Controlling for Quality Change in Digital Cameras Price Index

Kari Manninen
Bureau of Economic Analysis
7/7/2003

Abstract

In this paper we compare different hedonic and conventional quality adjustment methods in an uniform but somewhat unconventional descriptive framework. The main aim is to address questions on hedonic quality adjustment methods and their robustness in index compilation and to give an empirical example with digital camera prices. We will show how conventional quality adjusting methods may be treated parallel with hedonic ones and these methods may be evaluated similarly with regression based methods. The hedonic models in this paper are all used as forecast models, that is different e.g. from the BLS use of hedonic coefficients. Further, forecasting log prices enables us as to avoid some problems of structural models and, on our opinion, adds the method's robustness and feasibility.

Furthermore, we show how to decompose an unadjusted price index into a quality adjusted pure price change, a (separable average) quality correction term, and a covariance effect indicating the effect of covariance between the weights and estimated price change. This decomposition enables us to separate and quantify the effects of quality correction and weight effects in different hedonic models and compare them with other quality adjustment methods.

The empirical part of the paper is based on findings from a quarterly digital camera database including some 1,200 prices from over 250 different digital camera models over the years 1998 to 2002. The main findings indicate that relatively simple hedonic models may be sufficient for accurate quality controlling even in high technology product price indices. Further, if compared with a matched model framework, the collection of characteristics data for hedonics may not need to exceed the precision already needed in the matched model. This suggests that it may be feasible to use hedonic indices even in high frequency index compilation. To validate this, we claim that additional cross sectional explanatory power from a set of added quality characteristics in hedonic models may have only marginal longitudinal effects in the index series. We calculate various hedonic models and concentrate more on sensitivity in index variability than the model specification.

Contents:

Abstract 2
Introduction 4
1. The Economy 4
1.1 Defining the economy 4
1.2 Full information price index 7
2. Conditional expectation and hedonic price indices 9
3. The data 12
3.1 The sample design 12
3.2 Data reservations 13
4. Average prices and characteristics 15
4.1 Average price 15
4.2 Quality characteristics 16
5. Matched model and classification method 18
5.1 A matched model index 18
5.2 The classification method 19
6. Hedonic methods 20
6.1 Decomposition of hedonic price index 21
6.3 Time indictor models 22
6.4 Relaxing the temporary restrictions 24
7. Results 26

Appendix 1. Average prices and the classification method 28
Appendix 2. Decomposition of hedonic geometric indices 30
Appendix 3. Some regression results 35
Appendix 4. Model 3 and 4 coefficients 36

Introduction

In the first two sections, we provide the basis for hedonic regression models and hedonic price indices. We start with a description of an economy and discuss how one could measure the change in total output value in terms of price and volume indices. The strength and also the weakness of the approach is that it is strictly descriptive and does not make any assumptions about market behavior while allowing all possible market interactions to be reduced into a joint distribution of all quality characteristics. We shown how this same theoretical base may be used for apparently different quality correction methods.

The third and fourth sections describe the empirical data used for calculations of different hedonic regressions and quality corrected price indices for digital cameras. We also discuss variations of two methods; a classification method, and a matched model method. These two methods are tied together because of their similarities in quality adjustment. In section four, available quality characteristics are introduced and their properties discussed. We show how binary and other classification variables are used to make analysis of variance –type interference possible, although it is rarely used this way in price index compilation.

Estimated hedonic models are provided in section six together with a hedonic decomposition. Three types of hedonic regression models are calculated based on assumptions on temporary restrictions given to the regression coefficients of quality characteristics. Alternative pooling strategies and estimation methods are shortly discussed with practical consequences of temporary restrictions.

In the last section, we provide a summary of the results. We use different models and compare the resulting price indices.

1. The Economy

Empirical research on hedonic price indices often starts with a theoretical motivation. The same is true in this paper. Our viewpoint is more descriptive than economic, and we do not tackle a large part of fundamental questions involved in general regression analysis. Instead, we are more interested in analyzing the effects of different models in resulting indices rather than the assumptions underlying economic theories. In consequence, all price indices are viewed as based on estimating conditional expectations of prices at two consecutive periods.

1.1 Defining the economy

We start with a set of characteristics (quality variables) that fully describes every good and service produced and purchased in the economy $\{X_0, X_1, ..., X_K\}$[1]. The set includes variables specifying technology, inputs, price, quantity and all transaction-specific

[1] This finite but arbitrary large set of quality characteristics fully describes all goods and services. Hence, it also describes conditions for every transaction made in finest detail. The value of these characters form the basis for same quality. In the sequel it is assumed that the characteristic space can be represented by a finite set of numerical characteristics.

characteristics needed. We will call the first character price and denote $X_0 \equiv P$. All possible combinations of these characteristics form a set called the characteristic space and is denoted by Ω. The numerical representation of the characteristic space is clearly not a continuous one, but complicated combination of categorical and continuous variables within their proper domains. The elements $a_i^t \in \Omega$, $a_i^t \equiv a_i\left(P = p, X_1 = x_1, ..., X_K = x_K, T = t\right)$ representing the possible transactions of goods and services at certain time are called items or elementary outcomes. They are the possible outcomes of economic activity at any period. An outcome of the economy at time *t* is a set of realized items at the period in question denoted by $a^t \equiv \left\{a_1^t, ..., a_{n^t}^t\right\}$. A sequence of these periodic outcomes is called the economy. The set Ω remains intact from period to period but the size of each outcome – the number of items – varies.

Since the economy consists of materialized price - quality characteristics observations, it could be described as a transaction economy. To put it more formally, it is a sequence of multiple outcomes in characteristics space.

Definition 1: The Economy

The economy is an ordered sequence of outcomes $\left\{a^t\right\}_{t\in T}$ where

i) each outcome is a set of individual items $a^t = \left\{a_1^t, ..., a_{n^t}^t\right\}$ with each item $a_i^t \in \Omega$,

ii) n^t is the number of items at time $t \in T$, $T \subset \mathrm{N}$,

iii) the aggregate value of the economy at time *t* is the sum of all individual items' prices $V^t = \sum_{i=1}^{n^t} p_i^t$.

This definition is purely descriptive. The value is a simple sum of all transactions. Although we examine all transactions, by conditioning on some characteristics denoting the type of transaction we could restrict ourselves to study e.g. just consumer purchases. Later it will be shown that by partitioning the characteristics space into divergent subgroups the total value may also be calculated by summing over the sums of the partition[2]. The price indices may be constructed accordingly. For now, we may assume that the partition is so dense that no two transactions (items) may be considered identical. So far, no mechanism that determines the outcomes is provided. Loosely following the lines of Pakes (2002) we describe the mechanism as a reduced form relationship defined by whatever generates the relationship between all characteristics, including price. Still, it may be useful to think how to measure this economy and present it in terms of price and volume evolution. The fundamental problem is to somehow decompose the change in the aggregate value of the economy at two consecutive periods into changes in prices and changes in volume. For example, we want the value relative to be presented as a product of a price index and volume index:

(1) $V^t / V^{t-1} = P_{t-1}^t \times Q_{t-1}^t$.

[2] The partition would be based on values of a set of characteristics.

From the definition it is clear that the total value of the economy may change with number of transactions and quality attributes changing. Since the value is based only on the monetary measure of price, we want the price index to measure just price. As is customary, the price index will describe the pure price change while volume index counts for changes in everything else i.e. quantity and quality. Hence, we feel it is natural to try to find just a price change measure. The volume index in (1) is determined as a residual. The entity of volume is rather obscure but we do not want to define the value as product of price, quantity and quality. Although this is possible, and corresponding price-, quantity- and quality indices could be defined, we choose not to do so[3]. Instead, we direct our attention to the price index.

Now, let's assume there exists a probability structure described by a cumulative joint distribution $F^t\left(p, x_1, \ldots, x_K\right) \equiv F\left(p, x_1, \ldots, x_K \middle| T = t\right)$. As was noted before, this conditional distribution is determined by all market conditions of supply and demand, changing consumer preferences and production technology, etc. It is full, reduced characterization of the relationships between all characteristics, including price. Also, let's assume that the economy at time *t*, as described above, is a realization of a stochastic process determining the economy[4]. This way we may treat the observations in Ω as a sample (or universe) drawn from the joint distribution.

As our goal is to measure the changes in aggregate value of the economy in terms of price and volume indices, we move onto the price indices. Rather than wanting to describe the joint distribution or the stochastic process in detail, we are more interested in conditional distribution of price

(2) $F^t\left(p \middle| x\right) \equiv F\left(p \middle| X_1 = x_1, \ldots, X_K = x_K, T = t\right)$.

We will base all our price indices on this distribution. Quality correction methods introduced later assume, explicitly or implicitly, something about the structure of this distribution. For example, if no relationship between price and quality characteristics exists, we could use a relative of simple average prices as a measure of price index:

(3) $P_{t-1}^t = \left(\frac{1}{n^t}\sum_{i=1}^{n^t} p_i^t\right) \Big/ \left(\frac{1}{n^{t-1}}\sum_{i=1}^{n^{t-1}} p_i^{t-1}\right)$ or $P_{t-1}^t = \exp\left[\left(\frac{1}{n^t}\sum_{i=1}^{n^t} \ln p_i^t\right) - \left(\frac{1}{n^{t-1}}\sum_{i=1}^{n^{t-1}} \ln p_i^{t-1}\right)\right]$.

Clearly, in a more realistic case this would not count for any changes in the characteristics. Another solution would be to classify the items into somehow similar groups and base the price index on group averages. This would be the same as to assume only some categorical variables to affect the price. However, as literature of quality change in price indices show, this does not seem to work very well either. Something else is needed.

---

[3] See Pursiainen (2000) for index number analogy for three factor indices.

[4] In this paper, we do not elaborate further either on the definition of the economy or on the structure and properties of the stochastic process defined in $\Omega \times T$.

Whatever the structure of the distribution, we propose a price index based on a conditional expected price change. For practical purposes, we will examine changes in log-price and want to base our price index on expected log-price changes. Let's denote the conditional expected price change as

$$(4)\quad \begin{aligned} E\left(\ln \dot{p}\middle|H\right) &\equiv E\left[\left(\ln p\middle|H,T=t\right)-\left(\ln p\middle|H,T=t-1\right)\right] \\ &= E\left(\ln p\middle|H,T=t\right)-E\left(\ln p\middle|H,T=t-1\right) \\ &\equiv E\left(\ln p^{t}\middle|H\right)-E\left(\ln p^{t-1}\middle|H\right), \end{aligned}$$

where the conditioning set H defines the desired point in the characteristics space we want to condition with. Natural choices for H would be $H=\left\{E\left(X_1,...,X_K\right)\middle|T=t\right\}$, $H=\left\{E\left(X_1,...,X_K\right)\middle|T=t-1\right\}$ or $H=\left\{E\left(X_1,...,X_K\right)\middle|T=t,t-1\right\}$. Equally well we could condition on realized average quality $H=\left\{\left(\bar{x}_1^t,...,\bar{x}_K^t\right)\right\}$ or on all data $H=\left\{x_1^t,...,x_{n^t}^t\right\}, x_i^t=\left(x_{i1}^t,...,x_{n^t k}^t\right)$.

1.2 Full information price index

Let's assume we know the marginal distribution but do not have the realizations of the process, i.e we do not have the transactions of the economy. We could use the above expected log-price change and define our price index e.g. as

$$(5)\quad P_{t-1}^{t}=\exp E\left(\ln \dot{p}\middle|E(X^{t-1})\right),$$

where $E\left(X^{t-1}\right)\equiv E\left(\left(X_1,...,X_K\right)\middle|T=t-1\right)$ is the expected quality vector at time t-1. Alternatively, we could condition on expected quality at period t or expected quality over the two periods $E\left(\bar{X}\right)\equiv E\left(\left(X_1,...,X_K\right)\middle|T\in(t,\,t-1)\right)$.

The distribution at time *t*, F^t may depend on the materialized path of the economy until time *t-1*. Especially over a long period one would expect this to be the case. For example, a technological innovation at time t-1 affects the conditional distribution and hence all future (expected) states of the economy. Since there is nothing stochastic about already realized states of the economy, one might want to base the price index also on the actual transactions.

Now, let's assume we also know all transactions at both periods. What would be a natural measure for the overall price change with full information?

A solution is to condition on the materialized quality data. We could ask what the average price change would be for, let's say all period t-1 items. Now, for each period t-1 item we may form an expected log-price change

(6) $E\left(\ln \dot{p}_i \left| x_i^{t-1}\right.\right) \equiv E\left(\ln p_i^t \left| x_i^{t-1}\right.\right) - \ln p_i^{t-1}$,

and define the price index as

(7) $P_{t-1}^t = \exp\left(\sum_{i=1}^{n^{t-1}} E\left(\ln \dot{p}_i \left| x_i^{t-1}\right.\right)\right)$.

Similarly, we could reprice all period t transactions at time t-1 and define the price index as

(8) $P_{t-1}^t = \exp\left(\sum_{i=1}^{n^t} E\left(\ln \dot{p}_i \left| x_i^t\right.\right)\right) = \exp\left(\sum_{i=1}^{n^t}\left(\ln p_i^t - E\left(\ln p_i^{t-1}\left| x_i^t\right.\right)\right)\right)$.

We will get back to these indices in section 3 in more detail. First, we want to consider some practicalities in constructing the price indices as above.

We obviously do not have the pleasure of knowing the joint distribution or the conditional price distribution of all transactions in the economy. Also, the "*everything affects everything while everything is changing*" -approach is clearly not very fruitful in actual index compilation. Without too much loss of generality we argue that some quality characteristics affect only certain type of items. We interpret each original quality character X_k as a vector $X_k = \left(X_{k1},...,X_{km^k}\right)$ and assume that e.g.

(9) $F\left(p^t \left| x_1\right.\right) = F\left(p^t \left| x_1,...,x_K\right.\right), \forall\left(x_2,...,x_K\right) \in \Omega_{-(P,X_1)}$.

If this holds for all groups in the partition, then

(10) $F\left(p^t \left| X\right.\right) = F\left(p^t \left| X_1\right.\right) \times \cdots \times F\left(p^t \left| X_K\right.\right)$

and we can examine each group separately and determine the price index for each group. Of course, first we need to decide in which product groups prices might be calculated individually. Second, we need to gather sample data on prices and on potential quality characteristics. It is by no means clear that we should not examine (instead of an individual product) a group of products (e.g. home electronics) together and estimate the hedonic function for the whole group instead of just one product. For our purpose this practical and important question is, however, not relevant and the discussion in the next section does not make explicit the level of aggregation.

2. Conditional expectation and hedonic price indices

We now turn to the estimation of the conditional expectation (of log-price). The distribution $F^t\left(\ln P \middle| H\right)$ from which the expectation is derived will not be discussed further and the model behind all types of hedonic indices is described simply by[5]

(11) $\left(h(P) \middle| x,t\right) = g^t\left(x\right) + \varepsilon^t$

where the transformation function *h* is not generally known[6]. Because of some nice properties of logarithm function and congruence with geometric means, we use natural logarithm as the transformation function for price in all hedonic regression models. To simplify the notation we denote $\log P = p$. Empirical evidence also supports this choice in many cases[7]. Now, as in previous section we are interested in the systematic part of (11), i.e. the conditional expectation

(11') $E\left(\ln P \middle| x,t\right) = E\left(p \middle| x,t\right) = g^t\left(x\right).$

The function *g* gives an estimate of log-price for each point in quality space Ω_{-p}. At this point, nothing is assumed on the time specific functions g^t or the independent variables in *x*. Individual transformation functions (that may differ from identity or logarithm function) of original quality characteristics are applied to end up with vector $x_i^t = \left(1, x_{1i}^t, \ldots, x_{Ki}^t\right)$. The K characteristics are treated as "measured in transformed form" and the question of model specification is not discussed further[8].

One should note that this representation allows us to use flexible functional forms such as translog and quadratic models[9]. In all models, we assume the estimated function to be linear with respect to the parameters of the transformed original quality variables. Now, let function *f* be simply an estimate of the unknown relation g^t:

(12) $f^t(x) \equiv est\left[g^t\left(x\right)\right] = est\left[E\left(\log P \middle| x,t\right)\right].$

For our purposes we will use the following linear (with respect to parameters) functional form:

(13) $f^t(x) = \hat{\beta}^{t\prime} x,$

[5] One could also argue, as usually is done, that the "real" relation is an inversed model from which we transform this one just for sake of ease of estimation.
[6] This approach of estimated function follows Vartia and Koskimäki (2001).
[7] See e.g. Diewert (2002) and a summary treatment in Triplett (2002) or IMF(2002).
[8] For functional forms and economic approach to hedonic indices, see Rosen (1974), Diewert (2001) and Triplett (2002). Here the question of model selection and functional form is not discussed further.
[9] See Diewert (2001)

where the ($K+1$) β–vector includes a constant term.

For the purposes of this study the estimation method used does not have to be OLS as long as it forces the sum of residuals to zero at each period[10]. The forecast error connected with each observation is $e_i^t = p_i^t - f^t(x_i^t) = p_i^t - \hat{\beta}^{t} x_i^t$. We denote $\hat{p}^t_{x_i^{t-1}} = \hat{\beta}^{t} x_i^{t-1} = f^t\left(x_i^{t-1}\right)$ as period t estimated price for period t-1 observation i, and similarly $\hat{p}^{t-1}_{x_i^t} = \hat{\beta}^{t-1} x_i^t = f^{t-1}\left(x_i^t\right)$ as period t-1 estimated price for period t observation i. Also, as in the previous section, we denote the estimated log-price change as $\dot{p}$ for both $\hat{p}^t_{x_i^{t-1}} - p_i^{t-1} = \dot{p}_i$ and $p_i^t - \hat{p}^{t-1}_{x_i^t} = \dot{p}_i$.

The hedonic price indices using the above hedonic regression models are all weighted geometric means of estimated log-price changes of either base or reference period observations. The estimated log-prices are derived either using the reference period hedonic model for base period observations or base period model for the reference period observations. We have chosen to include also the weights in the table below, although in the empirical part we will only use equally weighted means. We call these weighted indices hedonic log-Laspeyres (1) and "log-Paasche" (2) respectively[11]. By taking geometric mean of the two indices, we get the formula for hedonic Törnqvist index. The equally weighted geometric mean is also called Jevons index.

Table 1.1 Hedonic logarithmic indices

Index	Formula	Weight	Estimated price
Log-Laspeyres (1)	$\log P^t_{t-1} = \sum_{i=1}^{N^{t-1}} w_i^t \left(\hat{p}^t_{x_i^{t-1}} - p_i^{t-1}\right)$	Period *t-1*	Period *t* using period *t-1* observations
"Log-Paasche" (2)	$\log P^t_{t-1} = \sum_{i=1}^{N^t} w_i^t \left(p_i^t - \hat{p}^{t-1}_{x_i^t}\right)$	Period *t*	Period *t-1* using period *t* observations
Hedonic Törnqvist	$P^t_{t-1} = \sqrt{\exp(1) \times \exp(2)}$	both	both

We separate five different types of hedonic models based on the model specification. Regardless of the type, indices may always be calculated using the above formulas even if it may not be the most efficient way.

i) *The time indicator model* assumes that the function *g* changes in time only in the constant term. The quality adjusted price index may be constructed directly from the estimates of time coefficients.

ii) *The classification model* or analysis of variance -type model uses only categorical variables. The quality adjusted price index may be constructed directly from class means implied by the classification (and cross classification).

[10] An example of trivial zero sum residual "estimation" of *g* is a relation *f* that gives the observed value for each observation *i*.

[11] The "Paasche formula" is really just reversed Laspeyres.

iii) *The patched model* uses imputation for those observations that are missing for the period in question and real observations for models whose prices can be found in both periods[12].

iv) *Full imputation* uses two sets of prices for same models at two periods. (The observed period *t* value may also be the best estimate for period *t*.)

v) *Mixed models* may include elements of i), ii) and iv).

The aim of this exercise is twofold. First, we want to assess the effect of quality adjustment on the price index and to show how the model specification affect the quality adjustment and the index. Second, we want to study the implied accuracy of the hedonic index and compare the indices based on different types of models.

All indices are treated as changes between two consecutive periods. The index strategy is to chain the index over longer periods. In some cases, an equivalent fixed base period index would be identical and give exactly the same results. The source of weights is not specified here and it is just assumed that they sum to one $\sum_i w_i = 1$.

Before going into the hedonic models in more detail, we discuss the data for digital cameras.

[12] The term Patched model was used by Pakes (2001).

3. The data

The quarterly price data were recorded from various sources, mainly from issues of Journal of Popular Photography on microfilm. For two last quarters of 2002, the price data were collected from the internet at *www.pricescan.com*. The model specific quality characteristics were also compiled from different sources but mainly from the website *dpreviw.com*.

Total of over 1300 prices of 288 different digital camera models were collected. Some of these prices were averaged over the advertising retailers at the time of entry. Although the share of these multiple observations were not recorded, they account for approximately 10 to 15 per cent of the total sample. Regardless of the number of observations used for each recorded model price quote, they are treated as single observations.

3.1 The sample design

It was not feasible to use random sampling in the study because of scarce data sources. For the early years, all possible models with a price quote in the Journal were recorded. When the price collection was changed to internet, generally the lowest price was recorded and almost all available makes and models were included. There were no effort to follow same models, and market entries and exits occurred when new models were first advertised or they were no longer available. In other words, the sample was not designed to mimic any typical statistical agency approach.

After the data collection, we decided to exclude SLR[13] type digital cameras and also restrict the time sample to the years 1999 through 2002. The SLR-type digital cameras are used by professionals and they are much more expensive and have partly different properties and characteristics. The data from years 1996 and 1997 proved to be too limited.

The quality characteristics may have some variation between retailers. However, just one reference attribute is used throughout the lifespan of a model over all retailers. These differences could not be observed in detail and are most likely limited to different memory cards included in cameras. For this reason, memory was excluded from some of the regression models.

The number of priced models (and hence price observations) follow the number of adds in the Journal and the number of different models each manufacturer makes. One might argue that the data is self-weighting in a sense that the more models a manufacturer has, the larger its share is in the data. No explicit data was available for the weights and all indices are calculated as equally weighted geometric means. In digital camera case, we assume that the number of different models advertised is a proxy for market share of that make. This view may be challenged, but no other data was available to support or contradict this hypothesis.

[13] SLR - single lens reflex.

3.2 Data reservations

There are some issues in the data that might be a hindrance to the index:

1) *"Call for price".* A number of times the Journal adds do not show the price directly, and they could only be obtained by calling the retailer. This practice is still in use and may be set by either the manufacturer or the retailer.
2) *Weight data.* There were no data on model nor manufacturer turnover or other data on relative importance of manufacturers.
3) *Unbalanced samples.* The price sample is unbalanced towards the last year and the last two quarters when the price collection changed fundamentally.
4) *Limited data on the early years.*

These concerns were treated by the following practices, which are commonly used. However, their effects should always be estimated.

1) *"Call for price".* These prices were not collected at all. Our assumption is that the "Call for price" is more widely used with new introductory models than older models. Although the basis for reasoning is not relevant in our descriptive indices, one may argue that models entering the market have higher mark-ups while models exiting are being sold out with smaller or even negative markups. Since we are not referring to marginal cost –pricing, as we are only interested in the actual prices paid, we assume that there are no differences in the price determining processes (or marginal distributions) between the advertised and non-advertised prices.
2) *Weight data.* As already noted, since we had no data on quantities sold by model or by manufacturer, we assumed that the number of models advertised provides a proxy for the manufacturer weights. Thus, each model has an equal weight. Out of all observations Sony counted for most (18%) and then Olympus (15%), Kodak (13%), Fuji (12%), Canon (11%) and Nikon (8%) of total observations. The rest 13 manufacturers account for the rest 24 % of observations. This self-weighting seems reasonable – especially since there were not great differences in pricing between manufacturers.[14] However, it would be very interesting to see how model specific weights would have affected the series.
3) *Unbalanced samples.* There is a clear change in the number of observations per quarter when the data source was changed to the internet after the second quarter of 2002. There may be a systematic reduction in prices due to this change, since the data selected from Pricescan.com usually refer to "Best price" which is the lowest advertised price within a selection of online retailers. These prices usually do not include shipping and there may be a tendency for the low price retailers to charge more for the shipping than others. However, this is clearly not the case every time and the same may also be true for the Journal adds as well. The degree of possible bias from this has not been quantified. There were total of 1052 price quota for the years 1998 – 2002 and the distribution of observations is presented in table 3.1.

[14] As tested with classification models.

Table 3.1 The distribution of price observations

Quarter / Year	1998	1999	2000	2001	2002
Q1	19	41	79	75	50
Q2	20	34	44	29	94
Q3	27	54	63	83	155
Q4	37	31	55	41	124
Total	103	160	241	228	423

The change in the data collection is likely to have an effect also on the index. The prices collected from the internet are usually the lowest prices for each model and not averaged over advertised prices, as is the case in the price data from the Journal adds. By collecting overlapping prices, one could estimate the magnitude of a shift change but this exercise was not carried out. Another feature of internet purchases is the shipping and handling fees. Although not included in the prices, it could be argued that in some cases a part of the actual price is actually charged as shipping and handling, which is rather evident with regard to some special offers of other consumer goods. However, a quick sample did not confirm the negative relationship between price of the good and the handling fees for same models.

4) *Limited data on the early years*. There are additional 21 price observations for 1996 and 38 for 1997 that have not been used in the analysis, because these data include missing information on the characteristics. With a distinctive model or other method the series could be extended a few years back, but for this study this exercise was not carried out. Also, if the time period were further reduced to, say the last three years, differences between quality adjusted and unadjusted price indices would not appear so large[15].

These questions, as important as they may be, are not addressed further in this study. Our main purpose was to compare different methods of controlling quality and the index series they produce. Some of the questions could have been solved with some complementary data or additional data collection.

[15] This will become apparent later.

4. Average prices and characteristics

As a group, digital cameras (compact or ultra compact models) is a rather homogenous set of product varieties (compared to some other transactions in the economy). For many high technology goods, the average price does not seem to change very much but at the same time average (technological) quality characteristics change considerably. This is true for digital cameras too.

4.1 Average price

The price for a typical new digital camera model often starts with a stable introductory price (maybe set by the manufacturer) and then the dispersion of offer prices becomes larger in time. Often the highest asking price stays the same (or decreases moderately) while the lowest price declines sharply. In the data, there is just one price for a model at one time, but additional sales information would be interesting to obtain.[16]

Without controlling for quality changes the average price fluctuates around $500 until early 2001 and then drops to under $350 during the last year. These prices are direct arithmetic averages (geometric means are by virtue always smaller, but when presented as relative changes they part very little). Changes in quarterly average prices are presented in figure 4.1 as index series.

Figure 4.1 Average prices as index series (1998 = 100)

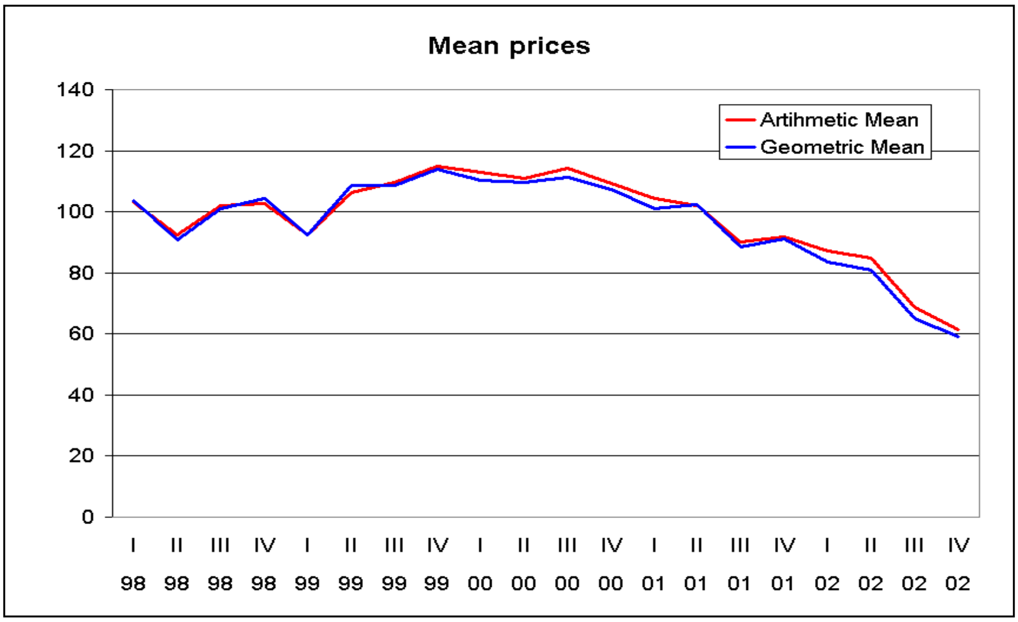


Figure 4.1 confirms the intuition one has with digital cameras. The prices of 'a good camera' have been around $500 for some years and just during the last few years the prices have

[16] See e.g. www.pricescan.com.This property brings some additional problems for the price index that were not accounted for in section 3. The data collected does not allow to take into account this behavior.

begun to drop, though less than the picture would indicate[17]. Of course, these prices do not count for the changes in performance and other characteristics of the equipment. Equally importantly, these average prices do not follow the prices of same camera models. The drop during the last few quarters may be partly explained with the change in data collection – both because the recorded price refers to the lowest price and also because there may be more ‘low end’ models in the data set.

In terms of the model in section 1, calculating unadjusted geometric means is equal to estimating a hedonic regression model

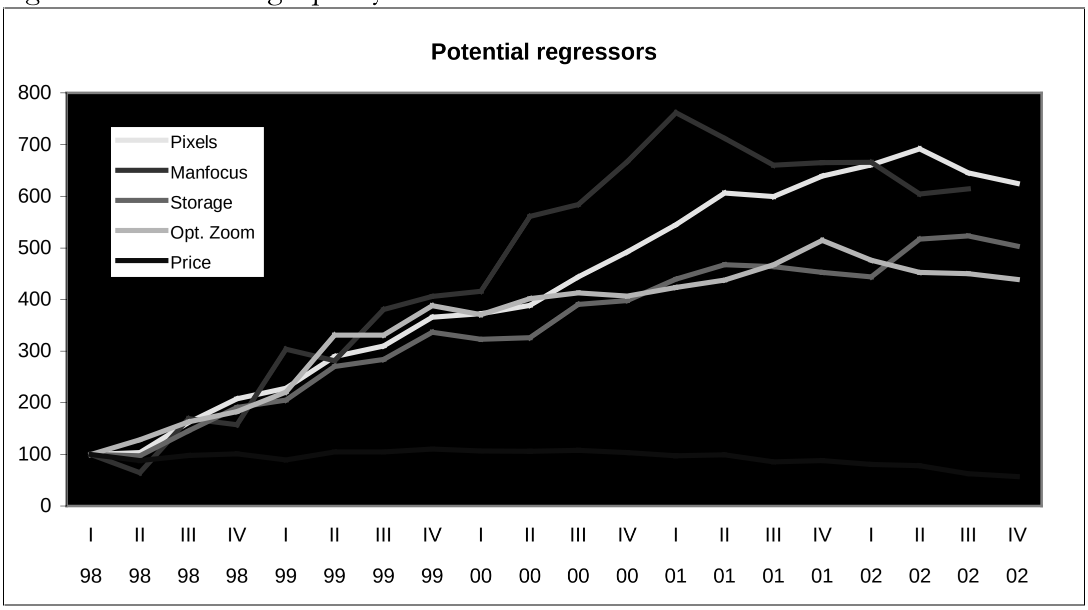
(13) $f^t(x) = \hat{\beta}^t x^t = \hat{\alpha}^t$.

This can be estimated from pooled data consisting all quarters and the resulting model fit measured in adjusted R^2 is little over 20% and most of individual t-test statistics for time indicator coefficients suggest that they do not deviate from *0*[18]. We call this a trivial model. It is actually used by some statistical agencies in imputing missing observations, though not usually explicitly in this form.

4.2 Quality characteristics

The regression models to be laid out in section 5 need input variables, namely regressors. Potentially good explanatory variables for hedonic regression would be such that have variability in time and relatively large spread within each time period. Additionally, they should be correlated with the price and not heavily with each other. In figure 4.2, there are some potential variables to be used in the regression models in form of a simple average relative change thus showing only the variability in time.

Figure 4.2 Some average quality variables as indices



[17] We believe the same goes with some other high-tech equipment, e.g. PC’s. For some years now you’ve got a ‘good computer’ – not the very state of the art, but the next best thing – for some $1500.

Changes in the average quality characteristics are large in time compared to the evolution of average prices. Of these variables, the manual focus is a binary variable and the index presentation should be interpreted as the evolution of the share of digital cameras having the feature in question. When used in average form e.g. in imputation, all binary variables should be interpreted similarly. All available variables in the data set are presented in table 2 in Appendix 1. The most useful are the ones indicating a sharpness of the picture, a memory capacity, an optical zoom ratio, and manual focus, an external flash and movie options.

[18] Actually tests for 0-assumption of indicators for just last two quarters would be rejected at 5% confidence level if the first quarter is used as a reference.

5. Matched model and classification method

The usual practice statistical agencies use in tackling the changes in quality comes as a by-product of the sampling or selection process. Just follow the prices of same goods in time! In case of digital cameras, the ‘same’ would mean the same store and same model of the same make. No measurable quality change to control for – by definition. This could be interpreted as using a tight classification where each model forms one class. The index is based on the class average prices. For the sake of illustration, we classify (ex post) the cameras to similar (homogenous) groups and calculate the group mean price changes. This classification method is actually often used by statistical agencies for missing observations or replacements and may be a good method in connection with some products.

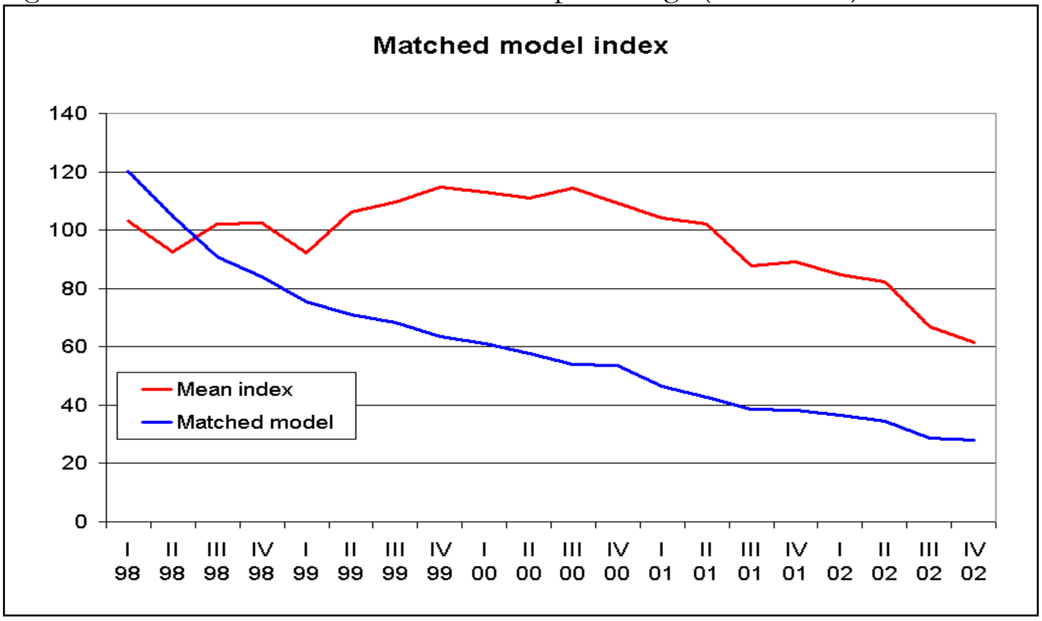
5.1 A matched model index

Before using the classification method, we first calculate a digital camera price index based on a matched model method. As mentioned above in section 3.1, the sampling frame was not meant to be used for calculation of a statistical agency –type matched model index. In a sense, this sampling could be described as “quarterly re-sampling”[19]. Since we did not initially plan to calculate the matched model index at all, it is provided only as an example and methodological criticism should not focus on inadequacy of this matched model method. Also, no patched models were compiled from the data.

There were observations from at least two quarters for almost all of the 288 models in the data set. However, the turnover of camera models was rather fast. The number of models for which price were found over more than 4 quarters’ period was 167 and over 6 quarters’ just 52. In traditional statistical agency practice, this would have meant a large number of replacement models to be found and an alternative method to account for those models at times of no price observation. Our matched model index does not include estimates for the missing models, either for the ones entering or exiting. The observed price change over more than one quarter is divided by the number of quarters and addressed only to the first quarter. With these reservations, the (log mean) index series are presented in figure 4.1 together with an unadjusted average price index. As can be seen, especially in the first six to eight quarters, the two methods differ considerably, and the matched model index is much smoother in decline. We will get back to some interpretation in section 6.

[19] It would not be true to claim that the samples were truly independent from one another since we used the same magazine having mostly the same advertisers over time. However, the notion of independence should not be too far from true and we would expect it to have only minor effect for the final outcome – the price index.

Figure 5.1 The matched model index and simple average (1998 = 100)



5.2 The classification method

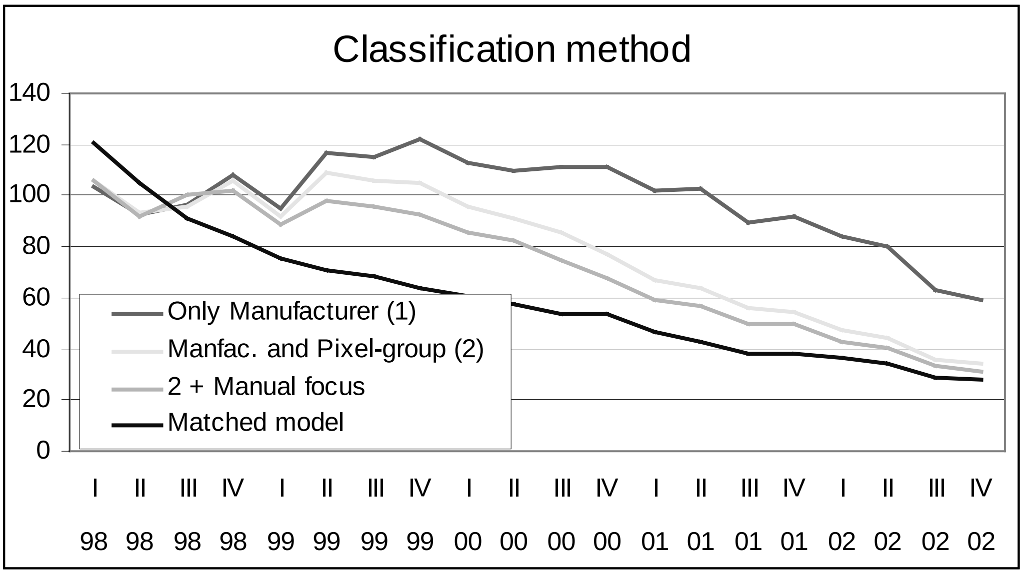
The classification method is based on classifying the cameras according to some rule – most likely by their characteristics – and calculating the class means. The matched model index above is an extreme case of this method. Each model is classified as its own group and 'empty' classes appear every time when the model is not found in the next period.

A more realistic case would be to use, for example, the manufacturer as a classification rule. With the classification, one hopes to remove as much as possible of the within class price variation at each time. This method is typically used in repeated measurement experiments in natural sciences and called analysis of variance. It is usually not called that way in the price index context and usually none of the available test statistics from the analysis are either estimated or presented.

The first classification model in figure 5.2 is based on the manufacturer. The index may be calculated from the changes in make-specific average prices[20]. The second model adds pixel group – a variable, which classifies the camera models into five categories according to the available accuracy of the picture (less that 1 megapixel, 2 Mp, 3MP and over 4 Mp cameras). The third model further classifies the data to models with or without a manual focus option (autofocus is the norm). As with unadjusted averages, we do not actually calculate the class means but instead use regression models (without cross effects).[21]

[20] The actually calculation is based on a time dummy regression model with indicator variables for each manufacturer.
[21] So we are actually not cross-classifying the models, but only using the 'main effects'.

Figure 5.2 The classification method



Compared to the matched model index, these indices have more volatility around the trend in the first two years. Adding the classifying factors clearly smoothens the series, but practical usability suffers because the number of class means to be calculated grows exponentially and empty classes start to appear.

The classification method may be presented in the notation of chapter 1 as follows:

$$(14)\quad f^{t}(x) = \hat{\beta}'^{t} x^{t} = \alpha + \beta_1 D_{MANUF} + \beta_2 D_{MF} + \cdots + \sum_{t=2}^{T} \gamma^{t} D_{t},$$

where Ds are binary variables for classifying variables and time indicator variables. When estimated from pooled data the R-squares for the three models above are 33, 62 and 69% respectively. As already mentioned, the index series may be constructed directly from the time indicator coefficients. However, the same resulting series is also reached by estimating the price for each observations and using the index number formulas presented in table 1.1.

6. Hedonic methods

As already noted in section 2, all models use natural logarithmic of price for the dependent variable and either logarithmic or identity transformation functions for the explanatory variables. In the first sub-section we introduce the index formula decompositions. I the next section we estimate models restricting the coefficients of all quality characteristics to be constant over time, while in 5.3 these restrictions are relaxed.

6.1 Decomposition of hedonic price index

This section provides an illustrative way of showing the quantitative effect of hedonic quality adjustment. The proof of formula (15) is given in appendix 2 which also discusses the implications with patched model index. Some of these decompositions are presented in section seven together with regression results from the data.

It can be show that, regardless of the type of the hedonic model, hedonic log-Laspeyres index can be presented as

$$(15)\ P_{t-1}^{t}(La) = \exp\left[\sum_{i=1}^{N^{t-1}} \left(w_i^{t-1} \left(\hat{p}_i^t - p_i^{t-1} \right) \right) \right]$$
$$= \frac{G(P^t)}{G(P^{t-1})} \times \exp\left(\hat{\beta}'^t \left(\bar{x}^{t-1} - \bar{x}^t \right) \right) \times \exp\left(\operatorname{cov}\left(\frac{w}{\bar{w}}, \dot{p} \right) \right)$$
$$= \frac{G(P^t)}{G(P^{t-1})} \times \exp\left(\hat{\beta}'^t \left(\bar{x}^{t-1} - \bar{x}^t \right) \right) \times \exp\left[\operatorname{cov}\left(\frac{w}{\bar{w}}, \left(\hat{\beta}'^t - \hat{\beta}'^{t-1} \right) x^{t-1} \right) \right] \times \exp\left[-\operatorname{cov}\left(\frac{w}{\bar{w}}, e^{t-1} \right) \right].$$

This decomposition, though slightly differently, was first presented in Koev and Suopera[22]. It says that the quality adjusted log-Laspeyres index may presented as a product of relative of geometric average prices at two periods, a separable quality correction term that depends on change in average quality, and a sample covariance term between mean adjusted weights and estimated pure log-price change. The last row of (15) breaks the covariance term into two parts; one that shows the effect of model differences at two periods and a term between the adjusted weight and period *t-1* estimation error. In case of a time indicator model, the first correlation term is zero, because the two models are same. The hedonic Törnqvist index in (16) (derived and discussed more in Appendix 2) uses the average regression coefficients over two periods and may be presented as

$$(16)\ P_{t-1}^{t}(To) = \sqrt{P_{t-1}^{t}(La) \times P_{t-1}^{t}(Pa)}$$
$$= \frac{G\left(P^t\right)}{G\left(P^{t-1}\right)} \times \exp\left(\bar{\hat{\beta}}' \left(\bar{x}^{t-1} - \bar{x}^t \right) \right) \times \frac{1}{2} \left(\operatorname{cov}\left(\frac{w^{t-1}}{\bar{w}^{t-1}}, \hat{p}^t - p^{t-1} \right) + \operatorname{cov}\left(\frac{w^t}{\bar{w}^t}, p^t - \hat{p}^{t-1} \right) \right).$$

With equal weights, as in the digital camera data, the whole covariance factor drops out and leaves just the quality adjustment factor. Also, in case of a time indicator model the coefficient estimate of β is not time specific and hedonic log-Laspeyres and hedonic Törnqvist indices are identical. As said, the quality correction term may be separated into effects of each quality characteristic and each effect may be presented separately:

$$(17)\ \hat{\beta}' \left(\bar{x}^{t-1} - \bar{x}^t \right) = \hat{\beta}_1 \left(\bar{x}_1^{t-1} - \bar{x}_1^t \right) + \ldots + \hat{\beta}_K \left(\bar{x}_K^{t-1} - \bar{x}_K^t \right).$$

[22] Koev and Suopera (2002).

It's clear that the quality correction factor depends on the regression model, but also – and equally importantly – on the changes of average quality. If some quality characteristics are good explanatory variables in cross sectional regression model, but the average values do not change, the quality correction factor does not affect the index! On the other hand, even large average quality changes may not affect the index, if there is only a small or no relationship between the price and the quality character[23].

What is important here is that we can compare the quality correction factors of different model types (classification, full hedonic and even patched model). Further, with equally weighted indices only changes in average quality matter.

It's also interesting that an equally weighted matched model index is identical to any hedonic index if calculated on the same data.[24] This, of course, assumes that there is no average quality change either, and hence no quality correction is needed in the first place! More interestingly, with entering and exiting product varieties, the degree of quality correction needed in addition to matched model may be factored out.[25]

6.3 Time indictor models

In this section all period t models are estimated as

$$\text{(18)}\quad \ln\left(\hat{p}_i^t\right) = \hat{\beta}^t x_i^t = \hat{\alpha} + \hat{\beta}_1 x_{i1}^t + \cdots + \hat{\beta}_K x_{iK}^t + \sum_{t=1}^{T} \hat{\delta}^t D_t \ ,$$

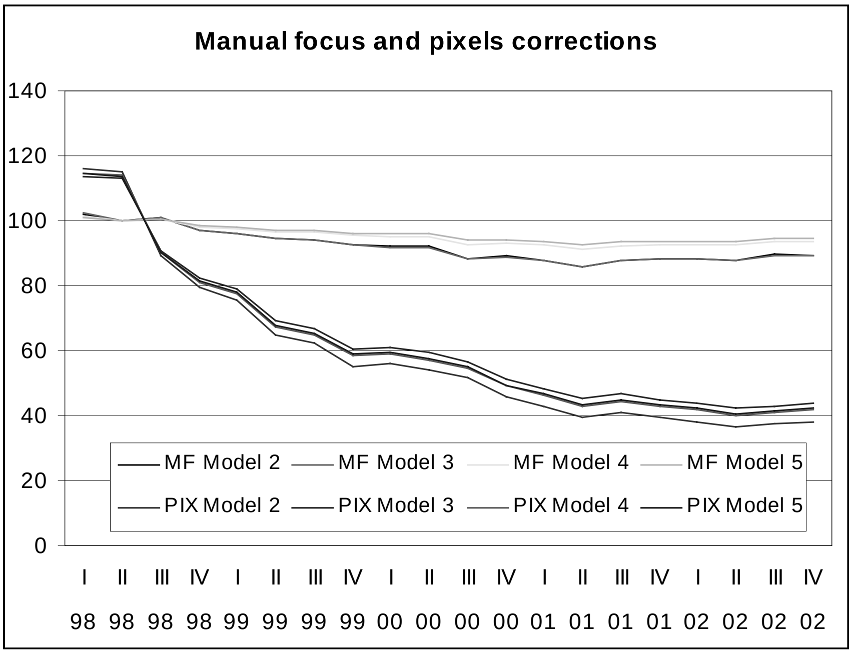
where an indicator variable D_t gets value 1 at period t and 0 otherwise. We will use four slightly different models to illustrate how the model selection affects the quality correction and the index. The models are:

Model 1: manual focus + ln(megapixels) [R-square 66%]
Model 2: manual focus + ln(megapixels) + ln(megabytes) [75%]
Model 3: manual focus + ln(megapixels) + optical zoom [77%]
Model 4: manual focus + ln(megapixels) + optical zoom + ln(megabytes) + external flash [80%]

Adding more quality characteristics increase the overall model fit somewhat, and the differences with quality correction factors from each model become even smaller. A summary table for the coefficients and quality correction terms of the below models are presented in appendix 4. In figure 6.1 quality correction terms for two characteristics, manual focus and pixels, are presented.

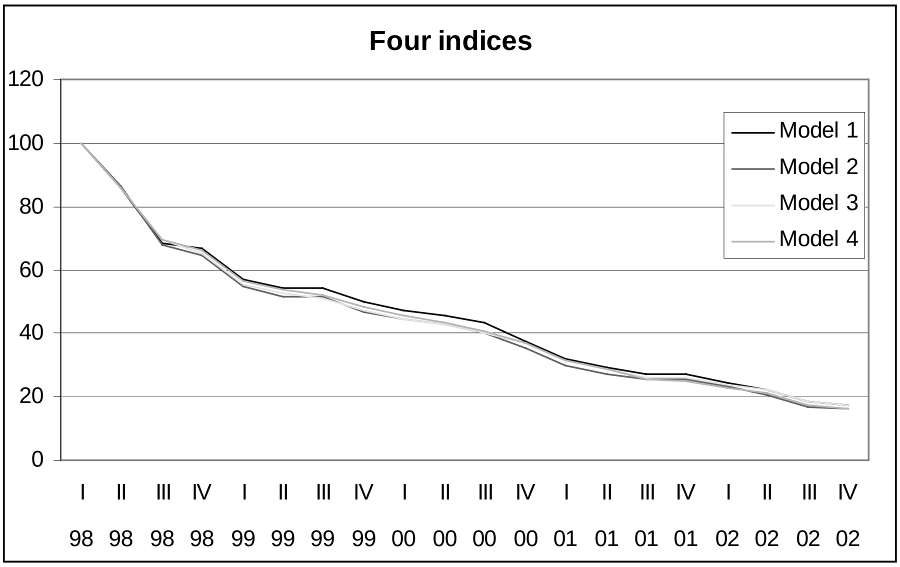
[23] It is debatable if it could then be called a quality character in the first place.
[24] As also stated by Triplett (2002) and Diewert(2001).
[25] See Appendix 2 for Patched model case.

Figure 6.1 Quality correction factors for four models



As one can see there are some variation between the four models when the OLS adjusts the hyperplane in price-quality coordination. The feature of forecast model is that by adding explanatory variables into the model the individual coefficients, and quality correction factors, adjust so that best overall fit is achieved. This means that the individual quality corrections factors "give" a part of their value to the new variable depending on the amount of multicollinearity it has with the variables already in the model. However, when taken together with all quality characteristics in the model, the total quality correction factor may have very little 'dispersion' between models, as in figure 6.2.

Figure 6.2 Quality adjusted log-Laspeyres indices for the four models



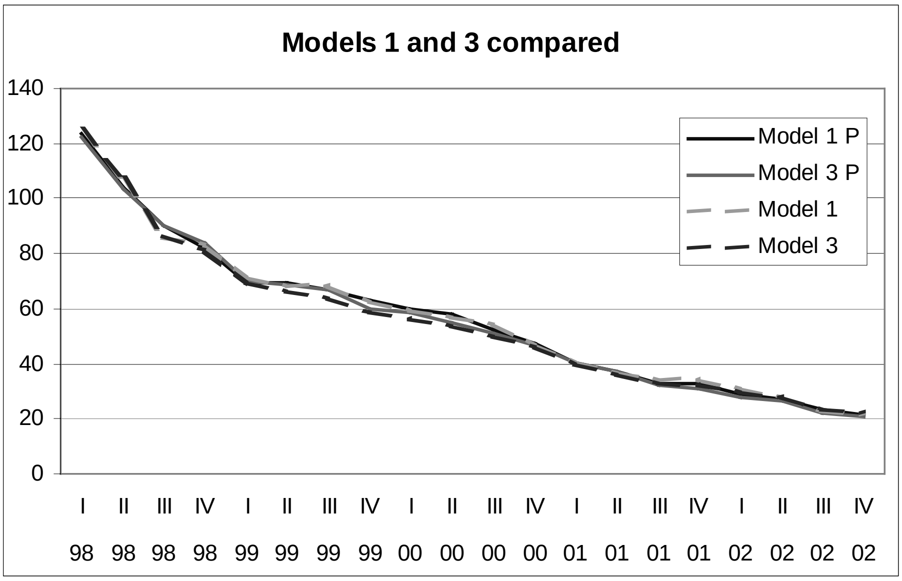
The overall picture of the quality adjusted price index for 1999 – 2002 does not change when adding new quality variables into the hedonic regression. Also, for practical reasons it may be feasible to collect even high frequency data on just few quality characteristics together with an existing price collection, e.g. the CPI.

6.4 Relaxing the temporary restrictions

In case of rapid quality change, the assumption that quality – price relation stays the same except from the constant term over a long period of time should be questioned. We will use two different methods to allow more flexibility in the models. The first one is to apply the time indicator model to data that pools data together only two consecutive periods, and estimate 19 different time indicator models (for all pairs). The second is to estimate separate models for each 20 quarters and impute the matching prices as suggested by the index formulas. With the latter we calculate different log-Laspeyres an "log-Paasche" indices using the model from period *t* and *t-1*, respectively and present the hedonic Törnqvist price index. We call these models pairwise pooled and full hedonic models.

As the results will show, again there are no large changes in the quality adjusted indices and thus we will use just two models, Model 1 and Model 3 from the previous section.
Now, the model R2:s vary between 60 and 85%. See a summary table of estimation results in Appendix 4. The resulting index series for pairwise pooled indices are presented in figure 6.3 together with ones from the previous section. The P refers to pairwise model, and as one can see, the two quality correction magnitudes are very similar with the completely pooled data models.

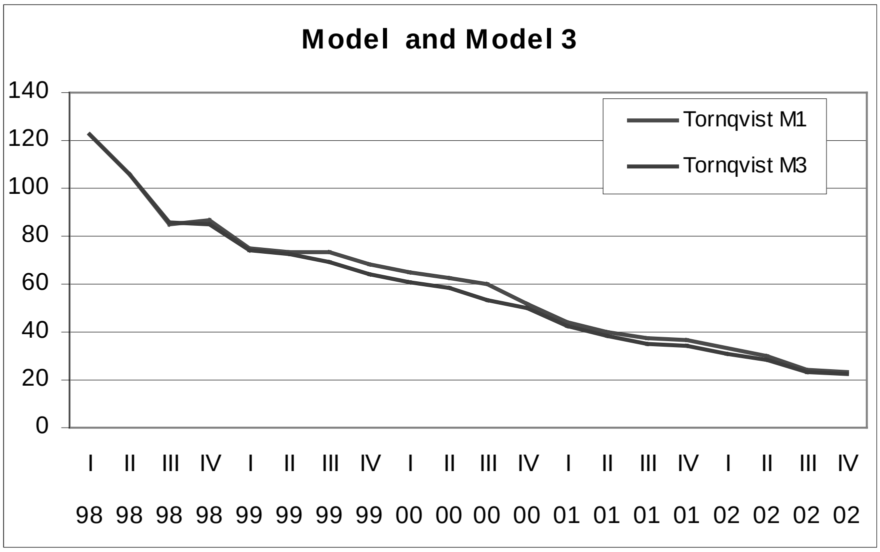
Figure 6.3 Pooled estimation models



In this case individual observations have much more effect for the coefficients, especially in the early years. Consequently, the quality correction factors for individual characteristics do vary little more, but the total quality correction factors are not affected as much, as expected. One could modify this method by adding the number of consecutive periods to the estimation, which would have a further smoothing effect but still gradually take into account possible changes in quality – price relation.

The second method could be called a true full hedonic method. Estimated models stay the same but instead of calculating just one index we will use all the data and present the resulting quality adjusted price indices as a Tornqvist index[26]. This would theoretically be the most comparable with an index based on time indicator model[27]. See result in the summary table in Appendix 4. The two indices are presented in figure 6.4.

Figure 6.4 Full hedonic models



While it may be difficult to see real difference in the quality adjusted price indices, there are some differences in the quality correction factors. The difference between the two model pairwise quality correction factors is some 8% at maximum. With the pooled data the difference is at most 6%. With individually estimated Fisher quality correction factors between the two model differ again at most some 8%.

[26] Difference of chained Laspeyres and Paasche are at most 7 index points and average to very close to 0.
[27] Since both use the data from the two periods to estimate the model(s). See appendix 1.

7. Summary

As shown above, regardless of different methods and hedonic models used to adjust for quality changes in digital camera price index, all reasonable models produce indices that come very close to each other. Since we apply a forecast model we prefer a simpler model if no substantial benefit is achieved from adding more variables into the hedonic model. If one has to choose between a simple model and more complicated one we think the simpler is better.

As a starting point we expect and allow the models differ between periods. Only if there is no strong evidence against changing coefficients, we may use time restricted models. If data allows the use of pooled regressions, hedonic models seem to be rather robust in choice of estimation and of model specification. The advantages from using somehow pooled data are that it gives more stable regression coefficients and 'smoothens' the quality adjusted index series. Especially with high frequency indices it may also make the hedonics more feasible since it demands less data. This is again not a bad thing.

Since relatively simple models may work well enough, hence large scale characteristics collection may not be necessary, and may not be any greater restraint for statistical agencies than a matched model approach.

As long as the matched model index does not produce outside the sample bias it works very well. However, if the joint distribution of characteristics changes in time, as it does with high technology products, frequent sampling is needed and quality changes in mismatches must be dealt somehow. Typical statistical agency procedures may not be suitable for simultaneously dealing with quality change and sampling. We argue that hedonic approach is a good and often feasible way to produce indices so that sampling may separated from the quality adjustment process.

References

Diewert, W. E. (2001), Hedonic Regressions: A Consumer Theory Approach, Discussion Paper No.: 01-12, University of British Columbia, Vancouver

Diewert, W. E. (2002), Hedonic Producer Price Indexes and Quality Adjustment, Discussion Paper No.: 02-14, University of British Columbia, Vancouver

Heravi, S. and Silver, M. (2002), On the Stability of Hedonic Coefficients and their Implications for Quality-Adjusted Price Change Measurement, Cambridge, Mass: NBER Summer Institute 2002.

Hulten C. R. (2002): Price Hedonics: A Critical Review. NBER Summer Institute 2002, Cambridge, Mass.

Feenstra, R.C. (1995), Exact Hedonic Price Indexes, Review of Economics and Statistics, LXXVII, 634-54.

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

IMF (2002): PPI Manual

Koev, E. and Suopera, A. (2002), Omakotitalojen ja omakotitalotonttien hintaindeksit 19985=100, Statistics Finland report.

Koskimaki, T. and Vartia, Y. (2001),

Vartia, Y.O. and Vartia, P.L.I. (1984), Descriptive Index Number Theory and the Bank of Finland Currency Index, Scandinavian Journal of Economics 86 (3), 352-64.

Pakes, A. (2002): A reconsideration of Hedonic Price Indexes with an Application to PC's, Harward University.

Pursiainen, H. (2000): Kolmen faktorin indeksit. University of Helsinki.

Rosen, S. (1974): Hedonic Prices and Implicit Markets: Product Differentation in Pure Competition

Silver, M. and Heravi, S. (2001), Scanner Data and the Measurement of Inflation

Triplett, J.E. (2002), Handbook on Quality Adjustment of Price Indexes for Information and Communication Technology Products, OECD Directorate for Science, Technology and Industry, Draft, OECD Paris

Appendix 1. Average prices and the classification method

Table A1.1 Average prices

Year	Quarter	Quarterly mean prices $	Mean index
98	I	528.9	103.0
98	II	474.5	92.4
98	III	524.1	102.0
98	IV	527.0	102.6
99	I	474.4	92.4
99	II	546.0	106.3
99	III	562.4	109.5
99	IV	590.1	114.9
00	I	580.5	113.0
00	II	569.4	110.9
00	III	587.0	114.3
00	IV	561.4	109.3
01	I	535.7	104.3
01	II	524.2	102.1
01	III	463.6	87.7
01	IV	471.6	89.2
02	I	447.3	84.6
02	II	436.0	82.4
02	III	353.4	66.8
02	IV	315.8	61.5

Table A1.2 Data set variabels

Variable	Description	type of measure
lnp	log of price	dollars
lnpix	log of sharpness	megapixels
lnsto	log of memory included	megabytes
movie	movie feature	0 - 1 variable
remote	remote control	0 - 1 variable
flash_ex	external flash	0 - 1 variable
manfocus	manual focus	0 - 1 variable
zoomo	optical zoom	scale of optical magnification
zoomd	digital zoom	scale of digital magnification
USB	usb connection	0 - 1 variable
serial	serial connection	0 - 1 variable
bat_re	battery recharger	0 - 1 variable
type	type of camera	compact, ultacomp, SLR-type
multires	choices of various resolutions	number, or '0 - 1 variable
ISO	number of different iso	number, or '0 - 1 variable
manufac	manufacturer	18 manufacturers

Table A1.3. Classification method

Year	Quarter	Only Manufacturer (1)	Manufacturer and Pixel-group (2)	2 + Manual focus
1998	I	103.4	105.4	106.0
1998	II	92.3	93.3	91.9
1998	III	96.4	95.4	100.4
1998	IV	108.0	105.8	101.7
1999	I	94.6	92.0	88.9
1999	II	116.5	108.6	97.9
1999	III	114.8	106.0	95.3
1999	IV	122.2	104.9	92.6
2000	I	112.8	96.0	85.4
2000	II	109.8	91.3	82.8
2000	III	111.3	85.3	74.9
2000	IV	110.9	77.3	67.9
2001	I	102.0	66.5	59.0
2001	II	102.5	63.9	56.4
2001	III	89.6	55.8	49.9
2001	IV	91.9	54.4	49.5
2002	I	83.9	47.7	43.1
2002	II	80.1	44.1	40.4
2002	III	63.3	36.0	33.5
2002	IV	58.7	33.9	31.5

Appendix 2. **Decomposition of hedonic geometric indices**

It was proposed that hedonic log-Laspeyres index may be decomposed as:

$$(A1)\ \ P_{t-1}^{t}(La) = \exp\left[\sum\nolimits_{i=1}^{N^{t-1}} \left(w_i^{t-1} \left(\hat{p}_{x_i^{t-1}}^{t} - p_i^{t-1} \right) \right) \right]$$

$$= \frac{G(P^t)}{G(P^{t-1})} \times \exp\left(\hat{\beta}'^{t} \left(\overline{x}^{t-1} - \overline{x}^{t} \right) \right) \times \exp\left(\operatorname{cov}\left(\frac{w}{\overline{w}}, \dot{p} \right) \right).$$

To see this, we develop the basic hedonic log-Laspeyres formula by adding and subtracting estimated equally weighted pure price change

$$(A2)\ \ P_{t-1}^{t} = \exp\left[\sum\nolimits_{i=1}^{n} \left(w_i^{t-1} \left(\hat{p}_i^{xt} - p_i^{t-1} \right) \right) \right]$$

$$= \exp\left[\sum\nolimits_{i=1}^{n} \left(\left(w_i^{t-1} - \frac{1}{N^{t-1}} \right) \left[\hat{p}_{x_i^{t-1}}^{t} - p_i^{t-1} \right] \right) + \frac{1}{N^{t-1}} \sum\nolimits_{i=1}^{n} \left(\hat{p}_{x_i^{t-1}}^{t} - p_i^{t-1} \right) \right]$$

$$= \exp\left[\sum\nolimits_{i=1}^{n} \left(\left(w_i^{t-1} - \frac{1}{N^{t-1}} \right) \left[\hat{p}_{x_i^{t-1}}^{t} - p_i^{t-1} \right] \right) \right] \times \exp\left(\frac{1}{N^{t-1}} \sum\nolimits_{i=1}^{n} \left(\hat{p}_{x_i^{t-1}}^{t} - p_i^{t-1} \right) \right).$$

Now manipulating the second part by again adding and subtracting period *t* average price it may be written as:

$$(A3)\ \ \exp\left[\frac{1}{N^{t-1}} \left(\sum\nolimits_{i=1}^{N^{t-1}} \hat{p}_{x_i^{t-1}}^{t} - \sum\nolimits_{i=1}^{N^{t-1}} p_i^{t-1} \right) \right]$$

$$= \exp\left(\hat{\beta}'^{t}\, \overline{x}^{t-1} - \overline{p^{t-1}} + \left(\overline{p^{t}} - \overline{p^{t}} \right) \right)$$

$$= \exp\left(\overline{p^{t}} - \overline{p^{t-1}} \right) \times \exp\left(\hat{\beta}'^{t}\, \overline{x}^{t-1} - \hat{\beta}'^{t}\, \overline{x}^{t} \right)$$

$$= \frac{G\left(P^t\right)}{G\left(P^{t-1}\right)} \times \exp\left(\hat{\beta}'^{t} \left(\overline{x}^{t-1} - \overline{x}^{t} \right) \right).$$

The first term is the relative of geometric means of the two periods' prices. It is an unadjusted or unit price index from period *t-1* to *t.* The second term is a multiplicative quality correction term that may be further factored into each characteristic. After estimating the regression coefficients, this equally weighted decomposition may easily be used for index calculation, since the quality correction term only depends on the average quality change.

The second term of (A1) may be written as

$$(A4)\quad \exp\left[\sum_{i=1}^{N^{t-1}}\left(w_i^{t-1}-\frac{1}{N^{t-1}}\right)\left[\hat{p}^{t}_{x_i^{t-1}}-p_i^{t-1}\right]\right]$$

$$=\exp\left[\frac{1}{N^{t-1}}\sum_{i=1}^{N^{t-1}}\left(\frac{w_i^{t-1}}{\overline{w}^{t-1}}-1\right)\left[\hat{p}^{t}_{x_i^{t-1}}-p_i^{t-1}\right]\right]$$

$$=\exp\left[\operatorname{cov}\left(\frac{w^{t-1}}{\overline{w}^{t-1}},\left[\hat{p}^{t}_{x_i^{t-1}}-p^{t-1}\right]\right)\right]$$

$$=\exp\left[\operatorname{cov}\left(\frac{w}{\overline{w}},\dot{p}\right)\right]$$

Also, to emphasize the hedonic model, we could further decompose the covariance term into *systematic and random* parts. Since the average weight depends on the number of observations, we may also write $\operatorname{cov}\left(w/\overline{w},\dot{p}\right)=\operatorname{cov}\left(N^{t-1}w,\dot{p}\right)$. To see the effect of selecting the model type this covariance term may further be

$$(A5)\quad \exp\left[\operatorname{cov}\left(\frac{w^{t-1}}{\overline{w}^{t-1}},\left[\hat{p}^{t}_{x_i^{t-1}}-p^{t-1}\right]\right)\right]$$

$$=\exp\left[\operatorname{cov}\left(\frac{w}{\overline{w}},\left[\hat{p}^{t}_{x_i^{t-1}}-\hat{p}_i^{t-1}-e_i^{t-1}\right]\right)\right]$$

$$=\exp\left[\operatorname{cov}\left(\frac{w}{\overline{w}},\hat{p}^{t}_{x_i^{t-1}}-\hat{p}_i^{t-1}\right)+\operatorname{cov}\left(\frac{w}{\overline{w}},-e_i^{t-1}\right)\right]$$

$$=\exp\left[\operatorname{cov}\left(\frac{w}{\overline{w}},\left(\hat{\beta}'^{t}-\hat{\beta}'^{t-1}\right)x^{t-1}\right)+\operatorname{cov}\left(\frac{w}{\overline{w}},-e_i^{t-1}\right)\right].$$

Now it's easy to see, that for any time indicator model, for which regression coefficients stay constant, the covariance is between the weights and forecast error. Error may of course depend on how many periods are used in the estimation of time indicator model.

To derive a symmetric hedonic index that makes use of both period weights and regression models we start with repricing period *t* observations using period *t-1*model. It is easy to show that this "Paasche type" formula is almost the same:

$$(A6)\quad P_{t-1}^{t}(Pa)=\exp\left[\sum_{i=1}^{N^{t}}\left(w_i^{t}\left(p_i^{t}-\hat{p}^{t-1}_{x_i^{t}}\right)\right)\right]=\cdots$$

$$=\frac{G\left(P^{t}\right)}{G\left(P^{t-1}\right)}\times\exp\left(\hat{\beta}^{t-1}\left(\overline{x}^{t-1}-\overline{x}^{t}\right)\right)\times\exp\left[\sum_{i=1}^{N^{t}}\frac{1}{N^{t}}\left(w_i^{t}-\frac{1}{N^{t}}\right)\left[p_i^{t}-\hat{p}^{t-1}_{x_i^{t}}\right]\right]$$

$$=\frac{G\left(P^{t}\right)}{G\left(P^{t-1}\right)}\times\exp\left(\hat{\beta}^{t-1}\left(\overline{x}^{t-1}-\overline{x}^{t}\right)\right)\times\exp\left(\operatorname{cov}\left(\frac{w}{\overline{w}},\dot{p}\right)\right).$$

Finally, the hedonic Törnqvist index may be obtained as a geometric average of the two. The traditional Törnqvist index formula uses arithmetic mean of weights as

$$(A7)\ P_{t-1}^{t}(T\ddot{o}) = \exp\left(\sum_{i=1}^{N} \left(w_i^t + w_i^{t-1} \right) \big/ 2 \left(p_i^t - p_i^{t-1} \right) \right).$$

But since we may have a different number of observations in the two periods we define the hedonic Törnqvist index as geometric mean of hedonic log-Laspeyres (A1) and the current period weighted "Paasche-type" index (A6). Using notation $\overline{\hat{\beta}}$ for average of the two period estimated regression coefficients it is rather straightforward to show that:

$$(A8)\ P_{t-1}^{t}(T\ddot{o}) = \sqrt{P_{t-1}^{t}(La) \times P_{t-1}^{t}(Pa)} = \frac{G\left(P^t\right)}{G\left(P^{t-1}\right)}$$

$$= \frac{G\left(P^t\right)}{G\left(P^{t-1}\right)} \times \exp\left(\overline{\hat{\beta}}' \left(\overline{x}^{t-1} - \overline{x}^{t} \right) \right) \times \exp \frac{1}{2} \left(\operatorname{cov}\left(\frac{w^{t-1}}{\overline{w}^{t-1}}, \hat{p}_{x^{t-1}}^{t} - p^{t-1} \right) + \operatorname{cov}\left(\frac{w^{t}}{\overline{w}^{t}}, p^{t} - \hat{p}_{x^{t}}^{t-1} \right) \right),$$

where the last term could be simplified - if the number of observations would stay the same – as

$$(A9)\ \frac{1}{2}\left(\operatorname{cov}\left(\frac{w^{t-1}}{\overline{w}^{t-1}}, \hat{p}^{t} - \hat{p}^{t-1} \right) + \operatorname{cov}\left(\frac{w^{t}}{\overline{w}^{t}}, \hat{p}^{t} - \hat{p}^{t-1} \right) \right) + \frac{1}{2}\left(\operatorname{cov}\left(\frac{w^{t-1}}{\overline{w}^{t-1}}, -e^{t-1} \right) + \operatorname{cov}\left(\frac{w^{t}}{\overline{w}^{t}}, e^{t} \right) \right)$$

$$= \operatorname{cov}\left(\frac{1}{2}\frac{w^{t-1}}{\overline{w}^{t-1}} + \frac{1}{2}\frac{w^{t-1}}{\overline{w}^{t-1}}, \hat{p}^{t} - \hat{p}^{t-1} \right) + \frac{1}{2}\left(\operatorname{cov}\left(\frac{w^{t}}{\overline{w}^{t}}, e^{t} \right) - \operatorname{cov}\left(\frac{w^{t-1}}{\overline{w}^{t-1}}, e^{t-1} \right) \right)$$

$$= \operatorname{cov}\left(w_T, \dot{p} \right) + \frac{1}{2}\left(\operatorname{cov}\left(\frac{w^{t}}{\overline{w}^{t}}, e^{t} \right) - \operatorname{cov}\left(\frac{w^{t-1}}{\overline{w}^{t-1}}, e^{t-1} \right) \right).$$

In (A9) the weights are now Törnqvist weights and also the regression coefficients used in the quality adjustment term follow the Törnqvist. Actually, any index formula based on log-change may be decomposed in the above way, just the covariance terms associated with the weighting scheme change.

Again, in case of any time indicator model (using estimated prices on both periods) or any equally weighted index, the last two terms disappear. Then the decomposition simplifies to

$$(A10)\ P_{t-1}^{t}(To) = \sqrt{P_{t-1}^{t}(La) \times P_{t-1}^{t}(Pa)} = \frac{G\left(P^t\right)}{G\left(P^{t-1}\right)} \times \exp\left(\overline{\hat{\beta}}' \left(\overline{x}^{t-1} - \overline{x}^{t} \right) \right),$$

from which it's clear that for time indicator models the Törnqvist index is identical to Laspeyres. This happens because in the covariance terms $\operatorname{cov}\left(w, \hat{p}^t - \hat{p}^{t-1}\right)$ the estimated price change is a constant (coefficient for the time dummy).

For equally weighted *patched model* it may be shown that the difference between full hedonic method depends on the average forecasting error for the missing models. Shortly, let the number of models in the index be constant N of which L are the same in both periods. For these models the observed matched price relative are used. For the rest N-L models new models are selected and they replace the old ones. Hedonic Laspeyres index (or more correctly, hedonic Jevons index) using replacements for missing N-L observations may be written as

$$(A11)\ P_{t-1}^{t}(La) = \exp\left[\frac{1}{N}\left(\sum_{i=1}^{L}\left(p_i^t - p_i^{t-1}\right) + \sum_{i=L+1}^{N}\left(\hat{p}_{x_{MIS}^{t-1}}^{t} - p_i^{t-1}\right)\right)\right]$$

$$= G\left(P^{t-1}\right)^{-1} \exp\left(\frac{1}{N}\left(\sum_{i=1}^{L}\left(p_i^t\right) + \sum_{i=1}^{N-L}\left(\hat{p}_{x_{MIS}}^{t}\right)\right)\right)$$

$$= G\left(P^{t-1}\right)^{-1} \times \exp\left(\frac{L}{N}\overline{p_{MAT}^{t}} + \frac{N-L}{N}\hat{\beta}'^{t}\,\overline{x}_{MIS}^{t-1}\right)$$

$$= G\left(P^{t-1}\right)^{-1} \times \exp\left(\frac{L}{N}\overline{p_{MAT}^{t}} + \frac{N-L}{N}\hat{\beta}'^{t}\,\overline{x}_{MIS}^{t-1} + \left(\frac{N-L}{N}\overline{p_{NEW}^{t}} - \frac{N-L}{N}\overline{p_{NEW}^{t}}\right)\right)$$

$$= \frac{G\left(P^{t}\right)}{G\left(P^{t-1}\right)} \times \exp\left(\frac{N-L}{N}\left(\hat{\beta}'^{t}\,\overline{x}_{MIS}^{t-1} - \overline{p_{NEW}^{t}}\right)\right)$$

$$= \frac{G\left(P^{t}\right)}{G\left(P^{t-1}\right)} \times \exp\left(\frac{N-L}{N}\left(\hat{\beta}'^{t}\,\overline{x}_{MIS}^{t-1} - \hat{\beta}'^{t}\,\overline{x}_{NEW}^{t} - \overline{e}_{NEW}^{t}\right)\right)$$

$$= \frac{G\left(P^{t}\right)}{G\left(P^{t-1}\right)} \times \exp\left(\hat{\beta}'^{t}\left(\overline{x}^{t-1} - \overline{x}^{t}\right) - \frac{N-L}{N}\overline{e}_{NEW}^{t}\right).$$

The last row becomes from the fact that for matched models the average quality change is zero and hence $\left(\overline{x}^{t-1} - \overline{x}^{t}\right) = \frac{L}{N}\left(\overline{x}_{MAT}^{t-1} - \overline{x}_{MAT}^{t}\right) + \frac{N-L}{N}\left(\overline{x}_{MIS}^{t-1} - \overline{x}_{NEW}^{t}\right)$. So, if the new models' average forecasting error does not considerably differ from zero the patched model is identical with full hedonic model. The magnitude of this error is easy to confirm.

The case where number of observations stays the same is limiting but interesting, since it is often used in statistical agencies. Also, if the replacements are chosen to be close substitutes, as is often the case, the average quality difference is likely to be small. Then the overall quality adjustment would be determined by the forecasting error (and the share of new models) for the new model and the direction and magnitude of quality correction would depend on the average forecasting error. Before drawing conclusions on market behavior of the new models, one should make sure there is good, non-estimation reason that $\bar{e}^t_{NEW} \neq \bar{e}^t_{MAT}$.

Of course, also the type of model and inclusion or exclusion of replaced models in the estimation of the regression coefficients affect the quality correction.

Appendix 3. Some regression results

Time indicator Model 4 estimation results. The model is estimated as

$$\ln\left(\hat{p}_i^t\right) = \hat{\beta}^t x_i^t =$$

$$\hat{\alpha} + \hat{\beta}_1 LNPIX + \hat{\beta}_2 LNSTO + \hat{\beta}_3 MANFOCUS + \hat{\beta}_4 ZOOMO + \hat{\beta}_5 FLASH\_EX + \sum_{t=10}^{28} \hat{\delta}^t Q_t$$

The SAS System 09:33 Monday, May 12, 2001

The REG Procedure
Model: Model 4
Dependent Variable: lnp

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	24	170.35142	7.09798	160.64	<.0001
Error	938	41.44658	0.04419		
Corrected Total	962	211.79800			

Root MSE	0.21020	R-Square	0.8043
Dependent Mean	6.05545	Adj R-Sq	0.7993
Coeff Var	3.47133		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	6.56490	0.07584	86.57	<.0001
lnpix	ln(PIXEL)	1	0.50031	0.02129	23.50	<.0001
Manfocus	Manfocus	1	0.13551	0.01722	7.87	<.0001
ZoomO	Opt Zoom	1	0.06494	0.00562	11.56	<.0001
lnsto	ln(STORAGE)	1	0.12207	0.01539	7.93	<.0001
Flash_ex	Ext flash	1	0.11777	0.01872	6.29	<.0001
Q10		1	-0.16623	0.09449	-1.76	0.0789
Q11		1	-0.34812	0.08709	-4.00	<.0001
Q12		1	-0.49367	0.08241	-5.99	<.0001
Q13		1	-0.69502	0.08191	-8.48	<.0001
Q14		1	-0.77254	0.08240	-9.38	<.0001
Q15		1	-0.80633	0.07965	-10.12	<.0001
Q16		1	-0.90879	0.08450	-10.76	<.0001
Q17		1	-0.93108	0.07829	-11.89	<.0001
Q18		1	-0.97396	0.08171	-11.92	<.0001
Q19		1	-1.07752	0.08038	-13.41	<.0001
Q20		1	-1.15870	0.08123	-14.26	<.0001
Q21		1	-1.29944	0.08044	-16.15	<.0001
Q22		1	-1.37906	0.08708	-15.84	<.0001
Q23		1	-1.50989	0.08095	-18.65	<.0001
Q24		1	-1.52072	0.08508	-17.87	<.0001
Q25		1	-1.59569	0.08398	-19.00	<.0001
Q26		1	-1.65902	0.08160	-20.33	<.0001
Q27		1	-1.83195	0.07983	-22.95	<.0001
Q28		1	-1.90307	0.08005	-23.77	<.0001

Appendix 4. **Model 3 and 4 coefficients**

Models 3 and 4 estimated from total and pairwise pooled data

Table A4.1 Parameter estimates for pairwise dummy model and full dummy model 3

Year	Qrt	Intercept	lnpix	Manfocus	ZoomO	Time D
98	I					
98	II	6.81	0.65	0.07	0.06	-0.167
98	III	6.52	0.52	0.13	0.07	-0.135
98	IV	6.33	0.42	-0.06	0.08	-0.079
99	I	6.25	0.43	0.04	0.07	-0.153
99	II	6.09	0.49	0.05	0.08	-0.025
99	III	6.05	0.54	0.14	0.07	-0.047
99	IV	5.99	0.53	0.21	0.07	-0.078
00	I	5.89	0.58	0.16	0.08	-0.037
00	II	5.84	0.60	0.15	0.08	-0.045
00	III	5.83	0.54	0.17	0.07	-0.085
00	IV	5.76	0.55	0.14	0.07	-0.089
01	I	5.66	0.60	0.14	0.07	-0.141
01	II	5.51	0.61	0.16	0.06	-0.089
01	III	5.41	0.59	0.08	0.09	-0.110
01	IV	5.28	0.57	0.10	0.10	-0.015
02	I	5.30	0.51	0.21	0.08	-0.083
02	II	5.15	0.53	0.26	0.09	-0.054
02	III	5.13	0.53	0.23	0.08	-0.186
02	IV	5.03	0.46	0.23	0.08	-0.063
Full model		**6.69**	**0.53**	**0.18**	**0.08**	

Table A4.2 Parameter estimates for pairwise dummy model and full dummy model 4

Year	Qrt	Intercept	lnpix	lnsto	Manfocus	ZoomO	Flash_ex	Time D
98	I							
98	II	6.68	0.65	0.18	0.00	0.04	0.00	-0.153
98	III	6.46	0.51	0.06	0.10	0.07	-0.12	-0.153
98	IV	6.28	0.41	0.04	-0.06	0.08	-0.02	-0.088
99	I	6.17	0.42	0.06	-0.02	0.07	0.03	-0.184
99	II	5.97	0.47	0.08	0.07	0.05	0.12	-0.039
99	III	5.78	0.34	0.11	0.16	0.07	0.22	-0.019
99	IV	5.75	0.35	0.08	0.18	0.10	0.20	-0.095
00	I	5.55	0.47	0.13	0.12	0.11	0.13	-0.015
00	II	5.60	0.53	0.08	0.09	0.12	0.09	-0.054
00	III	5.52	0.46	0.11	0.08	0.13	0.06	-0.089
00	IV	5.37	0.44	0.16	0.06	0.11	0.10	-0.074
01	I	5.40	0.55	0.14	0.11	0.04	0.11	-0.143
01	II	5.26	0.57	0.14	0.14	0.02	0.13	-0.087
01	III	5.22	0.64	0.09	0.05	0.06	0.11	-0.134
01	IV	5.11	0.68	0.07	0.04	0.06	0.13	-0.027
02	I	5.15	0.78	0.01	0.14	0.04	0.09	-0.094
02	II	4.88	0.68	0.10	0.15	0.06	0.08	-0.071
02	III	4.79	0.56	0.15	0.13	0.06	0.12	-0.170
02	IV	4.80	0.49	0.11	0.18	0.05	0.10	-0.073
Full model		**6.56**	**0.50**	**0.12**	**0.14**	**0.12**	**0.06**	

Table A4.3 Parameter estimates for full hedonic model 3

Year	Qrt	Intercept	lnpix	Manfocus	ZoomO
98	I	6.75	0.58		0.07
98	II	6.72	0.71	0.06	0.05
98	III	6.35	0.47	0.13	0.07
98	IV	6.25	0.38	-0.11	0.08
99	I	6.09	0.48	0.13	0.07
99	II	6.08	0.51	0.00	0.08
99	III	5.99	0.54	0.22	0.07
99	IV	5.91	0.51	0.20	0.07
00	I	5.84	0.61	0.14	0.08
00	II	5.79	0.57	0.17	0.08
00	III	5.77	0.53	0.16	0.07
00	IV	5.66	0.59	0.11	0.08
01	I	5.51	0.60	0.17	0.06
01	II	5.42	0.65	0.16	0.06
01	III	5.28	0.60	0.06	0.10
01	IV	5.30	0.52	0.18	0.09
02	I	5.22	0.50	0.24	0.08
02	II	5.06	0.55	0.28	0.09
02	III	4.98	0.53	0.20	0.08
02	IV	5.01	0.39	0.26	0.07

Table A4.4 Parameter estimates for full hedonic model 4

Year	Qrt	Intercept	lnpix	Manfocus	ZoomO	Insto	Flash_ex
98	I	6.63	0.59		0.05	0.16	
98	II	6.58	0.71	-0.01	0.03	0.19	
98	III	6.36	0.48	0.12	0.08	0.00	-0.05
98	IV	6.12	0.36	-0.11	0.09	0.08	-0.02
99	I	6.01	0.48	0.04	0.07	0.05	0.04
99	II	5.80	0.47	0.12	0.03	0.16	0.14
99	III	5.79	0.33	0.18	0.10	0.06	0.23
99	IV	5.58	0.34	0.18	0.10	0.12	0.18
00	I	5.50	0.52	0.09	0.12	0.13	0.11
00	II	5.72	0.58	0.09	0.13	-0.02	0.05
00	III	5.28	0.38	0.06	0.13	0.19	0.08
00	IV	5.40	0.53	0.07	0.07	0.14	0.10
01	I	5.26	0.57	0.15	0.02	0.14	0.13
01	II	5.19	0.57	0.13	0.03	0.14	0.13
01	III	5.08	0.66	0.01	0.07	0.08	0.10
01	IV	5.15	0.73	0.07	0.05	0.02	0.18
02	I	5.04	0.80	0.17	0.04	0.00	0.00
02	II	4.72	0.62	0.14	0.07	0.15	0.12
02	III	4.67	0.54	0.13	0.05	0.15	0.11
02	IV	4.88	0.43	0.24	0.05	0.06	0.07

Table A4.5 Average quality characteristics

Year	Qrt	lnp	lnpix	Manfocus	ZoomO	lnsto	Flash_ex
98	I	6.22	-0.96		0.59	0.85	
98	II	6.09	-0.93	0.09	0.75	0.82	
98	III	6.20	-0.49	0.06	0.96	1.22	0.06
98	IV	6.23	-0.23	0.15	1.08	1.49	0.04
99	I	6.11	-0.14	0.14	1.30	1.56	0.14
99	II	6.27	0.10	0.28	1.95	1.84	0.34
99	III	6.27	0.17	0.26	1.95	1.89	0.33
99	IV	6.32	0.33	0.35	2.28	2.06	0.42
00	I	6.29	0.35	0.37	2.18	2.02	0.34
00	II	6.28	0.39	0.38	2.36	2.03	0.35
00	III	6.30	0.53	0.51	2.43	2.21	0.41
00	IV	6.26	0.63	0.53	2.39	2.23	0.33
01	I	6.20	0.73	0.61	2.49	2.33	0.33
01	II	6.21	0.84	0.69	2.58	2.39	0.46
01	III	6.07	0.83	0.65	2.75	2.38	0.34
01	IV	6.10	0.89	0.60	3.03	2.36	0.34
02	I	6.01	0.92	0.60	2.80	2.34	0.26
02	II	5.98	0.97	0.60	2.66	2.49	0.22
02	III	5.76	0.90	0.55	2.65	2.50	0.18
02	IV	5.66	0.87	0.56	2.59	2.46	0.20

Table A4.6: Quality correction factors

Year	Qrt	Model 2P	Model 2Full	Model 4P	Model 4Full	Model 2	Model 4
98	I						
98	II	-0.04	-0.05	-0.02	-0.05	-0.03	-0.02
98	III	-0.24	-0.25	-0.26	-0.30	-0.27	-0.31
98	IV	-0.11	-0.16	-0.12	-0.18	-0.12	-0.13
99	I	-0.06	-0.06	-0.06	-0.09	-0.06	-0.06
99	II	-0.18	-0.20	-0.20	-0.26	-0.18	-0.20
99	III	-0.03	-0.03	-0.02	-0.04	-0.03	-0.03
99	IV	-0.13	-0.13	-0.14	-0.16	-0.13	-0.14
00	I	-0.01	-0.01	0.02	0.01	-0.01	0.02
00	II	-0.04	-0.04	-0.05	-0.05	-0.04	-0.05
00	III	-0.10	-0.10	-0.10	-0.12	-0.10	-0.10
00	IV	-0.06	-0.06	-0.04	-0.05	-0.06	-0.04
01	I	-0.08	-0.07	-0.08	-0.08	-0.08	-0.08
01	II	-0.08	-0.08	-0.10	-0.09	-0.09	-0.10
01	III	-0.01	0.00	0.01	0.00	0.00	0.02
01	IV	-0.06	-0.05	-0.06	-0.06	-0.06	-0.06
02	I	0.00	0.00	-0.01	0.02	0.00	-0.01
02	II	-0.01	-0.01	-0.04	-0.02	-0.01	-0.03
02	III	0.05	0.05	0.05	0.04	0.05	0.05
02	IV	0.02	0.02	0.02	0.03	0.02	0.02

These are derived as $\sum \hat{\beta}_1'\left(\overline{x}_1^{t-1} - \overline{x}_1^{t}\right) + \ldots + \hat{\beta}_K'\left(\overline{x}_K^{t-1} - \overline{x}_K^{t}\right)$

Table A4.7 Quality correction factors

Year	Qrt	Model 2P	Model 2Full	Model 4P	Model 4Full	Model 2	Model 4
98	I	117.7	118.5	117.6	121.5	119.1	120.2
98	II	113.4	115.0	114.9	117.6	115.2	117.5
98	III	89.1	89.9	88.7	87.8	87.7	86.3
98	IV	79.8	76.7	78.8	73.1	78.0	76.0
99	I	75.5	71.8	74.0	67.0	73.7	71.4
99	II	63.3	58.9	60.6	51.6	61.7	58.4
99	III	61.3	57.0	59.4	49.8	59.7	56.9
99	IV	53.7	50.0	51.5	42.3	52.4	49.4
00	I	53.4	49.7	52.4	42.7	52.1	50.3
00	II	51.2	47.9	49.9	40.8	49.9	47.9
00	III	46.4	43.4	45.0	36.3	45.2	43.2
00	IV	43.8	41.0	43.4	34.6	42.6	41.5
01	I	40.6	38.1	39.9	31.8	39.5	38.2
01	II	37.3	35.2	36.1	29.0	36.2	34.6
01	III	37.1	35.2	36.6	29.0	36.2	35.2
01	IV	35.0	33.6	34.5	27.5	34.2	33.2
02	I	35.0	33.6	34.2	28.0	34.2	32.9
02	II	34.6	33.1	33.0	27.3	33.8	31.8
02	III	36.4	34.7	34.6	28.6	35.6	33.5
02	IV	37.0	35.4	35.3	29.3	36.2	34.1