

The Measurement of Quality-Adjusted Price Changes

Mick Silver and Saeed Heravi

Cardiff Business School
Cardiff University
Colum Drive
Cardiff CF10 3EU
UK

Email: Silver@cardiff.ac.uk

Fax: (0) 1222 874419

Draft for conference: not to be quoted without permission.

This study is part of a wider study funded by the UK Office for National Statistics (ONS). We are grateful to the ONS for permission to reproduce some of this work in the form of this paper. The views expressed in the paper are those of the authors and not the ONS. Any errors and omissions are also the responsibility of the authors. Helpful advice has also been received from David Fenwick (ONS), Dawn Camus (ONS), Pat Barr (GfK Marketing) and valuable programming assistance from Bruce Webb (Cardiff University). The paper has also benefited from useful comments from Erwin Diewert (University of British Columbia), Christos Ioannidis (Brunel University), Jack Triplett (Brookings Institute) and Ralph Turvey (LSE). The usual disclaimers apply.

Abstract

This paper considers three approaches to estimating quality-adjusted price changes: the dummy variable approach from a hedonic regression, a superlative or exact hedonic index (SEHI) approach and a matching technique. The dummy variable approach is prevalent in the literature and has been used as independent estimates of quality changes when commenting on sources of error in consumer price indexes (Boskin *et al.*, 1996 and Hoffman, 1998). However, the availability of scanner data provides an opportunity to utilise data on the prices (unit values), volumes and quality characteristics of a much wider range of transactions and to consider methods less restrictive than the dummy variable approach. The practical use of SEHI and a matching techniques using scanner data is explored and the results from all three methods compared.

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

Keywords: Superlative index numbers; Cost of living indexes; COLI; Consumer price indexes; hedonic regressions; quality adjustment.

1. INTRODUCTION

A major source of bias in the measurement of inflation is held to be its inability to properly incorporate quality changes (Boskin *et al.*, 1996 and 1998; Diewert, 1996; Cunningham, 1996; Hoffmann, 1998; Abraham *et al.*, 1998; Blow and Crawford, 1999). This is not to say statistical offices are unaware of the problem. Price collectors attempt to match the prices of 'like with like' to minimise such bias. However, comparable items are often unavailable and methods of implicit and explicit quality adjustment are not always considered satisfactory (Reinsdorf *et al.*, 1995, Armknecht *et al.*, 1997 and Moulton *et al.*, 1998).

Against this is an extensive empirical literature concerned with the measurement of quality-adjusted price indexes at the product level. The main approach is the use of hedonic regressions (though see Blow and Crawford, 1999 for an exception) in which the price of a model, for example, of a personal computer is regressed on its characteristics. The data source are often cross sectional time series from catalogues, the coefficients on the time dummies being estimates of the changes in price having controlled for changes in characteristics. There is usually little by way of data on quantities, and thus weights, in these estimates. Yet estimates from hedonic regressions have been used to benchmark the extent of bias due to quality changes in consumer price indexes (CPIs) (Boskin *et al.*, 1996 and 1998 and Hoffmann, 1998).

In this paper we argue against the use of this widely adopted dummy variable approach. We set it against theoretical developments in the measurement of superlative, exact hedonic indexes (SEHI) by Fixler and Zieschang (1992) and Feenstra (1995). The SEHI approach provides measures of cost-of-living indexes (COLI) based in economic theory. COLI measure the ratio of the minimum expenditure required to maintain a given level of utility. The dummy variable approach is shown to be a restricted version of the SEHI approach. Concordant with the development of the theory for the SEHI approach has been developments in data availability. We utilise scanner data from Electronic-Point-of-Sale (EPOS) bar-code readings, which provide a sufficiently rich source to implement the SEHI approach and compare it with results from the dummy variable method.

Against all of this Turvey (1999a and 1999b) has proposed, on pragmatic grounds, a matched approach akin to that adopted by statistical offices. We show how the SEHI and matched approaches are related having respective pros and cons for the measurement of quality-adjusted COLI. We also provide estimates, using scanner data, of superlative quality-adjusted cost-of-living indexes using all of these three approaches and compare the results.

It is worth noting that scanner data are now available in Europe and North America for a wide range of consumer durables and fast moving goods. The coverage of the data is often quite extensive, being supplemented by store audits for independent stores without bar code readers (see Hawkes and Smith, 1999). Market research agencies including ACNielsen and GfK Marketing Services collate and supply such data. Its use for validation and other purposes is now recommended for the compilation of consumer price indexes by Boskin *et al.*, (1996) and Eurostat (1998) and for direct use by Silver (1995). Given the existence of such rich data we can move on from estimates based on the prevalent, though restrictive, dummy variable

approach. There is thus a need to consider an appropriate methodology for the practical measurement of quality-adjusted COLI now we have such data.

The paucity of work on measuring quality-adjusted price indexes using this rich source is worth noting for the exceptions. There have been a number of studies using matching where prices of items with a particular specification are compared with their counterparts over time. These include Silver (1995) and Lowe (1998) for television sets and Reinsdorf (1996), Bradley *et al.*, (1998), Haan (1998), Dalen (1998) and Hawkes and Smith (1999) for selected food products. The matching used here is at a highly disaggregated level matching individual item codes with their counterparts over time. There have been fewer studies which compare the results of alternative methodologies: Silver (1999) for TVs, using the dummy variable and SEHI approaches and Kokoski *et al.* (1999) for audio products, using the dummy variable and matching approaches. Studies, especially those using the dummy variable approach, invariably focus on a single methodology with little interest in the relationship between methods. In this study we show how all three approaches are related and contrast the results for the case of washing machines.

In section 2 we outline the three methods of measuring quality-adjusted price indexes and show how they are related. Section 3 provides a description of the data, the application in this study being to monthly data on washing machines in 1998. The implementation of the three methods and their results are also outlined in section 3. Conclusions on the appropriate method to measure quality-adjusted price changes using scanner data are in section 4.

2. QUALITY-ADJUSTED PRICE INDEXES: THREE APPROACHES USING SCANNER DATA

This section outlines three methods for measuring quality-adjusted price changes using scanner data: the dummy variable hedonic method, a SEHI approach and a matching technique. The data and methods differ from those employed by statistical offices. They mainly use data obtained by price collectors matching items each month and recording and comparing their prices. If comparable items are unavailable in any month direct estimates or imputation procedures exist, though these are not always satisfactory.

(a) Hedonic regressions and dummy variables

The hedonic approach involves the estimation of the implicit, shadow prices of the quality characteristics of a product. Products are often sold by a number of manufacturers who brand them by their ‘make’. Each make of product is usually available in more than one model, each having different characteristics. A set of ($z_k = 1, \dots, K$) characteristics of a product are identified and data over $i=1, \dots, N$ product varieties (or models) over $t=1, \dots, T$ periods are collected. A hedonic regression of the price of model i in period t on its characteristics set z_{tki} is given by:

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

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

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

The theoretical basis for the method has been derived in Rosen (1974) where a market in characteristic space is established (see also Triplett, 1988 and Arguea *et al.*, 1994). There are a plethora of studies of the above form as considered by Griliches (1990), Triplett (1990) and Gordon (1990), but subsequently including Nelson *et al.* (1994), Gandal (1994 and 1995), Arguea *et al.* (1994), Lerner (1995), Berndt *et al.* (1995), Moulton *et al.* (1998), Hoffmann (1998) and Murray and Sarentis (1999). An issue of specific concern is the choice of functional form to be used. There has been support for, and success in, the use of the linear form, including Arguea *et al.* (1994), Feenstra (1995), Stewart and Jones (1998), and Hoffmann (1998). The semi-log formulation has also been successfully used including Lerner (1995), Nelson *et al.* (1997), Moulton *et al.* (1998) and Ioannidis and Silver (1998). Studies using, and testing for, more complex functional forms have generally been applied to housing (Rasmussen and Zuehlke, 1990 and Mills and Simenauer, 1996) with some success, the limited studies on consumer durable goods (using flexible functional forms and neural networks – Curry *et al.*, 1999) showing little benefit.

The data source used may be scanner data, but is often specialist magazines or mail order catalogues. The approach is not without problems. First, it implicitly treats each model as being of equal importance, when some models will have quite substantial sales, while for others sales will be minimal. Second, the prices recorded are not the transaction price averaged over a representative sample of types of outlets, but often a single, unusual supplier.

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

(b) Superlative and Exact Hedonic Indexes (SEHI)

Konüs (1939) and Diewert (1976) define a theoretical cost-of-living index (COLI), P_c as the ratio of the minimum expenditure 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.

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

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:

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

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

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

Following Feenstra (1995, *proposition 7*), consider an economic agent’s indirect utility function from consuming one unit of product i while spending her remaining wealth on the numeraire good. This function can be expressed in the familiar Gorman form as:

$$V_i = \ln(w) - \ln(q_i) + \varepsilon_i \quad i=1, \dots, N \quad (4)$$

where w is the consumer’s wealth, q_i denotes the quality adjusted price for product i and is defined as $q_i = \phi_i(p_i, z_i)$, where $\phi_i(\cdot)$ is defined over $R^{k+1} \rightarrow R$; the money price of the product i is denoted by p_i , and $z_i \in R_+^K$ denotes a vector of characteristics. Consumer heterogeneity is captured additively by the random variable ε_i . There are M consumers all equally likely to purchase product i , but they receive different realisations of ε_i .

It is assumed that $\partial \phi_i / \partial p_i > 0$ and $\partial \phi_i / \partial z_{ik} \leq 0$. As $\phi_i(\cdot)$ is monotonic in p_i , it can be inverted to obtain $p_i = \pi_i(q_i, z_i)$. The *marginal value of characteristic k* in product i is defined by $\partial \ln \pi_i(q_i, z_i) / \partial z_{ik}$, which is the change in the price p_i that a consumer would be willing to pay for a change in characteristic z_{ik} , keeping the quality-adjusted price q_i and therefore utility, V_i , constant. Feenstra (1995), *proposition 1*, using the results derived by McFadden (1983), shows that a utility function of the form given in equation (4) “...is consistent with individual utility maximisation, in the sense that demand from the representative consumer equals total expected demands from individuals..”.

In this context the aggregate expenditure needed to achieve a given aggregate utility, U_t , is given by :

$$E(p_t, z_t, U_t) = M (U_t G(q_{1t}^{-1}, \dots, q_{Nt}^{-1})^{-1}) \quad (5)$$

Denote the marginal value of a characteristic in each time period by $\beta_{it} = \partial \ln \pi_i(q_{it}, z_{it}) / \partial z_{it}$. As characteristics change over time, bounds for the exact index can be constructed using these values. The base and current period weighted quality-adjusted bounds, when $\ln \pi_i(q_{it}, z_{it})$ is concave in z_{it} , are given by (Feenstra (1995), *proposition 7*):

$$\prod_{i=1}^N \left(\frac{p_{it}}{\hat{p}_{it-1}} \right)^{S_{it}} \leq \frac{E(p_t, z_t, U)}{E(p_{t-1}, z_{t-1}, U)} \leq \prod_{i=1}^N \left(\frac{\hat{p}_{it}}{p_{it-1}} \right)^{S_{it-1}} \quad (6a)$$

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

$$\hat{p}_{it-1} \equiv p_{it-1} \exp\left[\sum \beta_{ikt-1} (z_{ikt} - z_{ikt-1})\right]$$

$$\hat{p}_{it} \equiv p_{it} \exp[-\sum \beta_{kt} (z_{ikt} - z_{ikt-1})] \quad (6b)$$

which are prices in periods $t-1$ and t respectively adjusted for the sum of the changes in each quality characteristic weighted by its respective marginal value derived from a semi-log hedonic regression, s_{it} and s_{it-1} are the shares in total value of sales of product i in periods t and $t-1$ respectively.

An arithmetic aggregation for a linear hedonic equation is given by:

$$\left[\frac{\sum_{i=1}^N x_{it} p_{it}}{\sum_{i=1}^N x_{it} \hat{p}_{it-1}} \right] \leq \frac{E(p_t, z_t, U)}{E(p_{t-1}, z_{t-1}, U)} \leq \left[\frac{\sum_{i=1}^N x_{it-1} \hat{p}_{it}}{\sum_{i=1}^N x_{it-1} p_{it-1}} \right] \quad (7a)$$

where $\hat{p}_{it} \equiv p_{it} - \sum \beta_{kt} (z_{ikt} - z_{ikt-1})$

$$\hat{p}_{it} \equiv p_{it-1} + \sum \beta_{kt} (z_{ikt} - z_{ikt-1}) \quad (7b)$$

where Laspeyres and Paasche are upper and lower bounds on ‘true’, economic theoretic COLIs: x is quantity sold, p is price, and z a vector of characteristics with associated marginal values β_{kit} derived from linear hedonic regressions over $i=1 \dots N$ product varieties (models) for each period t . Changes in the quality of models are picked up via changes in their characteristics ($z_{kt} - z_{kt-1}$) which are multiplied by estimates of their associated marginal values β_{kt} . With sales data available, the vector z can be the sales-weighted average usage or mix of each characteristic in each period. Note that \hat{p}_{it} corrects the observed prices, p_{it} for changes in the characteristics between the two periods, corresponding to the “explicit quality adjustment” described by Triplett (1990: 39).

Cost-of-living indexes (COLIs) are defined in economic theory as *exact* if they equal the ratio of expenditure required to maintain constant utility, when the (representative) consumer is facing changing prices. Different index number formulae have been shown to have an exact correspondence to the functional form of the consumer’s expenditure function. A superlative index in the Diewert (1976 and 1978) sense is one that corresponds to a flexible functional form for the expenditure function. Laspeyres and Paasche price indices correspond to fixed coefficient Leontief forms and act as upper and lower bounds on superlative index numbers, one such index being the geometric mean of the two, Fisher’s ‘ideal’ index. The base and current period weighted geometric means are exact for (correspond to) utility maximising (representative) consumers with constant elasticity of substitution. They act as upper and lower bounds respectively on a superlative Törnqvist¹ index, the Törnqvist index being in turn, exact for (corresponds to) a flexible translog utility function. Fisher’s and Törnqvist indexes are thus quite special in that they are *superlative*, though Diewert (1997) has also shown the two formulae to be superior to many others from an axiomatic approach with Fisher’s in particular, satisfying more ‘reasonable’ tests than its competitors.

There has been a need for practical procedures to measure COLI superlative indexes when current period weights are not available. Moulton (1996) and Shapiro and Wilcox (1997) advocates an ingenious approach whereby the results for a superlative index are predicted using imposed values of the elasticity of substitution, 0.7 being a good basis for this. Such an approach might take the form of:

$$\left[\sum w_{t-1} \left(\frac{p_t}{p_{t-1}} \right)^{1-s} \right]^{1/(1-s)} \quad (8)$$

With scanner data we do not need such estimates. We have “real-time” measures of the base and current period quantities and values.

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

(c) Matching

We finally consider the process of matching. The aim is to devise a method using scanner data which allows for quality adjustment by comparing only ‘like’ with ‘like’. This is akin to the process used by price collectors for statistical offices in the compilation of CPIs, but the matching is electronic using scanner data. Scanner data have a code to describe each model of a good. The code can be extended to include the type of outlet in which it is sold, in order that a particular model of a good in a particular type of outlet is matched against its counterpart in successive periods. Since individual retailers often have unique codes for the same model, the matching is in practice closer than by ‘model and outlet type.’ The problem with such matching is missing observations. For scanner data they arise when there is no transaction in that outlet (type) in a period, possibly because the item is no longer being sold or is on display, but no one bought it².

We use an example similar to Turvey’s (1999b) to explain the problem of missing observations. Consider the case of replacement with overlap for five items (V,X,Y,Z) over 5 months, (0 to 4), as shown in Table 1. The price index for period 0 compared with period 1 (0:1) involves price comparisons for W, X and Y. For (1:2) and (2:3) it involves all four items, for (2:3) the same four items but not v_3 , while for (3:4) it involves V, X,Y and Z. The sample composition changes for each comparison as items die and are born. The results for each comparison are chained to provide a single index for the whole period. Turvey (1999a) notes that an advantage of the method for CPI compilation is that it preempts the need for difficult or arbitrary judgements as to quality differences when items are replaced. For example, in period 3 the price collector would be unaware that there is no w_4 . One method of allowing for the demise of W is to use the price differential between it and its replacement, say V, in period 3 as an indicator of quality differences. However, if we do not know of its demise in period 3, we move on to period 4 without a price quote on v_3 and a difficult basis for the quality adjustment required to incorporate v_4 . “The sample will grow in size when new products appear and shrink when old products disappear, changing in composition through time.” (Turvey, 1999a, page 4). He advocates a

chained geometric mean of matched observations.

Table 1 Illustration of matching

Month					
Product	0	1	2	3	4
V				v_3	$\leftarrow v_4$
W	w_0	$\leftarrow w_1$	$\leftarrow w_2$	$\leftarrow w_3$	
X	x_0	$\leftarrow x_1$	$\leftarrow x_2$	$\leftarrow x_3$	$\leftarrow x_4$
Y	y_0	$\leftarrow y_1$	$\leftarrow y_2$	$\leftarrow y_3$	$\leftarrow y_4$
Z		z_1	$\leftarrow z_2$	$\leftarrow z_3$	$\leftarrow z_4$

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

The fixed base equivalent requires estimates for the price of products V and Z in period 0. This can be undertaken by retrapolating on the basis of price movements of products whose observations are *not* missing. Article 6 of EC Regulation 1749/96 requires that extensive use be not made of such procedures in CPI compilation. Explicit estimates of quality adjustments are preferred, though are time-consuming and their reliability may be variable. We follow Turvey (1999a and 1999b) in adopting the chained formulation.

The method involves some loss of information, which is naturally to be regretted. However, first it is undertaken to allow constant quality comparisons. Price comparisons of unmatched data suffer from being affected by changes in the quality mix of the product. Second, weighted indexes will be less prone to error from the omission of such data given their relatively low sales volume. However, countering this is that it is on the death of a product that price changes are unusual and these are the very ones lost. Some care, however, is needed in such statements. Table 1 illustrates how the loss of a matched observation takes place. For item W in Table 1, w_3 is used in period 3 for the period 3 to 2 comparison. So it is not lost here. However, it is lost in the matching for the period 4 to 3 comparison. It is tempting to argue that this loss is unimportant. It relates to a meaningless comparison since it does not exist in period 4 and there is thus no basis for a price comparison. However, economic theory would assert otherwise.

The economic of new goods is quite clear on the subject. If a new good is introduced it is not sufficient to simply wait for two successive price quotations and then incorporate the good. This would ignore the welfare gain to consumers as they substitute from old technology to new technology. Such welfare gains are

inseparately linked to definition of COLI defined as indexes which measure the expenditure required to maintain a constant level of utility (welfare). There exists in economic theory and practice the tools for the estimation of such effects (Hicks, 1940 and Diewert, 1990). This involves setting a 'virtual' price in the period before introduction. This price is the one at which demand is set to zero. The virtual price is compared with the actual price in the period of introduction and this is used to estimate the welfare gain. Hausman (1997) provides some estimates for the introduction of a new brand of Apple-Cinnamon Cheerios. He concludes: "The correct economic approach to the evaluation of new goods has been known for over fifty years since Hick's pioneering contribution. However, it has not been implemented by government statistical agencies, perhaps because of its complications and data requirements. Data are now available. The impact of new goods on consumer welfare appears to be significant according to the demand estimates of this paper; the CPI for cereal may be too high by about 25 percent because it does not account for new cereal brands. An estimate this large seems worth worrying about."

Shapiro and Wilcox (1997) p144 have the same concerns:

"This problem can be solved only by estimating the consumer surplus created by the introduction of each new item. Hausman (1994) argues that this must involve explicit modelling of the demand for each new item. Although explicit modelling demand may be of dubious practicality for widespread implementation in the CPI, strategic application in a few selected cases might be worthwhile."

The very same argument applies to the demise of good *W* in period 4 in Table 1. Again a 'virtual' price is required. Scanner data is rich enough to provide the means for doing so. It is not beyond the realms of practicalities to devise fairly rough, and possibly robust, methods to consider this on a practical basis.³

Matching using scanner data has been undertaken by Lowe (1998) for television sets, matching average prices for a given screen sizes. Kokoski *et al.* (1999) matched prices of models of audio products using bi-monthly data over a two year period. The comparison only included models for which prices were reported in the reference and comparison period, thus leading to a greater loss of data than the above method.

(d) Correspondence Between the Methods

(i) Matched versus SEHI

There is an interesting and useful correspondence here. Consider equation (7). The \hat{p} is the price (or unit value) of model i (in a given outlet) in period t having adjusted it by changes in its quality characteristics between period $t-1$ and t , the change in each characteristic being weighted by its associated marginal value in period t . If we are matching there is no such adjustment necessary. Matching does, however, have its failings in that we lose information. The SEHI formulation, *as undertaken in practice*, aggregates not over each model, but over meaningful characteristics. For washing machines for example, we might use makes and outlet types. The Laspeyres formulation of equation (7) is

$$\frac{\sum_{k=1}^j x_{kt-1} \hat{p}_{kt}}{\sum_{k=1}^j x_{kt-1} \bar{p}_{kit-1}} \quad (9a)$$

where we define j characteristics that are present in most models of the product in each period, where $k=1 \dots j \dots K$ characteristics, say dummy variables for makes and store type. The x and \bar{p} in equation (9) are now the average prices and total quantities *for each make in each store type*. Within a make and store type ($k=1 \dots j$) for each period t ,

$$\hat{p}_t = \bar{p}_t - \sum_i \beta_t (z_{it} - z_{it-1})$$

and $\hat{p}_{t-1} = \bar{p}_{t-1} + \sum_i \beta_{t-1} (z_{it} - z_{it-1})$ (9b)

The aggregation in equation (9a) being over the sales-weighted adjusted average prices for a particular make in a particular store type where z_{it} and z_{it-1} are the sales weighted averages in each period of the $k=j+1, \dots, k$ characteristics other than makes and outlets (e.g. spin speed), β_k their marginal values and \bar{p}_{it} the (sales weighted) mean price. The \hat{p}_k are thus quality-adjusted within each of the groups being aggregated in equation (9a). The more quality characteristics we aggregate over in the body of equation (9a), the less characteristics are used in determining \bar{p} in (9b). Equation (9a) should collapse down to the matched method, when aggregating over all characteristics. So why restrict the aggregation in (9a) to only makes and outlet types? The answer is that in doing so we use all the data. In Table 1 we lost v_3 for the period 3:2 comparison, but regained it for 4:3.

Note how there is a minimal loss of information in the SEHI formulation given by equation (9) since each model has a make and store type. If we aggregated over all characteristics in equation (9a), with no adjustments in (9b) we would have a matching process with some models having no price data for either period t or $t-1$ in any two-way comparison. These would be excluded. However, by allowing the aggregation over a limited number of k characteristics where k is defined to include models available in both periods, we lose no information. The adjustment for the variables not included in this weighted aggregation takes place in \hat{p} . There is a trade-off. The more quality variables in the weighted aggregation, the more chance of losing information. We consider this in the empirical section.

Both the SEHI and matching approaches allow all forms of weighting systems, Laspeyres, Paasche, Fisher etc., to be used to gain insights into such things as substitution effects. The SEHI formulation uses statistical estimates of product 'worth' to partial out quality changes for the characteristics excluded in the aggregation in (9a), rather than the more computational, and accurate, matching. The differences between the methods depend on the reliability of the \hat{p} adjustment process, in terms of both the extent of changes in characteristics (Δz) and the values of β , and the relative loss of observations in bringing these characteristics into the aggregation process. It is an empirical matter and we will investigate this.

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

(ii) Direct versus SEHI

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

3 DATA AND IMPLEMENTATION

a Data

(i) Scope and coverage

The study is for monthly price indexes for washing machines in 1998 using scanner data. Scanner data are compiled on a monthly basis from the scanner (bar code) readings of retailers. The electronic records of just about every transaction 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 anti-competitive 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 Silver (1995), whose use is supported for example by Lowe (1998) for Canada, Moulton, LaFleur and Moses (1998) and Boskin *et al.* (1995) for the US and Eurostat – as Astin and Sellwood (1998 p297-298) note in the context of HICP:

“Eurostat attaches considerable importance to the possible use of scanner data for improving the comparability and reliability of HICPs 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. This is a very rich formulation since it allows us to estimate, for example, the marginal value of a characteristic in a particular month and a particular type of outlet and apply these to changes in the usage of such stores. Not all makes are sold in each type of outlet. In January 1998, for example, there were 266 models of washing machines with 500 observations, that is each model was sold on average in 1.88 types of outlets.

The coverage of the data is impressive both in terms of transactions and features. For the UK for example in 1998, there were 1.517 million transactions involving 7,750 observations (models/outlet types) worth £0.55 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.

(ii) The variables

The variable set includes:

Price - the unit value (Balk, 1996) of all transactions for a model in an outlet type in a month.

Volume 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 condensers (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.

(b) Implementation of Each Method

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

(i) The dummy variable approach

Both linear and semi-log formulations were used having, as outlined earlier, had some support in theory and practice for indexes based on the arithmetic and geometric means respectively (Feenstra, 1995). The \bar{R}^2 for the respective forms were relatively high at 0.83 and 0.82. A Box-Cox transformation was used for testing functional form the estimated λ being 1.003 with $SE(\lambda)=0.024$ favouring the linear form⁴. A Bera-McAleer test based on artificial regressions was, however, inconclusive⁵.

The t -statistics for θ_1 , θ_2 were 13.18 and 36.1 respectively. The F-statistics for the null hypotheses of the results of coefficients all equal to zero was rejected for both functional forms at 314.6 and 297.2 respectively for linear and semi-log and p -values of 0.0000.

The coefficients were almost invariably statistically significant with appropriate signs and magnitudes (results available from authors). There was some evidence of heteroskedasticity with the Breusch-Pagan test statistics of 27.7 and 9.0 for the linear and semi-log forms respectively, both exceeding the critical value of Chi-squared (3 degrees of freedom) = 7.815. However, the estimator remains unbiased and the t -statistics were adjusted to be heteroskedastic-consistent using a procedure by White (1980).

The regressions were estimated on a data set that excluded models with sales of 30 or less in any month and a minimal number of models with extreme prices arising from variables not included in the data, such as stainless steel washing machines. A failing of the dummy variable approach is that models with only one transaction are given the same importance in the regression as a model with, say 10,000 transactions. The choice of 30 was based on some experimentation. The loss in the number of observations was quite severe for washing machines from 7,750 to 3,957, while the loss in terms of the *volume* of sales was minimal, from 1.517 million to 1.482 million.

(ii) Superlative and Exact Hedonic Indexes (SEHI)

First, it is necessary to decide which quality-related variables are used in the aggregation in equation (9a), and which for the adjustment in (9b). As a first step we use 'makes' for the aggregation, but later extend this. The β_t estimates are then derived from *monthly* hedonic regressions and multiplied by changes in the sales-weighted change in the mix of quality characteristics to provide an adjustment to average prices (\bar{p}) for use in the main body of equation (9a).

The β coefficients are required for each of the remaining $j+1, \dots, K$ quality-related variables in (9b) *in each month*. The specification of the regression equations estimated for this purpose used all variables, to avoid omitted variable bias, with only the relevant β_t coefficients being used to generate \bar{p} . The specifications were

therefore similar to the direct method except that separate regressions were estimated for each month⁶.

The mean \bar{R}^2 over the 12 monthly hedonic regressions was 0.842⁷. The coefficients were almost invariably statistically significant and of reasonable magnitude with the appropriate sign⁸. As with the dummy variable method the regressions were estimated using the linear and semi-log forms. The coefficients from the linear form were used to derive quality-adjusted prices for use in an arithmetic framework – that is for Laspeyres, Paasche and (superlative) Fisher SEHI (equation 7). The coefficients from the semi-log form were used to calculate base and current period weighted geometric means and (superlative) Törnqvist SEHI, as given in equation (6) and footnote (1).

As explained in Section 2 the SEHI approach has the advantage over matching of minimal loss of data. However, the more variables included in the aggregation in (9a), the more information loss, as either of p_{it-1} or p_{it} becomes unavailable in any period for comparison. In the limiting case of all variables being included the method collapses to the matched approach. If we aggregate over makes, or even ‘makes and outlet types’, there is very little loss of data in terms of the number of observations and volume of sales. Aggregating only over the 21 makes leaves us with 99.67% of observations and 99.97% of sales volume. Extending the aggregation to the 21 makes and 4 outlets, 84 combinations still has little loss of data – 99.08% of observations and 99.92% of sales volume. Any manufacturer operating in a particular outlet type continues to do so on a monthly basis. Extending the aggregation further to 24 spin-speeds, i.e. over 2,016 combinations reduces the coverage to 95.9% of observations and 99.6% of sales volume.

(iii) Matching

The matching procedure used incurred further loss of data: only 83% of observations were used, though the missing ones were models in outlets which were being discarded with low sales, the volume of sales used in the matching being 97.8%

The extent of the matching is illustrated for washing machines in 1998 in Table 2. There were for example, 429 matched comparisons of a particular model in a specific outlet type in February 1998. These were selected from 500 and 498 observations available in February and January 1998 respectively. In total there were 6,020 matched comparisons for 1998 which compares with 7,750 available in 1998 or, more fairly, $7,750 - 500 = 7,256$ to exclude the January figures since the matched comparisons are over 11 monthly comparisons as opposed to 12 months data.

Table 2, Data on Matching for Washing Machines, 1998

	Number of observations*		Volume of sales (thousands)	
	Unmatched	Matched	Unmatched	Matched
January	500	-	126.2	-
February	488	429	111.4	115.1
March	605	425	134.0	118.6
April	625	510	113.3	120.7
May	647	527	112.5	111.3
June	711	555	137.1	122.5

July	744	620	116.3	124.9
August	711	627	123.0	118.3
September	717	606	150.4	135.1
October	695	602	129.2	138.5
November	643	566	124.8	125.8
December	664	553	138.6	129.8

*There were, for example, 429 matched comparisons between February and January 1998 taken from 488 observations (model in a specific outlet type) in February and 500 in January 1999. While the number of matched observations will exceed those unmatched the volume of sales may not do so.

This difference of 1,236 observations are price data that exist in either one of period t or period $t+1$, but do not have a counterpart to enable a comparison. Since they are models just born or about to die, they should have low sales and thus their omission should not unduly affect the index¹⁰. The total sales volume of matched comparisons was 1.3605 million compared with 1.3906 million (unmatched but excluding January) – a difference of about 30 thousand sales or about 2% of sales. From Table 2 the monthly variation can be deduced. The worst loss of information was in the March to February comparisons: from $(111.4 + 134.0)/2 = 122.7$ thousand to 118.6 thousand - a loss of 3.3%. For the September to October and October to November comparisons the losses were less than 1%.

A Unit value index is given by:

$$\left[\frac{\sum p_t x_t}{\sum x_t} \right] / \left[\frac{\sum p_{t-1} x_{t-1}}{\sum x_{t-1}} \right] \quad (10)$$

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

4. RESULTS

Table 1 and Figure 1 provide results using the matched approach for several different formulae, the results being qualitatively similar for the SEHI approach.

First, the two superlative indexes, Fisher and Törnqvist are almost identical, as expected. Second, Laspeyres and Paasche provide outer upper and lower bounds respectively to the superlative indexes, the extent of the substitution being about 1.35% over the year, as consumer substituted away from machines with relatively high price increases. Note that because these indexes are chained on a monthly basis, they understate the substitution arising from their fixed base index counterparts. They allow the basket to be updated each month, the substitution in each month being compounded over the year. Finally, the geometric base and current period weighted indexes are upper and lower inner bounds respectively on the superlative indexes since they incorporate some substitution effect (Shapiro and Wilcox, 1997). All of this is as predicted by economic theory. Figure 1 and Table 1 also show the unit value index, an index defined in section 3 to be unaffected by changes in the quality mix of models. While the index shows only a slight overall fall in prices over the year of about 1%, and increases in other months compared with January 1998, quality-adjusted price changes have fallen by just under 10% over the year. These superlative matched indexes effectively adjust for changes in the quality mix of purchases being based on computational matching as opposed to statistical models. They lose

however, 2% of the data by sales volume. We now consider SEHI indexes which lose only 0.4% of sales volumes.

Figure 2 and Table 2 provides results for different approaches to measuring quality-adjusted price indexes using an arithmetic formulation. The estimates from a linear model using the dummy variable approach show a fall of 6%, quite different from the other estimates. In section 3 we found the hedonic regression to have a relatively high \bar{R}^2 with signs and values of the coefficients being as expected on *a priori* grounds¹¹. The linear formulation was supported by a Box-Cox test, though the results from a semi-log formulation (Table 3) are very similar. By conventional standards these estimates are quite acceptable. The stark difference between the results from other approaches is more likely to be a result of the absence of a weighting system for the dummy variable approach. If prices of more popular models are falling faster than unpopular ones, the weighted matched and SEHI approaches will take this into account, while the dummy variable method will not. The results from SEHI and matched estimates are quite similar, a difference of about 0.4% and 1% for Laspeyres and Fisher over the year. In this case study the loss of data for the matching at about 2% by volume, was relatively low giving confidence in the results.

Figure 3 and Table 3 repeats the exercise using formulations based on a geometric mean. Again the dummy variable semi-log estimates, while closer to the other estimates, are still over 2% above the superlative estimate over the 12 months. However, the respective matched and SEHI results show similar patterns with differences of about 0.75 to 1.0 percentage over the year.

Finally Figure 4 shows the results for SEHI at different levels of aggregation for Fisher indexes. As we expand the weighted price changes in the body of formula (9a) from just makes to makes within each outlet type, and then further by spin-speed. The SEHI approaches the matched index.

In summary this paper uses scanner data to show how to measure quality-adjusted price changes. It casts doubt on the use of the dummy variable approach. It also argues for a matched approach as a special case of the theoretically based SEHI approach, the matched approach being based on computational matching and not being subject to the ideosyncracies of the econometric estimation of hedonic indexes (Griliches, 1990 and Triplett, 1990). Caution is however advised when the loss of data in matching is severe. In such a case an empirical investigation into the trade-off between including variables in the aggregation and the resulting loss of data is advised.

Notes

1. The Törnqvist index is given by:

$$\prod \left(\frac{P_t}{P_{t-1}} \right)^{\left(\frac{w_t + w_{t-1}}{2} \right)}$$

$$\text{where } w_t = \frac{P_t x_t}{\sum P_t x_t} \text{ and } w_{t-1} = \frac{P_{t-1} x_{t-1}}{\sum P_{t-1} x_{t-1}}$$

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

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

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

5. The Bera-McAleer test involves obtaining predicted values $\log(\hat{z})$ and (\tilde{z}) from a semi-logarithmic and linear formulation respectively. Artificial regressions are then computed using $\exp\{\log(\hat{z})\}$ and $\log(\tilde{z})$ on the left-hand-side and the residuals from each of these regressions, \hat{v}_1 and \hat{v}_0 , are included in a further set of artificial regressions:

$$\log(z) = \beta_0 + \beta_1 X_1 + \theta_0 \hat{v}_1 + \varepsilon_1;$$

$$z = \beta_0 + \beta_1 X_1 + \theta_1 \hat{v}_0 + \varepsilon_2$$

Using t-tests, if θ_0 is accepted we choose the log-linear model and if θ_1 is accepted we choose the linear model, the test being inconclusive if both are rejected or both accepted.

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

7. Test statistics here are illustrative being based on semi-log and linear models for the data as a whole, though they are indicative of the results for individual months (available from authors). Estimates for log-log models were not feasible given the large number of dummy variables on the right-hand-side of the equation.
8. Results are available from authors.
9. If, for example, the matched item had sales in period t of 100 and in period $t+1$ of 50, and a new model was launched in period $t+1$ with sales of 10, the matched volume would be $150/2 = 75$ and the unmatched 60 in period $t+1$.
10. The data are transactions over the month, so recently born or dead model may only have been available for part of the month in question and have relatively low sales.
11. It may be argued that weighted least squares using volume of sales (or its square) may be appropriate. This changes the specification of the regression. Yet the importance given in the regression to a model of washing machines with sales of 10,000 in a month, when determining the slope coefficient, is the same as that given to a model of washing machines with only one transaction. In any event estimates using WLS were no closer to the other approaches.

References

- Abraham K, Greenless J and Moulton, B (1998) Working to Improve the Consumer Price Index, *Journal of Economic Perspectives*, 12 (1), 23-36.
- Arguea NM, Haseo C and Taylor GA, (1994) Estimating Consumer Preferences using Market Data – An Application to US Automobile Demand, *Journal of Applied Econometrics*, 9, 1-18.
- Armknrecht, PA, Lane, WF and Stewart KJ (1997) New Products and the US Consumer Price Index. In Timothy F Bresnahan and Robert J Gordon (Eds) *The Economic of New Goods, Studies in Income and Wealth Vol.58*, NBER Chicago and London: University of Chicago Press.
- 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, Paper presented at the Ottawa Group Meeting, Bureau of Labour Statistics, Washington DC.
- Berndt, ER, Grilches, Z., and Rappaport N J (1995), Econometric estimates of price indexes for personal computers in the 1990s, *Journal of Econometrics*, 68, 243-68.
- Blow L and Crawford I (1999), *Cost of Living Indices and Revealed Preference*, London: The Institute of Fiscal Studies.
- 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.
- Bradley, R Cook, B, Leaver, SG and Moulton, BR (1998) An Overview of Research on Potential Uses of Scanner Data in the U.S. CPI, *Proceedings of the Third Meeting of the International Working Group on Price Indices*, edited by B. Balk. Research Paper no. 9806, Statistics Netherlands, 169-182.
- Cameron J and Collins A, (1997) Estimates of Hedonic Ageing For Partner Search, *Kyklos*, 50, 3, 409-418.
- 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 (1999) Hedonic Regressions: Misspecification and Neural Networks, *Mimeo*, Cardiff Business School.
- Dalen, 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.
- Diewert, WE (1976) Exact and Superlative Index Numbers, *Journal of Econometrics*, 4 115-45.
- Diewert, WE, (1978), Superlative Index Numbers and Consistency in Aggregation, *Econometrica*, 46, 883-900.
- 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 (1996) Sources of Bias in Consumer Price Indexes, *Discussion Paper No: DP-96-04*, School of Economics, University of New South Wales.
- Diewert, WE, (1997), *Commentary on Shapiro, MD and Wilcox, DW. (1997), Alternative Strategies for Aggregating Prices in the CPI, Federal Reserve Bank of St Louis Review*, 79, 3, 113-125.
- Eurostat (1999) Meeting of the Working Party on Harmonisation of Consumer Prices, Quality Adjustment, HCPI 98/227, Luxembourg: Eurostat B-3.
- Feenstra RC (1995), Exact Hedonic Price Indexes, *Review of Economics and Statistics*,

LXXVII, 634-54.

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

Forsyth, FG, and Fowler, RF (1981), The Theory and Practice of Chain Price Index Numbers, *Journal of the Royal Statistical Society, A*, 144, 2, 224-47.

Gandal, N (1994) Hedonic Price Indexes for Spreadsheets and Empirical Test for Network Externalities, *RAND Journal of Economics*, 25, 1, 160-170.

Gandal, N (1995) Competing Compatibility Standards and Network Externalities in the PC Market, *Review of Economics and Statistics*, LXXVII, 599-608.

Gordon, RL (1990) *The Measurement of Durable Goods Prices*, Chicago: University of Chicago Press.

Griliches, Z, (1990) Hedonic Price Indexes and the Measurement of Capital and Productivity: Some Historical Reflections In ER Berndt and JE Triplett (Eds). Fifty Years of Economic Measurement: The Jubilee Conference of Research in Income and Wealth, NBER *Studies in Income and Wealth*, Vol.54, Chicago University of Chicago Press.

Haan, J de and Oppendoes, 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 Timothy Bresnahan and Robert J Gordon *The Economics of New Goods*, Studies in Income and Wealth 58, Chicago and London: University of Chicago.

Hawkes W and Smith R (1999) Improving Consumer Price Measurements Through the use of Scanner Data and Market Segmentation, Draft Paper for Presentation at the Eurostat/ONS Cardiff University Conference on the Measurement of Inflation, Cardiff.

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

Hoffman 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 (1998) Chained, Exact and Superlative Price Changes: Estimates from Micro Data. *Proceedings of the Third Meeting of the International Working Group on Price Indices*, Edited by B. Balk, Research Paper 9806, Statistics Netherland, 215-224.

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

Kmietowicz, ZW and Silver MS (1980) New Products and the Index of Industrial Production, *Journal of Development Studies*, Vol.16, No.4, July.

Kokoski M, Waehrer K and Rozaklis P (1999) Using Hedonic Methods for Quality Adjustment in the CPI: The Consumer Audio Products Component, Draft Paper for Presentation at the Eurostat/ONS Cardiff University Conference on the Measurement of Inflation, Cardiff.

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

Lerner, J (1995) "Pricing and Financial Resources: An Analysis of the Disk Drive Industry, 1980-88," *The Review of Economics and Statistics*, LXXVII 585-598.

Lowe R (1998) Televisions: Quality Changes and Scanner Data, Paper Presented at the Fourth Meeting of the International Working Group on Price Indices, Bureau of Labor Statistics, Washington DC.

McFadden, D, (1983), Econometric models of probabilistic choice in Charles F. Manski and Daniel McFadden (Eds.) *Structural Analysis of Discrete Data with Econometric Applications* (Cambridge, MA: MIT Press 1983), 198-272.

Moulton B (1996) Bias in the Consumer Price Index: What is the Evidence? *Journal of Economic Perspectives*, 10, 4, 139-177.

Moulton, B, LaFleur, TJ, Moses, KE (1998) Research on Improved Quality Adjustments in the CPI: The Case Televisions. Paper Presented at the Fourth Meeting of the International Working Group on Price Indices, Bureau of Labor Statistics, Washington DC.

Nelson, RA, Tanguay, TL, and Patterson, C.D., (1994), "A Quality-Adjusted Price Index for Personal Computers," *Journal of Business and Economics Statistics*, 12, 1, 23-31.

Rasmussen, DW and Zuehlke, TW (1990), On the Choice of Functional Form for Hedonic Price Functions, *The Review of Economics and Statistics*, November, 668-675.

Reinsdorf M (1996) Constructing Basic Components for the U.S. CPI from Scanner Data: A Test Using Data on Coffee, Working Paper 277, Bureau of Labor Statistics, Washington, D.C.

Reinsdorf M (1998), Divisia Indexes and the Representative Fourth Meeting of the International Working Group on Price Indices, Washington DC: Bureau of Labor Statistics, pp.210-235.

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

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

Rosen S (1974) Hedonic Prices and Implicit Markets: Product Differentiation and Pure Competition, *Journal of Political Economy*, 82 34-49.

Shapiro, MD and Wilcox, DW (1997) Alternative Strategies for Aggregating Prices in the CPI, *Federal Reserve Bank Review*, 79, 3, 113-126.

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 (1998) Bias in the Compilation of a CPI when Models of Different Vintages Coexist in the Market. Paper presented at the Ottawa Group Meeting, Bureau of Labor Statistics, Washington DC.

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.

Stewart, KG and Jones, JCH (1998), Hedonic and Demand Analysis: The Implicit Demand for Player Attributes, *Economic Inquiry*, 36, 2, April, 192-202.

Szulc, BJ (1983) Linking Price Index Numbers, in *Price Level Measurement*, 537-566, Ottawa, Statistics Canada.

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

Triplett, JE, (1990), Hedonic methods in statistical agency environments: an intellectual biopsy, In ER Berndt and JE Tripplett (eds) *Fifty Years of Economic Measurement: The Jubilee Conference on Research in Income and Wealth, NBER Studies in Income and Wealth*, Vol 56, Chicago University of Chicago Press.

Turvey, R (1999a) Incorporating New Models into a CPI: PCs as an example. Paper to be presented at the Eurostat Conference on the Measurement of Inflation, Cardiff Business School, ONS, and Cardiff 1999.

Turvey R (1999b) Elementary Aggregate Indexes, Draft Manual: www.turvey.demon.co.uk, revised April.

White, H, (1980) A Heteroskedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity, *Econometrica*, 48, pp.721-746.

Table 1, Matched Quality-Adjusted Price Indexes by Formulae

	Laspeyres	Paasche	Fisher	Unit values	GMcurrent	Gmbase	Tornqvist
January	100.000	100.000	100.000	100.000	100.000	100.000	100.000
February	99.489	99.016	99.252	98.959	99.058	99.452	99.255
March	99.520	98.772	99.145	100.609	98.836	99.449	99.142
April	98.822	98.076	98.448	101.862	98.196	98.696	98.446
May	97.908	97.049	97.478	102.042	97.246	97.715	97.481
June	96.455	95.129	95.790	100.089	95.452	96.148	95.800
July	95.780	94.211	94.992	101.863	94.628	95.378	95.002
August	94.970	92.890	93.924	101.278	93.420	94.471	93.944
September	94.241	91.953	93.090	100.275	92.532	93.690	93.109
October	94.061	91.710	92.878	101.673	92.340	93.461	92.899
November	93.387	90.697	92.032	99.888	91.382	92.732	92.055
December	92.063	89.337	90.690	98.933	90.074	91.363	90.717

Table 2, Quality-Adjusted Price Indexes Based on Arithmetic Means.

	Hedonic (SEHI) by make and store type				
			Dummy	Matched	
	Laspeyres	Fisher	variable	Laspeyres	Fisher
	SEHI	SEHI	Linear	Matched	Matched
January	100.000	100.000	100.000	100.000	100.000
February	99.015	98.860	100.022	99.489	99.252
March	98.894	98.960	99.315	99.520	99.145
April	98.504	98.571	99.978	98.822	98.448
May	97.821	97.916	100.164	97.908	97.478
June	95.936	95.980	100.512	96.455	95.790
July	95.850	95.771	101.114	95.780	94.992
August	95.184	94.892	101.086	94.970	93.924
September	94.782	94.300	101.591	94.241	93.090
October	94.857	94.318	101.655	94.061	92.878
November	93.579	93.007	101.825	93.387	92.032
December	91.679	91.626	102.582	92.063	90.690

Table 3, Quality-Adjusted Price Indexes Based on Geometric Means.

	Hedonic (SEHI) by make and store type				
			Dummy	Matched	
	Gmbase	Tornqvist	variable	Gmbase	Tornqvist
	SEHI	SEHI	Semi-log	Matched	Matched
January	100.000	100.000	100.000	100.000	100.000
February	99.280	99.052	100.081	99.452	99.255
March	99.055	99.033	101.854	99.449	99.142
April	98.718	98.654	99.970	98.696	98.446
May	98.000	97.989	99.337	97.715	97.481
June	96.214	96.145	98.786	96.148	95.800
July	96.124	95.999	97.349	95.378	95.002
August	95.295	94.978	97.313	94.471	93.944
September	94.933	94.477	95.970	93.690	93.109
October	94.989	94.525	96.375	93.461	92.899
November	93.968	93.423	95.393	92.732	92.055
December	92.151	91.741	93.940	91.363	90.717

Figure 1, Matched quality adjusted price indexes for washing machines

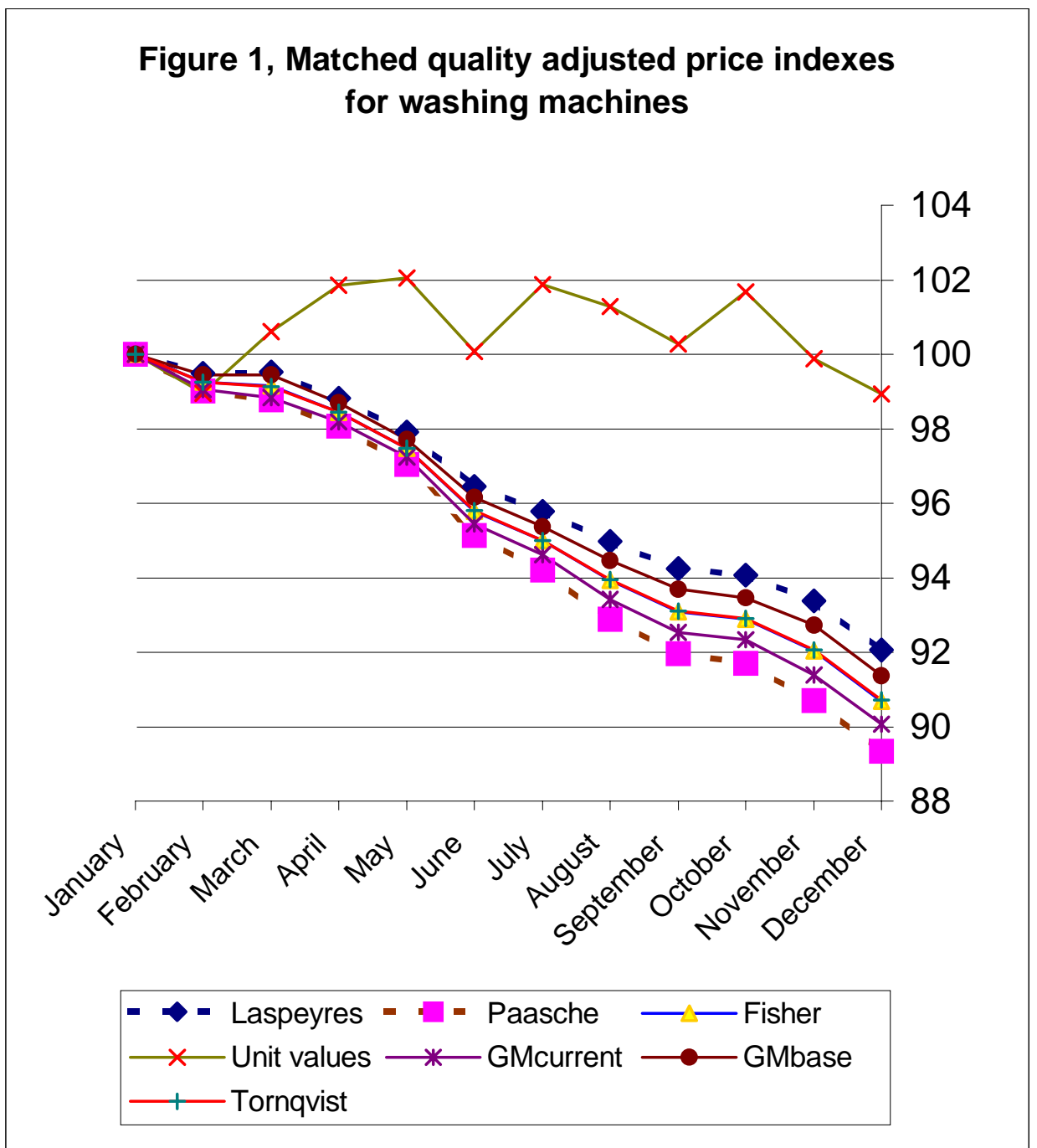


Figure 2, Quality-Adjusted Price Changes Indexes Using Arithmetic Means

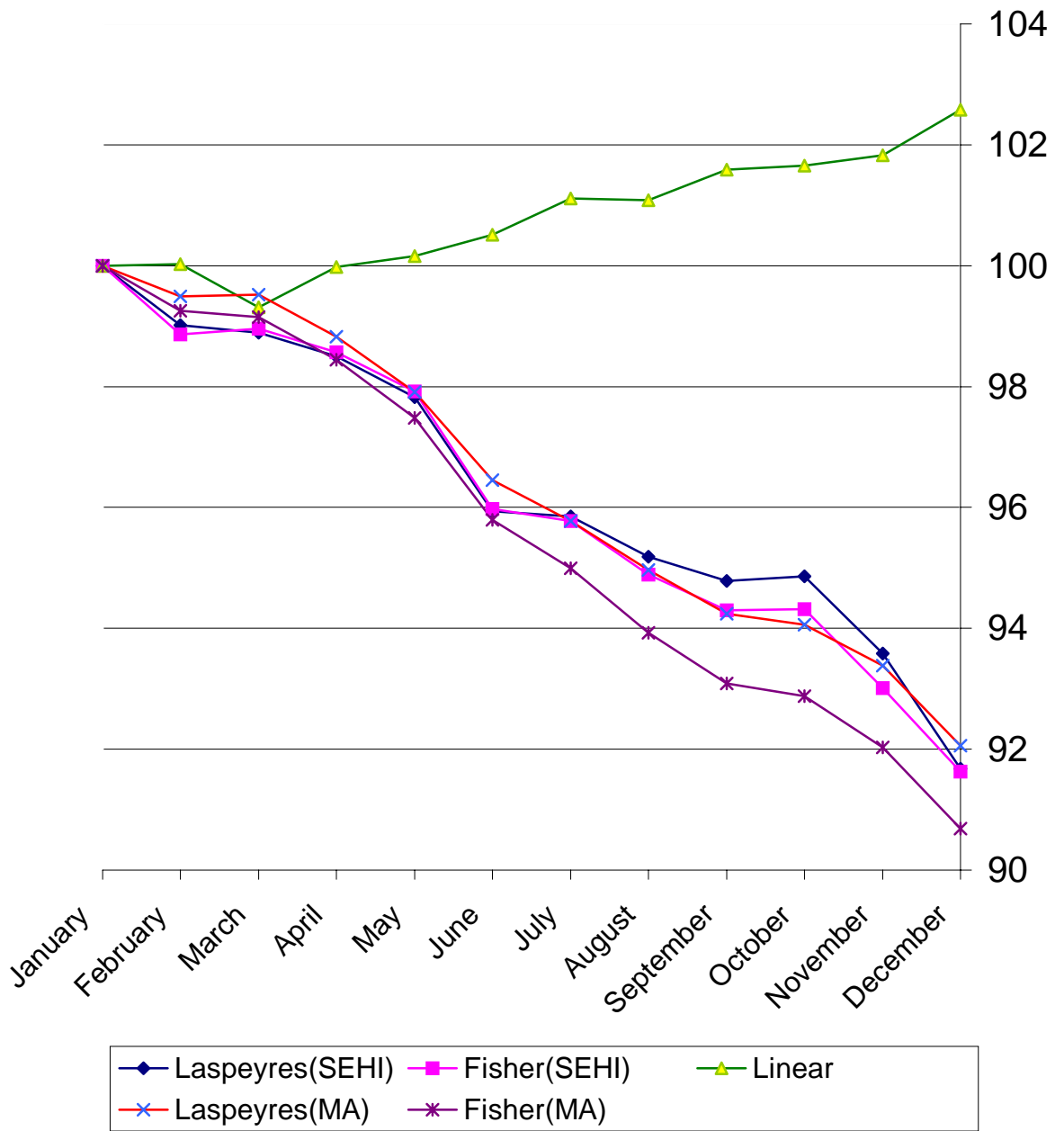


Figure 3, Quality-Adjusted Price Indexes Using Geometric Means

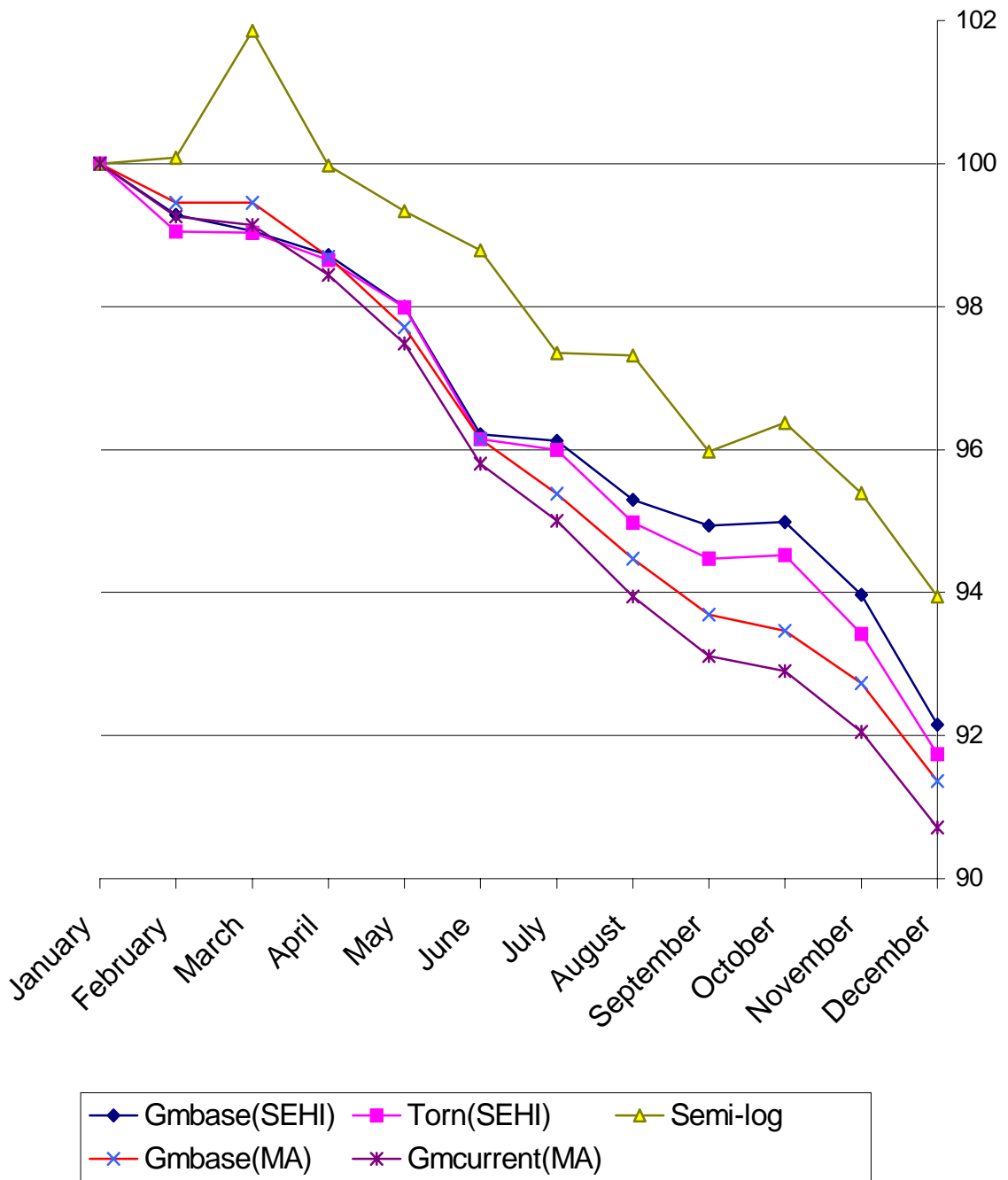


Figure 4, SEHI at different levels of aggregation

