the 'hat matrix' outlined in section 3.

Table 6 provides the average leverage and residuals for matched and unmatched observations. Separate regressions were undertaken for each month using here the semi-logarithmic hedonic form and a WLS by sales volume estimator. The mean and standard deviation of the residuals and leverage indices (see section 3) were calculated on a (volume) weighted and unweighted basis. It is clear from Table 6 that residuals from unmatched new models are higher than matched ones, while residuals from unmatched old models are much lower. The residuals generally have high standard deviations. The results for leverage are also very clear. Unmatched observations have nearly twice the (unweighted) leverage than matched ones – it is not just that they have high residuals not falling on the hedonic surface, but their influence in the estimation of the parameters of the regression equation is much greater, and their exclusion more serious.

[Table 6 about here]

7. STRATEGIES: CHAINING AND HEDONICS

There are two alternatives worth considering for products with rapid turnover of models. They both involve abandoning the use of matched models and have price collectors sampling from all models each month, or taking data from catalogues, web sites, retailer lists and market research agencies. The hedonic aproach uses all the data in each month as outlined above. An alternative is to refresh the sample by updating it biannually, quarterly or preferably, monthly. The use of monthly, chained indexes has been championed by Turvey (1999) more recently and used in Silver and Heravi (2001). It does not use all the data. For example, in Table 1 a chained index for January compared with March would be made from the product of two links: January and February and February and March. In the first case models 1,2,5 and 6 would be used, but not p_{42}; for the February to March link models

1, 4 and 5 would be used, but not p_{22}, p_{33}, p_{62} and p_{73}. . There is thus a preference for hedonics since each observation is used in the estimation. Furthermore, a chained index measures something quite different to, and is not strictly comparable with, a hedonic index as undertaken here. The chained indices are path dependent while fixed base CPI methodology for these within year monthly series are not. Hedonic indices as undertaken here compare prices in January directly, for example, with those in October, though have the advantage of being able to take a chained form built up from binary comparisons, if required. Table 7 provides results for indices with samples rotated on a biannual, quarterly and chained monthly basis.

It was noted in section 2(b) that hedonic equations could also be used for direct, explicit imputations or 'patching' of missing old unmatched prices. Triplett (2001) provides a thorough account of such approaches. We estimated hedonic regressions in each month and identified when a model when missing and its nearest replacement using the following routine. W*ithin each outlet type* in each month a search was made for the best match first, by matching brand, then in turn by type, width and spin speed (see section 4(b)). If more than one variety was found, the selection was according to the highest value of transactions (expenditure). In our example in Table 1, varieties 1, 3 or 4 and 5 or 7 would replace varieties 2 and 6 in March respectively. Any differences between the characteristics of the new replacement model and the old were ascertained. The subset of the *k* characteristics that distinguished the old and replacement new models were ascertained and the coefficients from the hedonic equations were applied to the difference in the values of the characteristics to first, correct the new model's price to adjust it to the old model's characteristics. This adjusted new price in the current period was compared to the old price in the base period. And then a similar procedure was undertaken to correct the old models' price to make it comparable to the new one. A geometric mean was taken of the two resulting estimates and used for the models' price change (see Silver and Heravi, 2000 and Triplett, 2001 for details). The results are given in Table 7. This allows our hedonic index to be compared with best practice matching with hedonic adjustment (patching), the latter benefiting from some limited updating of the sample by way of the forced, quality adusted replacements. The results of Table 7 will be discussed in the next, summary section.

[Table 7 about here]

9. CONCLUSIONS

The results are exploratory in the sense that they arise from an experimental formulation. 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, though 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 and de Haan and Opperdoes, 1998). However, the concept of 'missing' prices used here is not completely appropriate since a price collector may, for example, find a price missing for a variety in an outlet in one city, while other price collectors may find price quotes for the same variety in different stores/locations. The experiment would only treat prices as missing if there were no transactions anywhere for the product variety. Furthermore all missing prices were treated in the same way irrespective of whether they were permanently or temporarily missing. Scanner data does allow a search to see if the variety returns, though our data is aggregated at the outlet type level and 'missing' in our sense refers to no further transactions being conducted for that model in one of four store types.

The matched models approach can fail because of its use of a static sampling universe. Because the matching excludes unmatched new and unmatched old models from the sample, while hedonic estimates can include them. Long run Laspeyres comparisons over the period of 12 months were found to seriously degrade the active base period sample and its coverage of the current population of transactions. There are a number of responses to this. First, biannual, quarterly and monthly sample rotation using imputations did much to improve the coverage of the current new transactions, though not the old ones. Though, second, the use of the short run, modified Laspeyres did little to ameliorate the poor coverage of the sample. Finally, the use of hedonic indices has the advantage of using the full sample.

The above was concerned with the degraded coverage of the dynamic universe and response to limit the degradation. However, this degradation is only a necessary condition for bias. Matched models indices will only differ from hedonic indices if prices of unmatched new and unmatched old models differ from matched ones. The prices of unmatched and matched models were found to differ, as were their vintage and quality. Even when quality adjusted using hedonic regressions, prices of unmatched old models were found to be lower than matched ones, there also being evidence of higher prices for unmatched new models. These new unmatched ones are the dynamic sample CPI compilers never see because they are outside of the coverage of the initial selection. The next question was whether these differences mattered with regard to the measurement of quality adjusted price changes?

Quality adjusted prices fell faster for the matched sample, than the matched and unmatched samples. The differences were substantial. The exclusion of unmatched data seriously overstated these price falls. We repeat some of these results in Table 7 for clarification. The simple geometric mean and semi-log OLS estimator using unmatched data naturally provide similar results of an about 10 per cent fall. Similar results arose when 'best practice' explicit hedonic adjustments to forced, non-comparable replacements for missing old unmatched models. When weighted estimators were used, the falls in prices generally become less marked at about 8 per cent. This was similar to the result from a chained matched Laspeyres, which also takes account of the weights. Yet when all the data were used – matched and unmatched –we have falls in prices of under 7 per cent without weights and under 5 per cent with weights. As predicted from equation (5) the combined effect of the relatively low unmatched old models and high priced unmatched new models combine to limit the fall in prices of the matched models.

The analysis of unmatched new and old models and matched models was further developed to see if any differences arose from their residuals from a common hedonic surface and/or their leverage. The residuals from unmatched new models were higher than matched ones, while residuals from unmatched old models were much lower. Unmatched observations had nearly twice the (unweighted) leverage than matched ones - their influence in the estimation of the parameters of the regression equation was much greater, and their exclusion more serious.

The matched models method may also fail because of the effect the quality adjustment procedures used for dealing with missing values have on the sample space. The procedures used for selecting comparable replacements were argued to have possible serious adverse effects. One way round this is to rotate or update the sample more regularly, or use the short-run modified formulation. Biannual rotation provided similar indices to the long run Laspeyres, though with quarterly and monthly (chained) rotation the results were more favourable. Yet the disparities between the results of the hedonic indices, and the matched, chained matched and matched with hedonic adjustments on forced replacements, clearly shows the need for including a sample of unmatched new and old prices when compiling indices. This study has shown why the hedonic approach provides a useful approach to meet this need and the matched models approach is ill suited to today's dynamic markets.

References

Abraham K, Greenless J and Moulton, B (1998) Working to Improve the Consumer Price Index, *Journal of Economic Perspectives*, 12 (1), 23-36.

Aizcorbe A, Corrado, C and Doms M (2000) Constructing Price and Quantity Indexes for High Technology Goods, Industrial Output Section, Division of Research and Statistics, Board of Govenors of the Federal Reserve System, July.

Arguea, N M, Haseo, C and Taylor, GA, (1994) Estimating Consumer preferences Using Market data: An Application to US Automobile Demand. *Journal of Applied Econometrics*, vol. 9, pp. 1-18.

Armknecht, PA (1996) Improving the Efficiency of the US CPI, International Monetary Fund *Working Paper*, 96/103.

Armknecht, PA, Lane, WF and Stewart KJ (1997). New products and the US Consumer Price Index. In *The Economic of New Goods*, *Studies in Income and Wealth*, (eds. T.F. Bresnaham and R.J. Gordon), Vol.58, pp. 375-396, NBER Chicago and London: University of Chicago Press.

Armknecht, PA and Maitland-Smith, F (1999), Price Imputation and Other Techniques for Dealing with Missing Observations, Seasonality and Quality Changes in Price Indices, *IMF Working Paper 99/78*, Statistics Department, International Monetary Fund, Washington, D.C.

Astin, JA and Sellwood DJ (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, edited by B. Balk. Research Paper no. 9806, Statistics Netherlands, 291-308.

Balk, B (1998) 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), pp.112-120, Washington DC: Bureau of Labor Statistics.

Baxter, M (editor), (1998) *The Retail Prices Index: Technical Manual.* Office for National Statistics, London: HMSO.

Berndt, ER, Grilches, Z, and Rappaport NJ (1995) Econometric Estimates of Price Indices for Personal Computers in the 1990s, *Journal of Econometrics*, vol. 68, pp. 243-68.

Berndt ER , Ling D and Kyle MK (2001)The Long Shadow of Patent Expiration: Do RX to OTC Switches Provide an Afterlife? Forthcoming in Mathew Shapiro and Rob Feenstra (eds.), *Scanner Data and Price Indexes*, National Bureau of Economic Research, Studies in Income and Wealth, vol. 61 Chicago: University of Chicago Press.

Boskin MS (Chair) 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 DC.

Boskin MS, Delberger ER, Gordon RJ, Griliches Z, Jorgenson DW (1998) Consumer Prices in the Consumer Price Index and the Cost of Living, *Journal of Economic Perspectives*, 12 (1), 3-26.

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, Morgan, P and Silver, M (2001) Hedonic Regressions: Misspecification and Neural Networks, *Applied Economics,* 33, 659-671.

Dalén, J (1998) Experiments with Swedish Scanner Data, *Proceedings* of the Third Meeting of the International Working Group on Price Indices, edited by B Balk. Research Paper no. 9806, Statistics Netherlands, 163-168.

Davidson R and MacKinnon JG (1993) *Estimation and Inference in Econometrics*, Oxford, Oxford University Press.

Diewert, WE (1976) Exact and Superlative Index Numbers. *Journal of Econometrics*, vol. 4, pp. 115-45.

Diewert, WE (1980), Aggregation Problems in the Measurement of Capital. In *The Measurement of Capital*, edited by Dan Usher, Studies in Income and Wealth, Vol 45, National Bureau of Economics Research, Chicago: University of Chicago Press, pp.433-528.

Diewert, WE (1990) The Theory of the Cost-of-Living Index and the Measurement of Welfare Change in *Price Level Measurement* ed. WE Diewert, 79-147, Contributions to Economic Analysis, 196, Amsterdam: Holland.

Diewert, WE (1993), The Early History of Price Index Number Research. In WE Diewert and AO Nakamura (editors), Essays in Index Number Theory, volume I, North Holland: Amsterdam, 33-65.

Diewert, WE (1996) Sources of Bias in Consumer Price Indices, *Discussion Paper No: DP-96-04*, School of Economics, University of New South Wales.

Diewert, WE (2001) Hedonic Regressions: A Consumer Theory Approach. Forthcoming in Mathew Shapiro and Rob Feenstra (eds.), *Scanner Data and Price Indexes*, National Bureau of Economic Research, Studies in Income and Wealth, vol. 61 Chicago: University of Chicago Press.

Feenstra RC (1995), Exact Hedonic Price Indices. *Review of Economics and Statistics*, vol. LXXVII, pp. 634-54.

Feenstra, RC and Diewert, EW (2001), Imputation and Price Indices: Theory and Evidence from the International Price Program, National Bureau of Economic Research, *Working Paper.*

Fenwick, D, Ball, A, Silver, M and Morgan, PH (2000) Price Collection and Quality Assurance of Item Sampling in the Retail Price Index: How Can Scanner Data Help? Paper presented at the NBER/CRIW seminar on Scanner Data and Price Indexes. Forthcoming in Mathew Shapiro and Rob Feenstra (eds.), *Scanner Data and Price Indexes*, National Bureau of Economic Research, Studies in Income and Wealth, vol. 61 Chicago: University of Chicago Press.

Fixler, D and Zieschang, K (1992). Incorporating Ancillary Measures of Processes and Quality Change into a Superlative Productivity Index. *Journal of Productivity Analysis,* vol. 2, pp. 245-267.

Greenspan, A (2001) The Challenge of Measuring and Modeling a Dynamic Economy. Remarks at the Washington Economic Policy Conference of the National Association for Business Economics, Washington D.C., March 27th.

Gordon, RL (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, NBER (eds. ER Berndt and JE Triplett), pp.185-206. Chicago University of Chicago Press.

Haan, J de and Opperdoes, E (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, edited by B Balk. Research Paper no. 9806, Statistics Netherlands, 191-202.

Hausman JR (1997) Valuation of New Goods Under Perfect and Imperfect Conditions. In *The Economics of New Goods*, Studies in Income and Wealth 58, NBER (eds. T Bresnahan and RJ Gordon), pp.209-248, Chicago and London: University of Chicago.

Hicks, JR (1940) The Valuation of the Social Income. *Economica*, vol. 7, pp. 105-24.

Hoffmann, J (1998) Problems of Inflation Measurement in Germany, *Economic Research Group of Deutsche Bundesbank,* Discussion Paper 1/98, Frankfurt: Deutsche Bundesbank.

Ioannidis C and Silver, M (1999) Estimating Hedonic Indices: an Application to UK Television Sets, *Journal of Economics, Zeitschrift fur Nationalokonomie*, 69,1, 71-94.

Koskimäki T and Vartia Y (2001), Beyond Matched Pairs and Griliches-Type Hedonic Methods for Controlling for Quality Change in CPI Sub-Indices: Mathematical Considerations and Empirical Examples on the use of Linear and Non-Linear Models with Time-Dependent Quality Parameters, Paper presented at the 6th Meeting of the (Ottawa) International Working Group on Price Indices, Canberra, Australia 2-6 April.

Kokoski, M, Waehrer, K and Rozaklis, P (1999). Using Hedonic Methods for Quality Adjustment in the CPI: The Consumer Audio Products Component. In *Proceedings* of the Measurement of Inflation conference, (eds. M. Silver and D. Fenwick), Eurostat/ONS/Cardiff University: Cardiff Business School.

Kokoski, Mary F, Brent R Moulton and Kimberly D Zieschang, (1999) Interarea Price Comparisons for Heterogeneous Goods and Several Levels of Commodity Aggregation, in A Heston and R Lipsey eds., *International and Interarea Comparisons of Prices, Output and Productivity*, Committee on Research on Income and Wealth, NBER, Chicago: University of Chicago Press.

Konüs, AA, (1939) The Problem of the True Index of the Cost-of-Living. *Econometrica*, January, 10-29 – first published in Russian in 1924.

Lane W (2001) Addressing the New Goods Problem in the Consumer Price Index, Paper presented at the 6th Meeting of the (Ottawa) International Working Group on Price Indices, Canberra, Australia 2-6 April.

Lowe, R (1998) Televisions: Quality Changes and Scanner Data. In *Proceedings* of the Fourth Meeting of the International Working Group on Price Indices, (ed. W Lane), pp.5-20, Washington DC: Bureau of Labor Statistics.

Moulton, B and Moses, KE (1997), Addressing the Quality Change Issue in the Consumer Price Index, *Brookings Papers on Economic Activity* 0(1): 305-49.

Moulton, B, LaFleur, TJ, Moses, KE (1998) Research on improved quality adjustments in the CPI: the case televisions. In Proceedings of the Fourth Meeting of the International Working Group on Price Indices (ed. W Lane), pp.77-99, Washington DC: Bureau of Labor Statistics.

Pakes A (2001), New Goods, Hedonics, and Price Indices; With an Application to PCs, *Mimeo,* Harvard University and NBER.

Reinsdorf, Marshall, Paul Liegey and Kenneth Stewart (1995), New Ways of Handling Quality Change in the US Consumer Price Index. *Working Paper* No.276, Bureau of Labor Statistics.

Reinsdorf M (1996) Constructing Basic Components for the U.S. CPI from Scanner Data: A Test Using Data on Coffee, *Working Paper* No. 277, Bureau of Labor Statistics, Washington, D.C.

Reinsdorf, Marshall, and Brent R Moulton (1997), The Construction of Basic Components of Cost of Living Indices. In Bresnahan, TF and RJ Gordon, Eds., *The Economics of New Goods*, Chicago: University of Chicago Press, 397-423.

Reinsdorf, M (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), pp.210-235. Washington DC: Bureau of Labor Statistics.

Saglio, A (1995), Comparative Changes in Average Price and a Price Index: Two Case Studies, L.M. Duchare (editor), *Proceedings of the First Meeting of the International Working Group on Price Indices*, Ottawa: Statistics Canada.

Sellwood D (2001) Improving Quality Adjustment in Practice, Paper presented at the 6th Meeting of the (Ottawa) International Working Group on Price Indices, Canberra, Australia 2-6 April.

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.

Silver, M (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.

Silver, MS and Heravi S (2000) The Measurement of Quality-Adjusted Price Changes. Paper presented at the NBER/CRIW Seminar on Scanner Data and Price Indexes and forthcoming in Mathew Shapiro and Rob Feenstra (eds.), *Scanner Data and Price Indexes*, National Bureau of Economic Research, Studies in Income and Wealth, vol. 61 Chicago: University of Chicago Press.

Silver, MS and Heravi S (2001), Scanner Data and the Measurement of Inflation, *The Economic Journal* , 111 June, 384-405

Teekens, R and Koerts, J (1972) Some Statistical Implications of the Log Transformations of Multiplicative Models. *Econometrica*, vol. 40, no. 5, pp.793-819.

Triplett, JE (1988) Hedonic Functions and Hedonic Indices, in *The New Palgraves Dictionary of Economics,* pp. 630-634.

Triplett, JE (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, NBER (eds. ER Berndt and JE Triplett), pp.207-238. Chicago: University of Chicago Press.

Triplett J E (2001) *Handbook on Quality Adjustment of Price Indexes for Information and Communication Technology Products* , OECD Directorate for Science, Technology and Industry, Draft, OECD, Paris

Turvey R *et al.* (1989), *Consumer Price Indices* , Geneva: International Labour Organisation.

Turvey, R (1999) Incorporating New Models into a CPI: PCs as an Example In *Proceedings* of the Measurement of Inflation conference, (eds. M Silver and D Fenwick), Eurostat/ONS/Cardiff University: Cardiff Business School.

Table 1: Illustration of matching and approaches to quality adjustment

Outlet-type	Model	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}

Table 2, Number of observations and expenditure shares used by different methods

				Imputation – no replacements				**Sample rotation**		
				Long run			Short run	Biannual	Quarterly	Monthly chained
	Number of Observations Models in Outlets	Number of transactions thousands	Expenditure, £'000s	% of January observations used	% of January expenditure used	% of current Expenditure Used	% of current Expenditure Used	% of current periods expenditure used	% of current periods expenditure used	% of current periods Expenditure Used
January	500	126,171	45,727							
February	488	111,358	39,938	85.8	97.2	97.1	97.1	97.06	97.06	97.06
March	605	134,049	48,877	83.0	99.1	91.1	89.8	91.06	91.06	91.88
April	625	113,338	41,841	81.0	98.7	81.8	81.7	81.79	94.86	94.86
May	647	112,549	41,623	77.6	98.3	76.6	76.6	76.62	91.11	97.59
June	711	137,070	49,721	78.6	98.0	72.9	72.8	72.90	86.34	96.83
July	744	116,273	42,924	76.0	97.1	64.2	64.2	83.33	83.33	83.33
August	711	122,977	45,139	69.6	93.5	57.5	57.5	80.31	80.31	98.80
September	717	150,422	54,665	66.2	86.4	54.2	54.2	76.71	76.71	98.56
October	695	129,235	47,620	59.8	87.1	51.6	51.6	75.25	86.62	86.62
November	643	124,845	45,196	54.0	83.5	49.4	49.2	75.74	87.09	98.85
December	664	138,634	49,707	53.0	81.6	48.2	47.9	71.08	83.13	97.72

Table 3, Descriptive statistics on matched and unmatched sample

	Matched January Price	Matched current price	Unmatched old (exits) price	Unmatched new (entrants) price	Matched vintage	Unmatched old (exits) vintage	Unmatched new (entrants) vintage
February	436.00	437.59	457.38	493.27	95.4	94.6	95.3
March	438.73	440.37	440.55	485.14	95.4	94.6	95.3
April	441.13	438.48	430.11	504.80	95.5	94.6	95.8
May	441.69	437.76	429.86	508.12	95.4	94.9	96.2
June	444.64	427.30	418.47	507.19	95.4	94.8	96.5
July	440.37	421.45	434.81	492.00	95.5	94.8	96.8
August	441.73	420.48	432.91	479.23	95.5	94.9	96.9
September	441.01	416.62	435.17	480.30	95.5	95.0	97.0
October	439.90	415.82	437.72	474.43	95.6	94.9	97.0
November	439.18	408.00	438.87	462.24	95.6	94.9	97.1
December	441.91	403.95	435.80	447.65	95.7	94.9	97.2
Mean	**440.57**	**424.35**	**435.60**	**484.94**	**95.5**	**94.8**	**96.5**

	Matched spin speed	Unmatched old (exits) spin speed	Unmatched new (entrants) spin speed	Matched No. of models	Unmatched old (exits) No. of models	Unmatched new (entrants) No. of models
February	1095	1076	1097	846	71	59
March	1092	1095	1102	933	85	190
April	1089	1104	1127	953	95	220
May	1094	1086	1128	986	112	259
June	1100	1066	1137	1038	107	318
July	1097	1080	1122	1069	120	364
August	1095	1086	1132	1047	152	363
September	1097	1082	1123	1047	169	386
October	1093	1091	1113	1038	201	396
November	1114	1067	1109	994	230	373
December	1103	1080	1114	1013	235	399
Mean	**1097.1**	**1083.0**	**1118.5**	**996.7**	**143.4**	**302.5**

Table 4, Hedonic regression coefficients for unmatched new and old dummy variables

	Unmatched new			Unmatched old			Total	
	Coefficient	*t*-statistic	No. of obs.	Coefficient	*t*-statistic	No. of obs.	no. of obs.	$\overline{R}^2$
Semi-logarithmic - OLS								
February	0.006	0.17	59	-0.081	2.22*	71	846	0.83
March	0.006	0.28	190	-0.080	2.37*	85	933	0.80
April	0.035	2.04*	220	-0.070	2.26*	95	953	0.83
May	0.020	1.12	259	-0.100	3.24**	112	986	0.82
June	0.038	2.36*	318	-0.094	3.00**	107	1038	0.81
July	-0.007	0.25	364	-0.048	2.00*	120	1069	0.76
August	-0.002	0.15	363	-0.074	3.13**	152	1047	0.81
September	0.008	0.43	386	-0.043	1.80	169	1047	0.78
October	-0.014	0.69	396	-0.027	1.22	201	1038	0.76
November	-0.029	1.60	373	-0.032	1.60	230	994	0.79
December	-0.031	1.77	399	-0.025	1.32	235	1013	0.77

Table 4 continued
Semi-logarithmic – WLS by volume

	Unmatched new		Unmatched old	
	Coefficient	*t*-statistic	Coefficient	*t*-statistic
February	-0.042	2.29*	-0.039	1.83
March	0.040	1.67	-0.090	2.36*
April	0.098	3.22**	-1.220	5.08***
May	0.099	3.98***	-0.086	2.96**
June	0.080	3.52***	-0.064	2.80**
July	0.068	3.00**	-0.019	0.70
August	0.050	2.11*	-0.005	0.20
September	0.024	0.90	-0.079	3.15**
October	0.058	2.01*	-0.059	2.17*
November	0.014	0.54	-0.045	1.86
December	0.034	1.17	-0.084	3.02**

***, **, * denote statistically significant at a 0.1%, 1% and 5% level respectively for two-tailed t-tests.

Table 5, Hedonic regression coefficients for dummy variable on time for matched and unmatched samples

Semi-logarithmic – OLS

	Matched plus old and new		**Matched**		**Matched plus old**		**Matched plus new**	
	Coefficient:		Coefficient:		Coefficient:		Coefficient:	
	Time dummy	*t*-statistic	time dummy	*t*-statistic	time dummy	*t*-statistic	time dummy	*t*-statistic
February	0.017	1.52	0.008	0.75	0.016	1.50	0.008	0.77
March	0.011	0.96	0.000	0.02	0.010	0.79	0.001	0.08
April	0.010	0.84	-0.010	0.83	0.001	0.07	-0.001	0.08
May	0.013	1.16	-0.013	1.16	0.006	0.54	-0.005	0.45
June	-0.015	1.29	-0.050	4.40	-0.033	2.63	-0.033	3.01
July	-0.034	2.36	-0.055	4.07	-0.012	3.15	-0.049	3.40
August	-0.024	2.17	-0.059	4.87	-0.036	2.83	-0.048	4.34
September	-0.034	2.69	-0.070	5.21	-0.049	3.34	-0.053	4.14
October	-0.049	3.30	-0.078	4.94	-0.055	3.35	-0.066	4.34
November	-0.054	4.20	-0.080	6.02	-0.056	3.85	-0.076	5.59
December	-0.067	5.48	-0.099	7.63	-0.070	5.39	-0.089	6.84

Semi-logarithmic - WLS by volume

	Matched plus old and new		**Matched**		**Matched plus old**		**Matched plus new**	
	Coefficient:		Coefficient:		Coefficient:		Coefficient:	
	time dummy	*t*-statistic	time dummy	*t*-statistic	time dummy	*t*-statistic	time dummy	*t*-statistic
February	-0.005	0.59	-0.005	0.52	-0.004	0.47	-0.006	0.62
March	-0.003	0.40	-0.006	0.66	-0.006	0.69	-0.004	0.42
April	-0.004	0.34	-0.017	1.97	-0.016	1.89	-0.006	0.54
May	-0.007	0.60	-0.026	2.70	-0.025	2.58	-0.008	0.73
June	-0.026	2.27	-0.046	4.28	-0.045	4.05	-0.028	2.48
July	-0.025	2.04	-0.046	4.77	-0.045	4.83	-0.026	2.18
August	-0.043	4.32	-0.025	2.27	-0.045	4.75	-0.044	4.53
September	-0.027	2.35	-0.045	4.54	-0.043	4.55	-0.034	2.92
October	-0.023	1.76	-0.047	5.01	-0.046	4.83	-0.027	1.88
November	-0.038	3.19	-0.056	5.20	-0.055	5.38	-0.041	3.25
December	-0.047	3.20	-0.075	6.83	-0.074	6.78	-0.058	4.27

Table 5 continued

Linear - OLS

	Matched plus old and new		Matched		Matched plus old		Matched plus new	
	Coefficient:		Coefficient:		Coefficient:		Coefficient:	
	time dummy	*t*-statistic	time dummy	*t*-statistic	time dummy	*t*-statistic	time dummy	*t*-statistic
February	6.974	1.34	1.809	0.36	6.070	1.21	2.682	0.51
March	7.809	1.47	0.958	0.18	5.618	1.06	2.724	0.52
April	6.830	1.30	-3.115	0.60	1.448	0.28	2.740	0.53
May	6.523	1.23	-4.235	0.80	2.174	0.41	0.828	0.16
June	-3.301	0.60	-18.987	3.17	-12.057	2.08	-9.503	1.66
July	-5.707	0.97	-19.655	3.42	-13.584	2.38	-11.978	2.04
August	-9.998	1.85	-22.497	3.79	-13.435	2.28	-18.624	3.38
September	-11.438	1.94	-25.287	4.13	-18.202	2.96	-17.044	2.81
October	-15.341	2.51	-24.240	3.54	-17.756	2.72	-19.090	2.85
November	-19.426	3.23	-31.476	4.70	-20.738	3.22	-26.217	4.03
December	-30.739	5.23	-38.874	6.14	-31.723	4.83	-35.246	5.74

Linear - WLS by volume

	Matched plus old and new		Matched		Matched plus old		Matched plus new	
	Coefficient:		Coefficient:		Coefficient:		Coefficient:	
	time dummy	*t*-statistic	time dummy	*t*-statistic	time dummy	*t*-statistic	time dummy	*t*-statistic
February	-1.547	0.49	-1.799	0.53	-1.371	0.43	-1.975	0.59
March	-0.029	0.01	-1.485	0.45	-1.473	0.47	-0.105	0.03
April	0.008	0.00	-5.251	1.61	-4.852	1.56	-0.583	0.16
May	-1.151	0.29	-8.432	2.46	-7.944	2.39	-1.540	0.38
June	-8.864	2.18	-16.021	4.15	-15.557	3.99	-9.479	2.32
July	-8.566	1.98	-16.516	4.34	-16.146	4.46	-9.174	2.10
August	-8.976	2.25	-15.783	4.21	-15.380	4.16	-9.781	2.40
September	-9.297	2.33	-14.947	3.92	-14.163	3.94	-11.849	2.85
October	-7.065	1.57	-15.618	4.22	-15.225	4.14	-8.407	1.73
November	-12.804	3.06	-19.923	4.72	-19.377	4.78	-14.252	3.16
December	-16.112	3.22	-26.168	6.09	-25.793	6.11	-20.896	4.41

Table 6, Descriptive statistics of leverage for matched and unmatched observations

Semi –logarithmic, WLS by volume - dummy variable hedonic regression: January and current month

	Residuals						**Leverage**					
	Matched		**Unmatched old**		**Unmatched new**		**Matched**		**Unmatched old**		**Unmatched new**	
	mean	Standard Deviation	Mean	standard deviation	mean	Standard Deviation	mean	standard deviation	mean	Standard Deviation	mean	standard deviation
Unweighted												
February	0.0040	0.1450	-0.0589	0.2404	0.0022	0.2045	0.0434	0.0400	0.0999	0.1879	0.0938	0.1713
March	0.0043	0.1647	-0.0567	0.2302	0.0033	0.1716	0.0383	0.0344	0.0840	0.1701	0.0687	0.1250
April	0.0012	0.1515	-0.0529	0.2249	0.0174	0.1634	0.0379	0.0343	0.0769	0.1566	0.0735	0.1437
May	0.0059	0.1417	-0.0746	0.2475	0.0105	0.1887	0.0362	0.0310	0.0782	0.1580	0.0626	0.1159
June	0.0015	0.1546	-0.0843	0.2577	0.0224	0.1592	0.0356	0.0343	0.0710	0.1490	0.0548	0.1251
July	0.0043	0.1737	-0.0525	0.1920	0.0066	0.2350	0.0340	0.0312	0.0645	0.1201	0.0463	0.0792
August	0.0089	0.1496	-0.0657	0.2279	0.0080	0.1592	0.0350	0.0353	0.0570	0.1026	0.0500	0.1050
September	0.0039	0.1711	-0.0457	0.2324	0.0123	0.1730	0.0332	0.0294	0.0563	0.1029	0.0513	0.1041
October	0.0067	0.1906	-0.0359	0.2196	0.0062	0.1751	0.0336	0.0301	0.0542	0.1003	0.0507	0.1054
November	0.0119	0.1553	-0.0354	0.2146	0.0024	0.1680	0.0341	0.0276	0.0567	0.0929	0.0523	0.0973
December	0.0108	0.1523	-0.0335	0.2195	0.0036	0.1704	0.0341	0.0341	0.0534	0.0918	0.0535	0.1117
Mean	0.0058	0.1591	-0.0542	0.2279	0.0086	0.1789	0.0360	0.0329	0.0683	0.1302	0.0598	0.1167

Table 6 continued

	Residuals						**Leverage**					
	Matched		Unmatched old		Unmatched new		Matched		Unmatched old		Unmatched new	
	mean	Standard Deviation	Mean	standard deviation	mean	standard deviation	mean	standard deviation	mean	Standard Deviation	mean	standard deviation
Weighted												
February	-0.0153	0.0092	-0.0398	0.0075	-0.0661	0.0122	0.0282	0.0005	0.2920	0.1650	0.0613	0.0080
March	-0.0151	0.0099	-0.0227	0.0073	-0.0078	0.0204	0.0250	0.0005	0.7197	0.1421	0.0536	0.0070
April	-0.0227	0.0100	-0.0445	0.0104	0.0434	0.0158	0.0248	0.0005	0.4573	0.1936	0.0366	0.0076
May	-0.0187	0.0104	-0.0328	0.0133	0.0411	0.0156	0.0241	0.0005	0.3839	0.1812	0.0300	0.0032

June	-0.0129	0.0101	-0.0372	0.0090	0.0321	0.0169	0.0228	0.0004	0.3305	0.1730	0.0301	0.0051
July	-0.0053	0.0103	-0.0460	0.0119	0.0309	0.0167	0.0225	0.0004	0.1383	0.0390	0.0252	0.0017
August	-0.0148	0.0100	-0.0355	0.0067	0.0070	0.0169	0.0230	0.0005	0.0677	0.0134	0.0263	0.0024
September	0.0054	0.0136	-0.0718	0.0082	-0.0015	0.0224	0.0221	0.0005	0.0477	0.0110	0.0282	0.0043
October	0.0102	0.0144	-0.0602	0.0098	0.0122	0.0243	0.0225	0.0005	0.0519	0.0124	0.0283	0.0035
November	0.0005	0.0107	-0.0576	0.0062	-0.0166	0.0206	0.0238	0.0005	0.0399	0.0042	0.0283	0.0022
December	0.0068	0.0133	-0.0810	0.0093	-0.0105	0.0251	0.0228	0.0005	0.0367	0.0048	0.0306	0.0045
Mean	-0.0074	0.0111	-0.0481	0.0090	0.0058	0.0188	0.0238	0.0005	0.2332	0.0854	0.0344	0.0045

Table 7, Indices of quality adjusted price changes, January 1998=1.00

	Matched data:				**Sample rotation**				**All data**	
	ratio of geometric means	**semi-log OLS**	**semi-log WLS by volume**	**with hedonic adjustments***	**Laspeyres biannual**	**Laspeyres Quarterly**	**Laspeyres monthly chained**	**Fisher chained index**	**semi-log OLS**	**semi-log WLS by volume**
February	1.006	1.008	0.995	0.994	0.995	0.995	0.995	0.993	1.017	0.995
March	1.001	1.000	0.994	0.992	0.993	0.993	0.995	0.992	1.011	0.997
April	0.991	0.990	0.983	0.983	0.983	0.986	0.988	0.984	1.010	0.996
May	0.987	0.988	0.974	0.971	0.971	0.974	0.979	0.974	1.013	0.993
June	0.951	0.950	0.954	0.950	0.950	0.958	0.965	0.958	0.985	0.974
July	0.947	0.945	0.954	0.932	0.943	0.947	0.958	0.950	0.966	0.975
August	0.945	0.941	0.975	0.942	0.939	0.947	0.950	0.939	0.976	0.957
September	0.935	0.930	0.955	0.933	0.930	0.942	0.942	0.931	0.966	0.973
October	0.927	0.922	0.954	0.914	0.925	0.940	0.941	0.929	0.951	0.977
November	0.921	0.920	0.944	0.908	0.917	0.934	0.934	0.920	0.946	0.962
December	0.908	0.901	0.926	0.912	0.906	0.921	0.921	0.907	0.934	0.953

OLS and WLS dummy variable regression estimates include an adjustment to the coeffficient on the Semi-logarithmic form of half the squared standard error following Teekens, R. and Koerts, J. (1972)

* semi-logarithmic WLS with forced replacements quality-adjusted using hedonic coefficients

Table 8, Regression results for semi-logarithmic hedonic equation for 1998

Dependent variable: ln PRICE; OLS estimator

Variable	Estimated Coefficient	Standard Error
C	2.466***	7.01
Months (benchmark: January)		
February	0.017	1.52
March	0.011	0.96
April	0.009	0.85
May	0.013	1.16
June	-0.015	1.29
July	-0.034*	2.36
August	-0.024*	2.17
September	-0.034**	2.70
October	-0.049***	3.30
November	-0.054***	4.20
December	-0.067***	5.48
Characteristics		
Height (cms.)	-0.003***	3.88
Depth (cms.)DEP	0.010***	9.06
Width (cms.)WID	0.014***	6.08
Water consumption (litres)	-0.001**	2.85
Load capacity (kgms.)	-0.00005	0.43
Spin speed (rpm)	0.00074***	34.59
Drying capacity-W/Dryer (kgms.)	0.00084	1.42
Condensor-W/Dryer	0.077***	7.38
Energy consumption (kWh.per cycle)	0.00074***	4.06
Vintage	0.016***	6.85
Type of machine (benchmark:front loader washing machine (WM)		
Top loader	0.562***	14.89
Twin tub	-1.610***	35.98
W/Dryer	0.157***	10.50
WM with computer	0.147***	10.79
WD with computer	0.253***	8.80
Installation (benchmark: free standing)		
Built-under integrated	0.563***	35.99
Built-under	-0.007	0.19
Built-in integrated	0.307***	7.16
Outlet type (benchmark: multiples)		
Mass merchandisers	0.077***	12.31
Independents	0.089***	15.00
Catalogue	0.242***	36.25
Makes (benchmark: Bosch)		
AEG	0.188***	16.88
Siemens	0.193***	14.04
Hoover	-0.166***	18.83
Miele	0.475***	28.97
Candy	-0.215***	18.69
English Electric	0.002	1.04
Ariston	-0.114***	8.80
New Pol	-0.332***	5.01
Beko	-0.370***	22.71
Zanussi	0.052***	5.43
Electro	0.057**	3.16
Indesit	-0.173***	8.84

Neff	*0.180****	*4.63*
Philco	*-0.235****	*6.43*
Ignis	*-0.0946****	*4.77*
Creda	*-0.171****	*10.56*
Tricity/Bendi	*-0.150****	*13.50*
Hotpoint	*-0.092****	*8.25*
Servis	*-0.222****	*18.16*
Asko	*0.318***	*2.88*
Adm	*0.6918****	*5.20*
Ocean	*-0.123****	*4.02*

Number of observations †: 6694; Std. error of regression = 0.180218; Adjusted R-squared = 0.792; F-statistic (zero slopes) = 456.5
****, **, * denote statistically significant at a 0.1%, 1% and 5% level respectively for two-tailed t-tests.*