Pricing Over the Product Life Cycle: An Empirical Analysis

Daniel Melser* 
Department of Economics 
Monash University 
Caulfield, VIC 3145 
Australia

Iqbal Syed 
School of Economics 
University of New South Wales 
Sydney, NSW 2052 
Australia

Draft: 23 October 2006

Abstract: This paper explores the extent to which goods follow systematic pricing patterns over their life cycle. The theoretical literature, and anecdotal evidence, suggests that new products are often introduced at high prices which decline as the good ages while, older goods exit the market at a discount. We outline and apply a smoothing-spline approach to the estimation of life cycle pricing effects using data on two different types of goods; supermarket products (beer, canned soup and cereals) and high-tech goods (desktop and laptop computers, and personal digital assistants). We find evidence for the existence of large life cycle pricing effects though there is some heterogeneity across products. The existence of life cycle effects on prices has important implications for the estimation of inflation. If prices change due to life cycle factors then this needs to be accounted for in estimating an index. Using our pricing function, we show that significant bias can arise from the failure to accurately represent the age structure of sampled products.

Keywords: Product life cycle; Hedonic regression; Price index; Spline smoothing.

JEL Classification Codes: C43, C50, D00, E31.

* The authors gratefully acknowledge both the Bureau of Economic Analysis (BEA) and the James M. Kilts Center, Graduate School of Business, University of Chicago who made data available for this project. Any opinions expressed in the paper are those of the authors. Corresponding author: daniel.melser@buseco.monash.edu.au.
1. Introduction

In modern economies there is an extremely wide range of goods from which to choose. Moreover there is a regular churn with varieties appearing and disappearing from retailers shelves at great frequency. These trends are evident right across the product spectrum, from computers to digital cameras, to clothing and canned goods. The purpose of this paper is to look at whether there are trends in the prices of products over their life cycle, from entry to exit. That is, do new goods enter the market at relatively high prices, with retailers taking advantage of the novelty factor to earn a premium at introduction, or do they enter at rock-bottom prices in order to generate a sufficient number of consumer-trials to build a market following?

Strangely, while the theoretical work in this area of intertemporal price discrimination is well developed there has been little empirical analysis. This is a major omission because the existence or otherwise of systematic life cycle price trends has implications for price index construction and hence, the measurement of inflation and economic growth. If prices vary due to the age of a product, holding other factors constant, then the selection of a representative sample of items across the age spectrum will be vital in accurately representing price change. In this paper we investigate the potential magnitude of this new age-selection bias in price indexes.

Theoretical models of intertemporal price discrimination abound. These modelling efforts have focused on cases where sellers have some sort of monopoly pricing power so that prices deviate from marginal cost. Perhaps the first paper to address the question of intertemporal price discrimination was Stokey (1979) who noted that pocket calculators, the PCs of yesteryear, tended to enter the market at high prices which fell over time. She showed, however, that for a continuum of consumers and time periods, with a monopolist having the same rate of time preference as the consumers and perfect information, that it was not generally profitable to price discriminate. The more recent literature has focused on relaxing some of these assumptions and challenged this conclusion. Landsberger and Meilijson (1985) argued that price will tend to decrease over time if the monopolist has a flatter – less impatient – rate of time preference than consumers. Those with a high willingness to pay will buy the
product immediately at a higher price while those with lower reservation prices will wait until later. Koh (2006) shows that prices can either decrease or increase over time depending on the particular distribution of purchasers’ characteristics and rate of time preference. Imperfect information and liquidity constraints on the consumer-side can also have an effect on the results. Changing marginal costs of production as a product ages provide further motivation for life cycle pricing. For many high-tech goods costs of production are likely to fall leading to a declining price as the product ages. The introduction of competition later in the product cycle, perhaps induced by the entry of firms as they reverse-engineer a product or by patent expiration, provides further impetus for declining prices as a product ages. These interesting theoretical possibilities have not been accompanied by detailed empirical work aimed at identifying the features of life-cycle pricing.

In this paper we aim to resolve the uncertainty surrounding the path of prices over a product’s life. In this regard we build a flexible model of pricing over the life-cycle in section 2. Section 3 applies this methodology to two U.S. scanner data sets, first, for supermarket products; beer, canned soup, and cereals, and second, to the high-tech goods; desktop and laptop computers, and personal digital assistants (PDAs). These two expenditure classes provide an interesting contrast of the potential effects on price indexes of life cycle price trends, a topic which is also addressed in section 3. Section 4 concludes our discussion while some additional results are contained in the Appendix. In the remainder of this section we discuss the effect on price indexes of pricing over the life cycle.

1.1. Price Indexes and the Product Life Cycle

The existence of price trends over the product life cycle has important implications in economic measurement for the construction of price indexes. This has been emphasized in a number of recent reports on consumer price indexes (CPIs) such as the Boskin Commission (Gordon and Griliches, 1997), the Schultze Report (Schultze and Mackie, 2002) as well as the recent ILO Manual on CPIs. There are two main ways

---

1The CPI Manual was drafted by an international group of experts and compiled in consultation
in which life cycle pricing can influence the estimate of price change and hence require careful treatment.

The primary question is one of sampling representativity. If new goods, and goods close to disappearing, have different price trends from other products, then care must be taken in ensuring that a representative sample is taken across products in various stages of the life cycle.

Statistical agencies have often encountered problems in this regard. As a characterization statistical agencies have been slow to introduce new products and focused too much on mid- and late-life products. Armknecht (1997) gives the example of video recorders which were introduced into the U.S. CPI in 1987 when they sold for around $500 but the price change from 1978, when they first appeared on the market at around $1200, was ignored. The stories for cellular telephones (Hausman, 1999), air conditioners and microwave ovens are similar (Gordon and Griliches, 1997). The omission of products early in their life cycle has the potential to introduce bias into the index. In these cases there is a strong supposition that the products discussed had strongly declining prices early in their life hence there exclusion tends to bias the index upward.

While the omission of whole product categories, such as those discussed above, is important it is also likely that omitting new varieties of an existing set of products will matter. If firms, say, charge lower prices for new items in their product range of, say, canned soup, then the failure to incorporate this in the index may also lead to bias. This bias is fundamentally a sampling issue and not related to the conceptual basis of the index. Hence this point is just as relevant for fixed basket indexes as it is for cost-of-living indexes. In the latter case the objective is to hold utility fixed and compare the cost of obtaining a given level under two price regime. In the fixed basket approach a selection of goods is priced through time and substitution is not of concern (see Schultze and Mackie, 2002). In both cases a representative sample of new with the IMF, OECD, UNECE, Eurostat, The World Bank as well as the ILO. The Manual serves as a guide in the construction of price indexes and represents the conventional wisdom accumulated by some of the most experienced statisticians in this area and has a similar status to the System of National Accounts Manual.
and old goods must be taken in order to adequately represent changes in aggregate prices across the desired population of transactions.

Note that potentially the bias introduced by unrepresentative age-related sampling is large. Unlike other sources of index bias, such as that for quality change and new goods, which is confined to an often small subset of product categories and only a few items within those product categories, all goods have an associated age and hence the possibility of age-related price movements.

Another way in which the existence of age-related price effects can be important for price index construction is when goods enter and exit the index sample. Often overlap pricing methods are used which ‘splice’ the new product into the index in place of the old product at the relative prices for which they trade in the market.¹ This form of quality adjustment is used in the U.S. CPI each year where 20% of the sample is replenished as products are rotated into and out of the index (Abraham, 1997). However, this overlap price method will clearly be influenced by the stage of the respective products’ life cycles, which may serve to distort the comparison of relative prices.

The need for such quality adjusted changes in indexes is determined by the rapidity of turnover, both the appearance and disappearance, of products. Also, it seems plausible, at least a priori, that prices will exhibit more dynamic change early on in life, or just as they are about to exit the market, and be more settled during mid-life. For these reasons the share of entering and exiting products in the market will be important. Pakes (2003) notes that around 80% of computers disappear each year for his sample while we find below that around two-thirds of models are removed from the market per annum. Silver and Heravi (2002) noted that exits and entry of products reduced the sample coverage by about 50% over a twelve month period for white wear in the U.K. For coffee Reinsdorf (1999) found that there was around 20% attrition on an annual basis while Melser (2006), for a wide range of supermarket

---

¹If we want to replace item i with j in period 1 then we use the ratio, \( p_{i1}/p_{j1} \) to adjust the numerator in the comparison of prices, \( p_{j1}/p_{i0} \). For a more detailed discussion see for example ILO (2004, p. 106).
products, observed that around 15% of items disappeared over the year.

While the possibility of systematic life cycle effects has been recognized in the measurement literature (see for example, Silver and Heravi, 2002, and the references below) there has again been very little quantification of this phenomenon. One exception is an interesting paper by Berndt, Kyle and Ling (2003) (following earlier work by Berndt, Griliches and Rosett (1993)) who investigate the effect on the price of prescription drugs of patent expiration and the entry of generic producers. They come to the startling conclusion that prices for the established branded varieties tend to rise after patents ran out. This emphasizes that the empirical results in this area can potentially turn theory on its head. In another interesting paper, Haan (2004), a researcher at a statistical agency, was keenly aware of life cycle pricing issues, outlined a hedonic regression model which allowed for systematic effects for entering and exiting varieties. However, this was not empirically estimated.

In the absence of substantial empirical evidence there has been speculation about the likely path for prices as products age. For example, a passage from the ILO Manual argues that,

> It may be that the prices of old items being dropped are relatively low and the prices of new ones relatively high, and such differences in price remain even after quality differences have been taken into account (Silver and Heravi, 2002). For strategic reasons, firms may wish to dump old models, perhaps to make way for the introduction of new models priced relatively high. (ILO, 2005, p.100).

This view is shared by Triplett (2004, Chap. 4, p. 17) as well as Schultze and Mackie (2002, p. 162). While this hypothesis of price-skimming for new goods sounds plausible (and is consistent with much of the theoretical literature) so does the possibility that firms charge a low price for new goods so as to build a market share (Triplett, 2004, Chap. 4, p. 5). One reason for such a strategy, which does not appear to have been fully explored in the literature, is that firms must overcome consumers’ inertia and uncertainty about trying a new product and ‘tempt’ buyers to sample the variety by lowering its price relative to substitutes. The existing empirical literature gives us very few strong priors about the pattern of these pricing strategies. With this in
mind we seek to develop as flexible a model as possible of pricing over the life cycle in the next section.

2. Models of Life-Cycle Pricing

The objective of our estimation is to isolate the effects of age on the price of a product from the myriad of other contributing factors. The approach we adopt is a non-parametric spline smoothing model where we solve an optimization problem penalizing both for deviations from the data and irregularity (i.e. curvature) in the estimated function. This approach is somewhat novel in this context and allows for the flexible estimation of the desired parameters. We construct confidence intervals around the estimated functions by recasting the optimization problem within a maximum likelihood framework. The overall approach, as will be seen, gives an attractive mix of functional form flexibility and structure.

The basic problem that we faced was to construct a model which would control adequately for extraneous factors and enable us to identify the underlying influence of a product’s age on its price. A natural approach is a panel fixed effects model with particular attention to the specification of product-age effects. Suppose that \( p_n \) denotes the price of observation \( n = 1, \ldots, N \). For each time period, \( t = 1, \ldots, T \), there is a cross-section of varieties of each good, \( i = 1, \ldots, I \).\(^3\) Both dimensions can be represented by a set of fixed effects so we define the dummy variables; \( b_{ni} = 1 \) if observation \( n \) is from product category \( i \) and zero otherwise, while \( h_{nt} = 1 \) if the observation is from period \( t \) and zero otherwise. In addition to these factors we have a dependence of price on the age of a good. The fact that out data is finite, and hence we do not observe the date of birth and death of every single product, means we have two bits of age related information; the age from appearance, \( a_n \), for those goods that we have seen arrive on the market, and the age to disappearance, \( d_n \), for the goods

\(^3\)What we mean by a variety \( i \) is a unique bundle of characteristics, e.g. “a 14.25 ounce can of Campbells pumpkin soup.” The use of dummy variables provides a very flexible way of constructing a hedonic function (Diewert, 2003).
which we have observed to disappear in our data.\footnote{That is, if } Controlling for fixed effects using the dummy variables, with log-prices as the dependent variable and normalizing on the first period, gives the model:

$$\ln(p_n) = \sum_{i=1}^I \beta_i b_{ni} + \sum_{t=2}^T \delta_t h_{nt} + f_a(a_n) + f_d(d_n) + e_n, \quad n = 1, \ldots, N \quad (1)$$

The basic estimation problem revolves around how to best represent the effect of new and disappearing goods, $f_a(a_n)$ and $f_d(d_n)$ respectively, on price. A number of options are potentially available.

We could pursue a purely non-parametric approach where a dummy variable is included for each unique value of $a_n$ and $d_n$. This functional form is shown below where we introduce the dummy variables, $u_{anj} = 1 \ (u_{dnj} = 1)$ if $a_n = j \ (d_n = j)$ and zero otherwise, where $j = 1, \ldots, J$ are the values taken by $a_n$ and $d_n$, over the sample, as well as an index of these values.

$$f_a(a_n) = \sum_{j=1}^J \theta_{aj} u_{anj} \quad (2)$$

$$f_d(d_n) = \sum_{j=1}^J \theta_{dj} u_{dnj} \quad (3)$$

Alternatively, a fully parametric approach could be adopted where we hypothesize a functional form for $f_a(a_n)$ and $f_d(d_n)$. For example, we could use a polynomial, say of third order such as is shown below. A motivation for this functional form could be given on the basis that it represents a Taylor series approximation around the point, $a_n = 0$ and $d_n = 0$.

$$f_a(a_n) = \theta_{a0} + \theta_{a1} a_n + \theta_{a2} a_n^2 + \theta_{a3} a_n^3 \quad (4)$$

$$f_d(d_n) = \theta_{d0} + \theta_{d1} d_n + \theta_{d2} d_n^2 + \theta_{d3} d_n^3 \quad (5)$$

While the polynomial approach has some appeal it has the drawback that it places a lot of structure on the problem. Even a third order polynomial places a high de-
gree of restrictiveness upon the global nature of the functional form – forcing the function to have two turning points and a certain degree of smoothness. Perhaps the biggest problem with polynomial regression, from a practical perspective, is that the behaviour of the function at end points can be problematic in that it will tend to be rapidly increasing or decreasing. This poses problems for our analysis as we are particularly interested in the path of the pricing functions at the extremities. In contrast the fully non-parametric dummy variable approach has the advantage of great flexibility. In fact, it may place too little restriction on the model with the results being driven by sampling variability rather than the underlying data generating process. It seems reasonable to impose some continuity restrictions on the function as pricing life-cycle effects are likely to change slowly – the price of a good of age 5 is likely to be more similar, controlling for other factors, to the price of a good of age 4 and 6 than, say, 15. We can use this intuition to place some light-handed continuity restrictions on the dummy variable model.

With this in mind we adopt a spline smoothing approach where the life cycle functions take the form shown in (2) and (3) but we penalize, in a transparent fashion, for rapid changes in their values. This approach gives the best of both worlds in that we have a very flexible function which can provide a robust global approximation, rather than the local approximation of the Taylor polynomial, and is still smooth and hence more easily interpreted. Consider the penalized smoothing problem shown below,

\[
\min_{\beta, \delta, f_a, f_d} \sum_{n=1}^{N} \left[ \ln(p_n) - \sum_{i=1}^{I} \beta_i b_{ni} - \sum_{t=2}^{T} \delta_t h_{nt} - f_a(a_n) - f_d(d_n) \right]^2 + \lambda_1 \int \left| f_a''(x) \right|^2 dx + \lambda_2 \int \left| f_d''(x) \right|^2 dx \tag{6}
\]

The first objective of the optimization is fidelity to the data. In addition, we add a penalty for rapid changes in the curvature of the function. This is reflected

---

5The most well known use of splines in Economics is the Hodrick-Prescott filter (Hodrick and Prescott, 1980) though their use dates back to Whitaker (1923). Spline methods are becoming increasingly popular in the applied literature, particularly for hedonic regression and spatial models, see for example Bao and Wan (2004).
in the integral over the squared second derivative of the functions. The smoothing parameters, \( \lambda_1 \) and \( \lambda_2 \), represent the relative weights that we give to fidelity and smoothness. As \( \lambda_1, \lambda_2 \to \infty \) the selected functions will have no second-order curvature which implies that the estimators are linear (i.e. \( f_a(a_n) = \theta_{a_0} + \theta_{a_1}a_n \), and \( f_d(d_n) = \theta_{d_0} + \theta_{d_1}d_n \)). Note also that the spline smoothing model has the desirable property of nesting the non-parametric dummy variable approach. When \( \lambda_1, \lambda_2 \to 0 \) the functions will collapse to non-smooth dummy variable functions, as in (2) and (3), and potentially be very ‘jagged’. We discuss a method for choosing \( \lambda_1 \) and \( \lambda_2 \) below.

Green and Silverman, (2000) and Wahba (1990) show that the problem (6) has a unique solution and that efficient algorithms exist to solve the resulting system of equations. The minimizer is a natural cubic spline (Green and Silverman, 2000, pp. 13, 66) and, while a closed form solution exists, the size of the data sets used in our application necessitates an iterative procedure using a set of simultaneous equations.\(^6\)

The choice of the smoothing parameters \( \lambda_1 \) and \( \lambda_2 \) is somewhat arbitrary but also of some importance in determining the solution. A way around this subjectiveness is to use the method of Cross Validation (CV). This approach minimizes the ‘forecast error’ of the model in the sense that we withhold one observation from the model and compare actual and estimated values. That is, we minimize \( CV(\lambda_1, \lambda_2) \) which is shown in (8). For notational brevity we use \( y_n = \ln(p_n) \), and \( \hat{y}_{-n} \) to denote the estimated value of observation \( n \) when this observation is excluded from the data used to estimate parameters, \( \hat{y}_n \) is the estimated value from the full model. The alternative way of writing \( CV \) uses the matrix \( A \) such that, \( \hat{y} = Ay \).

\[
CV(\lambda_1, \lambda_2) = \sum_{n=1}^{N} [y_n - \hat{y}_{-n}(\lambda_1, \lambda_2)]^2 = \sum_{n=1}^{N} \left[ \frac{y_n - \hat{y}_n(\lambda_1, \lambda_2)}{1 - A_{nn}} \right]^2 \quad (7)
\]

The second expression for calculating \( CV \) is clearly less computationally intensive making it relatively straightforward to apply in practice. However, one problem with \( CV \), noted by Craven and Wahba (1979), is that it tends to give too much influence to outliers. They suggested a Generalized Cross Validation (GCV) score function which ascribes lesser weight to these high-influence observations.\(^6\)

\(^6\)The details of the problem and its solution are given in the Appendix for interested readers.
This robustness to outliers is an important advantage and hence we use the GCV approach to derive \( \hat{\lambda}_1 \) and \( \hat{\lambda}_2 \).

We use a somewhat novel approach to calculate standard errors for the semi-parametric model by recasting the model within a maximum likelihood framework. This approach is relatively new in the literature, and is due to Reeves et al. (2000). The method gives a familiar expression for the variance with the details left to the Appendix. Other approaches taken to the calculation of standard errors in the spline case have been to recast the smoothing model within a Bayesian framework, which involves adopting this conceptual framework, or estimating bootstrap confidence intervals by constructing sample replicates (see Wahba (1990) for discussion of both these approaches). While the latter approach is attractive it will be computational prohibitive for the large data sets used in the next section.

With this model of life cycle pricing we move to apply it to two large U.S. scanner data sets. Additionally, we undertake an investigation of the effects of life-cycle pricing on price indexes.

### 3. An Empirical Investigation

We apply the non-parametric smoothing methods discussed above to two data sets. The first is a supermarket scanner (or barcode) data set for the Dominick’s chain of food stores in the Chicago area.\(^7\) There are 96 stores included in the database with price data available at a monthly frequency from September 1989 to May 1997 – a period of almost eight years (though not all the products are available for all this period). We focus on prices for; beer, canned soup, and cereals.\(^8\)

---

\(^7\)The data is made publicly available, free of charge, by the James M. Kilts Center, Graduate School of Business, University of Chicago. The authors gratefully acknowledge the Center for making the data accessible in this way.

\(^8\)One point to note regarding the use of Dominick’s data is that it represents the sales of just a single chain of stores. Any pricing strategies that we find may in this case be characteristic to this retailer and atypical of pricing patterns more generally.
The other data set is also scanner data that was purchased by the BEA for research purposes and made available to the authors. This data set is particularly interesting because it includes the high tech goods; desktop and laptop computers as well as personal digital assistants (PDAs). These electronic goods provide an interesting contrast with the supermarket products also analysed and correspond most closely to the types of products for which the literature has modelled life cycle price discrimination. The high tech products are also of considerable interest because of the difficulty inherent in calculating price indexes for these rapidly changing and dynamic expenditure classes (Gordon and Griliches, 1997; ILO, 2004; Triplett, 2004). This data runs for a period of 3 years, from October 2001 to September 2004.

We provide some summary information in Table 1 on the data sets focusing particularly on aspects of the product life cycle. It can be seen that the high-tech goods have a particularly short life span. For these products which were observed for at least 3 consecutive months the average length of life was just 7 months for desktops, 6 months for laptops and a meagre 4 months for PDAs. The short life span reflects a highly skewed distribution of product lives. While some models last for years, most varieties appear and disappear in very short order, pulling the averages down. Another way to demonstrate the short life span of these high-tech products is to look at the probability that a particular variety will be available over a certain time range. For desktop computers, for example, the probability that a model will disappear next month is 28.13%, in 6 months time, 50.94%, and in a year 67.33%. Of all varieties recorded in our data, 2427 in all, 711 (31.8%) were new during the period while 426 (17.6%) disappeared, further indicating significant turnover. The trends are similar for laptops and PDAs.

There is a strong contrast between the high tech goods and the supermarket products. The latter have a longer life span, around 20 months for all those products which lived at least 6 months, and a significantly less dramatic disappearance rate. Though, with a disappearance rate of around 15% for canned soup and cereals and 25% for beer, over 12 months, it is still far from trivial.

---

9 We are particularly grateful to the BEA for allowing the use of their data in this project.
Table 1: Product Life Cycle Statistics

<table>
<thead>
<tr>
<th>Product</th>
<th>Probability that a Model Disappears from the Market in:</th>
<th>Number of Products</th>
<th>Average Length of Product Life (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 month</td>
<td>6 months</td>
<td>12 months</td>
</tr>
<tr>
<td>Desktops</td>
<td>28.13%</td>
<td>50.94%</td>
<td>67.33%</td>
</tr>
<tr>
<td>Laptops</td>
<td>26.34%</td>
<td>50.24%</td>
<td>66.15%</td>
</tr>
<tr>
<td>PDAs</td>
<td>9.68%</td>
<td>17.70%</td>
<td>28.23%</td>
</tr>
<tr>
<td>Beer</td>
<td>5.17%</td>
<td>14.98%</td>
<td>25.38%</td>
</tr>
<tr>
<td>Canned Soup</td>
<td>2.16%</td>
<td>8.42%</td>
<td>15.56%</td>
</tr>
<tr>
<td>Cereal</td>
<td>2.79%</td>
<td>9.86%</td>
<td>15.56%</td>
</tr>
</tbody>
</table>

(a) For all products for which we observed both the start and end of the life cycle and for which the life cycle was longer than 6 months for the supermarket products and 3 months for the high-tech goods.

As emphasized earlier the turnover rates are potentially important as it influences the amount of quality adjustment which statistical agencies are required to undertake in their matched samples. It also reflects the extent to which products are likely to be in dynamic price transition as more ‘action’ in prices is likely to take place either early or late in life. The statistics also indicate the amount of information that we have to work with in estimating the life cycle pricing function. For laptops, for example, we have around 813 varieties appearing over the sample and 489 disappearing, out of a total of 2,451. This gives a good amount of information by which to robustly estimate \( f_a(a) \) and \( f_d(d) \).

3.1. Some Results

Using the data, and smoothing spline model outlined above, we estimated the effect of age, from appearance and to disappearance, on the prices for each of our product categories. Turning first to the model diagnostics. In Table 2 we report the value of the smoothing parameters chosen using GCV as well as the sample size and length, and pseudo-\( R^2 \).

The model generally performed well. The use of a smoothing spline serves to better elucidate the relationship between price and age than the dummy variables, which
are quite volatile especially at the right-hand-side where there are fewer observations. The \textit{GCV} method leads to a range of values for the smoothing parameters. More smoothing occurs for the high-tech goods, where the price effects are larger, than for the supermarket products where $\lambda_1$ and $\lambda_2$ are small though never equal to zero – as they would be for the pure dummy variable model. For desktops $\lambda_2$ reached our upper bound, where the function became effectively linear, indicating that significant smoothing was required in this case. The Pseudo-$R^2$ values, the squared correlation coefficient of actual and predicted values, show how well the models explain the data. While the regression model does not have a strict statistical interpretation we can undertake approximate F-tests of parameter restrictions by looking at changes in the sum of squared errors. We find that in all cases there are statistically insignificant differences between the spline model and the unsmoothed age-dummy model. However, for canned soup and cereals, there is a statistically significant difference between a regression model without any age effects and the unsmoothed model. This indicates to us that the spline smoothing model does almost as good a job of explaining data \textit{within sample} as the dummy variable age specification.\textsuperscript{10} The removal of age effects reduces the explanatory power of the model in all cases though the large amount of variation explained by the other parameters in the model mean these changes are mostly insignificant. As we will see, however, the economic magnitudes of the parameters are often quite large.

The results from the model estimation are presented graphically in Figures 1–12. The Figures show the smoothing spline, the solid dark line, as well as the unsmoothed dummy variable estimates for the life cycle, the white line. The dashed lines represent our maximum likelihood confidence intervals, conditional on the value of smoothing parameter. The Figures record the influence of age, from appearance or to disappearance, on price over a 24-month period.

\{Insert Figures About Here\}

\textsuperscript{10}Note that our approach to choosing parameters, and the values of $\lambda_1$ and $\lambda_2$, was to focus on out of sample performance as in \textit{GCV}. In this regard we judged the spline model as superior to the unsmoothed model.
Table 2: Model Statistics

<table>
<thead>
<tr>
<th>Product</th>
<th>Number of Months</th>
<th>$\lambda_1$ (’000s)</th>
<th>$\lambda_2$ (’000s)</th>
<th>Number of Observations</th>
<th>Pseudo-$R^2$(^{(a)}) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Unsmoothed Model</td>
</tr>
<tr>
<td>Desktops</td>
<td>36</td>
<td>2.2</td>
<td>100.0(^{(b)})</td>
<td>21,505</td>
<td>88.9327</td>
</tr>
<tr>
<td>Laptops</td>
<td>36</td>
<td>2.7</td>
<td>6.8</td>
<td>19,937</td>
<td>88.9042</td>
</tr>
<tr>
<td>PDAs</td>
<td>36</td>
<td>1.9</td>
<td>0.3</td>
<td>4,274</td>
<td>89.6395</td>
</tr>
<tr>
<td>Beer</td>
<td>72</td>
<td>0.1</td>
<td>0.1</td>
<td>21,338</td>
<td>98.4176</td>
</tr>
<tr>
<td>Canned Soup</td>
<td>93</td>
<td>0.1</td>
<td>0.1</td>
<td>21,665</td>
<td>96.1294</td>
</tr>
<tr>
<td>Cereals</td>
<td>88</td>
<td>0.1</td>
<td>1.2</td>
<td>18,088</td>
<td>95.5206</td>
</tr>
</tbody>
</table>

\(^{(a)}\) The $R^2$ of a regression of estimated on actual log prices (i.e. the squared correlation coefficient)

\(^{(b)}\) An upper bound of 100,000 was placed upon this value which makes the function effectively linear.

\(^{(c)}\) Statistically significantly different from the unsmoothed model at the 99% level using a Pseudo F-test.

The pattern of life cycle pricing, as shown in the Figures, is interesting for a number of reasons. First, the pricing effects that we find are generally quite large. For laptops, for example, the price of a new product peaks at month 4 and then declines so that in month 12 the price is 10.04% lower. This is a significant price change due entirely to life cycle factors. This order of magnitude is fairly typical of the high-tech goods which show strong life cycle effects. However, the supermarket products exhibit considerably less pronounced swings in price due to age. New cereals, for example, decrease in price by 4.53% in the first 12 months of their life while canned soup falls by around 5.11%. New varieties of beer on the other hand appear to be relatively stable over the first year of life. One of the immediate implications of the existence of these life-cycle pricing effects is that hedonic functions which exclude ‘age’ as an explanatory variable may potentially be mis-specified.

Second, some stylized facts of life-cycle pricing emerge arise. For desktops and laptops prices generally tend to decline as products age. However, there appears to be an initial lip at around 4–6 months over which prices are flat or even weakly rising. While the same trends are apparent for PDAs the lip is more pronounced, products
come in cheap and rise in price up to the 9th month and then begin declining. This pricing-lip may reflect introductory sales used to generate customer trials of the new variety. Interestingly, there does not appear to be any strong trend amongst the high-tech products for prices to decline as a variety reaches the end of its life. While this does occur for desktops, in contrast, for laptops, and to a lesser extent for PDAs, prices rise as death approaches. For the supermarket products prices appear to decline upon entry though beer is flat. Upon exit prices tend to fall and often quite markedly for canned soup and cereal. Beer varieties, on the other hand, actually rise in price as they disappear from the market. This may reflect the strong part taste and brand loyalty play in this market so that something like the Berndt, Kyle and Ling (2003) result arises with a smaller number of consumers willing to pay a premium for their preferred brand.

Third, it does appear that the beginning and end of the life cycle do generally represent pricing extremes. All the high-tech products are at or near life-cycle pricing peaks or troughs except PDAs which are disappearing. Similarly for the supermarket products, in all cases except new-beer varieties, are either at the top or bottom of the pricing life-cycle. This confirms our a priori supposition that much of the dynamism in product pricing is seen early and/or late in the cycle.

Fourth, the approximate confidence intervals for the estimated life cycle function are generally fairly wide for the high tech products to reject the individual hypotheses that the life cycle effects are constant through time, i.e. the life cycle function is in fact flat. Only on the case of canned soup and cereals could we reject the joint hypothesis of no life cycle effects over the sample. This means we must be cautious in interpreting the results. In the case of the supermarket products the confidence intervals are generally tighter meaning we can reject the hypothesis of no life cycle effects in around half the cases. Unfortunately, while some of the estimated effects may not register as statistically significant, they are certainly of an economically significant magnitude. In the next section we show that life-cycle pricing effects of this magnitude can have a large influence on estimates of price change.
3.2. Implications for Price Indexes

The results above have some important implications for the construction of price indexes. If the stage of the life cycle is an important price-determining characteristic then this must be taken into account in constructing an index. More specifically, the sample of products included in the index must be representative of the population of prices in the market so that the index’s life-cycle price effects are the same as that of the target population.

Let us use the model of price outlined above and consider the implications for a simple price index between periods $0$ and $1$. To measure price change we compare the imputed prices of some basket of goods in the two periods – an approach which is called the imputations method in the hedonic regression literature (Triplett, 2004; Hill and Melser, 2006a, 2006b). We use a geometric mean aggregator over $M$ products with equal weights for each observation.$^{11}$

$$\hat{P}_{01} = \prod_{m=1}^{M} \left[ \frac{\hat{p}(b_m, h_{m1}, a_m + 1, d_m - 1)}{\hat{p}(b_m, h_{m0}, a_m, d_m)} \right]^{1/M}$$ (9)

The important feature of the index is that, while the characteristics $b_m$ are held fixed between periods, the age of each of the varieties must increase. That is, between periods $0$ and $1$ the products get one period further from introduction and one period closer to disappearing from the market. Using our logarithmic model of price outlined and estimated above in equations (1), (2), (3), and (6) we can write the index in the following way.$^{12}$

$$\hat{P}_{01} = \prod_{m=1}^{M} \left[ \frac{\exp(\sum_{i=1}^{I} \hat{\beta}_i b_m + \delta h_{m1} + \sum_{j=1}^{J} \hat{\theta}_{aj} u_{amj+1} + \sum_{j=1}^{J} \hat{\theta}_{dj} u_{dmj-1})}{\exp(\sum_{i=1}^{I} \hat{\beta}_i b_m + \delta_0 h_{m0} + \sum_{j=1}^{J} \hat{\theta}_{aj} u_{amj} + \sum_{j=1}^{J} \hat{\theta}_{dj} u_{dm})} \right]^{1/M}$$ (10)

$^{11}$We use the geometric mean because it has particular advantages when used in conjunction with the logarithmic price function adopted above (see Hill and Melser, 2006a), but its use, as well as the absence of weighting, does not change the substance of the point made.

$^{12}$Note that if we want an unbiased estimator then we should make some adjustment to this index due to the fact that we are taking the exponent of estimated parameters (see Garderen and Shah, 2002; Hill and Melser, 2006b). We will not illustrate for the case of these adjustments because it complicates the problem without diminishing our essential point.
It can readily be seen that the index, while immune from compositional change due to matching, will be influenced by the aging of products. In (13) the age distribution in the initial period clearly determines the age of products in the second period under a matched sample framework. The specific nature of the original age distribution will influence price change – an effect which is mediated by the size of the life cycle pricing parameters.

To consider the potential size of the sampling error we run a series of simulations using our estimated price function and data. We suppose that a sample is taken in some initial month and that this is kept constant over the remainder of the year. In this case note that while the initial sample may contain some varieties which are brand new to the market (i.e. have an age of 1) the following month can only contain goods which are at least 2 periods old, and so forth. That is, the sample must age and hence move along the price function. We examine how this index is influenced by aging over a 12-month period. Unfortunately, it is not possible to be definitive about the bias because we cannot calculate the ideal index. The problem is that our data sets are finite so we do not observe the birth and death of all products hence we do not know the actual (i.e. population) distribution of product ages in our data. Ideally an index would use the actual distribution of ages amongst products and record price change for these products. Given our inability to do this we instead undertake a number of simulations looking at various distributions for the ages of products in the base period (i.e. how products are scattered along the age spectrum) and then
compute the price change for this selection. The scenarios considered are:

(A) The empirical distribution of products for which we observed the entire life cycle. The features of this distribution are a high degree of skewness towards varieties which are new to the market and paradoxically also close to disappearing which reflects the relatively short life-span of products, at least for the high-tech products. (When a product disappears from the market a chained matched-sample approach is used.)

(B) A uniform distribution of products is sampled for a 12-month period over the ages, 1 month to 12 months for new goods. We further supposed that these products were uniformly distributed as being between 12 to 23 months (inclusive) to disappearing from the market. (These assumptions mean that no products disappear from the sample over the 12 months for which the index is calculated.)

(C) A uniform distribution of products is sampled over a 6 month period from the ages of 7 to 12 months for new goods and we supposed that these products had between 18 to 23 months (inclusive) to go in their life until they disappeared. (These assumptions mean that no products disappear from the sample over the 12 months for which the index is calculated.)

Scenario (A) gives our best estimate of price change from our data, given the caveat of limited knowledge expressed above. Scenarios (B) and (C) contrast in that the former includes the effects from goods which are very early and late in their product cycle while scenario (C) excludes all effects from products which are of age 6 months or less and within 6 months of exiting the market.

We present a summary of our results in Table 3. The life cycle effects are decomposed into those due to New Goods and Disappearing Goods, as shown in equation (13). These effects along with pure price change, give the Total price change over the 12-month comparison.

The results show that the effects of product age can be large in that the spreads between the price change under various sampling methods are big. In the case of
Table 3: Life Cycle Effects on Price Indexes\(^{(a)}\)

<table>
<thead>
<tr>
<th>Product</th>
<th>(A)</th>
<th>(B)</th>
<th>(C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desktops</td>
<td>-6.57</td>
<td>-2.12</td>
<td>-33.98</td>
</tr>
<tr>
<td>Laptops</td>
<td>-11.38</td>
<td>0.75</td>
<td>-46.17</td>
</tr>
<tr>
<td>PDAs</td>
<td>-1.80</td>
<td>-10.44</td>
<td>-30.46</td>
</tr>
<tr>
<td>Beer</td>
<td>0.86</td>
<td>-0.60</td>
<td>2.05</td>
</tr>
<tr>
<td>Canned Soup</td>
<td>-3.52</td>
<td>1.90</td>
<td>4.04</td>
</tr>
<tr>
<td>Cereal</td>
<td>-1.75</td>
<td>2.43</td>
<td>3.04</td>
</tr>
</tbody>
</table>

\(^{(a)}\) All numbers represent annual average percentage changes.

desktop computers, for example, under scenario (A), which reflects the empirical distribution of ages, we get an annual average price fall of 33.98%. The price falls under scenarios (B) and (C) are even larger. The reason for this is that the latter methods do not have such as strong weight on items which are relatively young, in the case of scenario (C) items younger than 6 months are not sampled at all. As the price function for appearing products slopes down from around 4 months onward, having greater weight on the lower (righthand) end of the price function leads to larger price falls. The spread between scenarios for laptops is –46.19% for (A) and –48.49% for (C) a small fraction of overall price change. PDAs, the most recent and dynamic of the high-tech products, show a very large difference of –30.46% for (A) compared with just –18.30% for (C). Generally, because the appearing products life cycling tends to slope down, the aging of products leads to a fall in the price of appearing goods. This is reflected in the fact that for all the high-tech products the contribution to price change from the aging of appearing products to the index is always negative.

The effects for the supermarket products are less dramatic. For beer, for example, we saw that the life cycle effects were relatively small for appearing products – i.e. the price function was essentially flat. Consequently the effect of different sampling patterns for new products was small. However, the effect of age to disappearance
for beer was larger and led to differences between total price change of over 2% between methods (A), and (B) and (C). Canned soup and cereal also show differences between methods of around 2%. While these differences are of a significantly smaller magnitude than for the high-tech goods, it makes up a large proportion of overall price change, around half. This indicates that even for fairly mundane product categories life-cycling pricing patterns are important for determining measured price change.

More generally, we have shown with our simulations that the effects of life cycle pricing are large enough so that different selection strategies will lead to different index numbers. While we can not be sure of the true price change in our data sets, because of its limitations, the existence of a big spread between index numbers under various methods emphasizes our point.

4. Conclusion

The purpose of this paper has been two-fold. First, to shed some light on the path of prices for commonly traded supermarket products and high-tech goods, over their life cycle. This is in the context of an ongoing modelling effort looking at intertemporal price discrimination. Second, we investigated the implications of these duration effects for the estimation of price indexes.

The results of our model, while not linked in a structural manner to the theoretical literature, do provide some support for their hypotheses. We generally found, most strongly for the high-tech goods, that prices decline as products age. Given the absence of more detailed information we were not able to identify the extent to which this was due to the various possible factors; declining costs, increased competition, customer segmentation or other reasons. One interesting feature of the results is that in the initial months of a good’s introduction they often appear to be discounted. This ‘pricing lip’ is then followed by a more protracted and significant decline in prices. Further theoretical work could focus on motivations for this apparent pricing strategy while empirical work could try to flesh out more of the details of the systematic patterns in life cycle pricing.

The key empirical finding was that the age of a product, either from its birth or to
its death, does contribute to its price level and that this has significant implications
for price indexes. In a simple simulation, where we held the sample of products fixed
for a year, we showed that bias could arise because of the effects of life cycle pricing.
This must caution statistical agencies to take steps to ensure that their samples of
products along the age spectrum are representative of the actual distribution. Failure
to accurately reflect product age in a sample of goods is likely to lead to a sizeable
error in the index.

5. Appendix

5.1. Estimation and GCV

Writing the model in matrix notation we define the log of prices by the vector, $y$, $B$ is
the matrix of product dummy variables with coefficients $b$, and $H$ is a matrix of
time dummy variables with coefficients $h$. We also use the incidence matrices $U_1$ and
$U_2$, for the smoothing functions $f_a(.)$ and $f_d(.)$ respectively, which record the values
of $u_{anj}$ and $u_{dnj}$, $j = 1, ..., J$, from equations (2) and (3). The corresponding vectors
of parameters are $g_1$ and $g_2$. The vector of errors is denoted $e$.

$$ y = Bb + Hh + U_1g_1 + U_2g_2 + e \quad (14) $$

Green and Silverman (2000) show that, because of the particular nature of the
penalty function, we can write the penalised sum of squares as in (6). The matrices
$K_1$ and $K_2$ are defined in Green and Silverman (2000).

$$(y - Bb - Hh - U_1g_1 - U_2g_2)^T(y - Bb - Hh - U_1g_1 - U_2g_2)$$
$$+ \lambda_1g_1^TK_1g_1 + \lambda_2g_2^TK_2g_2 \quad (15)$$

Given this expression for the penalised sum of squares in (15) we can write the first
order conditions as a set of simultaneous equations. These are shown below.

$$\begin{pmatrix}
B^T(y - U_1g_1 - U_2g_2) \\
H^T(y - U_1g_1 - U_2g_2)
\end{pmatrix} =
\begin{pmatrix}
B^TB & B^TH \\
H^TB & H^TH
\end{pmatrix}
\begin{pmatrix}
b \\
h
\end{pmatrix} \quad (16)$$

22
\[
\begin{pmatrix}
(U_1^T(y - Bb - Hh)) \\
(U_2^T(y - Bb - Hh))
\end{pmatrix}
= \begin{pmatrix}
(U_1^TU_1 + \lambda_1K_1 & U_1^TU_2 \\
U_2^TU_1 & (U_2^TU_2 + \lambda_2K_2)
\end{pmatrix}
\begin{pmatrix}
g_1 \\
g_2
\end{pmatrix}
\] (17)

These equations can be solved by using the backfitting technique where we start with some initial estimate, say \(b^0\) and \(h^0\), and use this to determine \(g_1^0\) and \(g_2^0\), and proceed iteratively until convergence is achieved. Green and Silverman (2000, p. 68) show that such a procedure will converge, and in practice it is likely to converge fairly quickly.

In order to use GCV for the selection of the smoothing parameters we are required to know \(A\) such that, \(\hat{y} = Ay\). Using equations (16) and (17) we can write \(A\) as,

\[
\begin{pmatrix}
B^TB & B^TH & B^TU_1 & B^TU_2 \\
H^TB & H^TH & H^TU_1 & H^TU_2 \\
U_1^TB & U_1^TH & U_1^TU_1 + \lambda_1K_1 & U_1^TU_2 \\
U_2^TB & U_2^TH & U_2^TU_1 & U_2^TU_2 + \lambda_2K_2
\end{pmatrix}^{-1}
\begin{pmatrix}
B^T \\
H^T \\
U_1^T \\
U_2^T
\end{pmatrix}
\] (18)

We require the trace of this large matrix but can simplify the problem by using the rule, \(\text{tr}(ZX^{-1}Z^T) = \text{tr}(X^{-1}Z^TZ)\).

5.2. Estimation of Standard Errors

In order to estimate the variance of the coefficients we re-write the model in equations (2), (3), and (6) in Maximum Likelihood form. With the assumption of Normality the log-likelihood, up to a constant, has the form,

\[
\ln L = \frac{1}{2\sigma^2} \sum_{n=1}^{N} \left[ \ln(p_n) - \sum_{i=1}^{J} \beta_ib_{ni} - T \sum_{t=2}^{T} \delta_tb_{nt} - \sum_{j=1}^{J} \theta_{aj}u_{aj} - \sum_{j=1}^{J} \theta_{dj}u_{dj} \right]^2 \\
+ \frac{\lambda_1}{2\sigma^2} \sum_{j=2}^{J-1} \left[ (\theta_{aj+1} - \theta_{aj}) - (\theta_{aj} - \theta_{aj-1}) \right]^2 \\
+ \frac{\lambda_2}{2\sigma^2} \sum_{j=2}^{J-1} \left[ (\theta_{aj} - \theta_{aj}) - (\theta_{aj} - \theta_{aj-1}) \right]^2
\] (19)
Where $\sigma^2$ is the variance of the error, $\lambda_1 = \sigma^2/\sigma_1^2$ and $\lambda_2 = \sigma^2/\sigma_2^2$ are given. To estimate the covariance matrix of the parameters, $\hat{V}$, we take the inverse of the outer product of first derivatives of the log-likelihood function (Greene, 2000). Inserting the estimator for $\sigma^2$, the average of squared residuals, gives,

$$
\hat{V} = \hat{\sigma}^2 (L^T L)^{-1}
$$

L = \begin{pmatrix}
B & H & U_1 & U_2 \\
0 & 0 & \lambda_1 Q_1^T & 0 \\
0 & 0 & 0 & \lambda_2 Q_2^T
\end{pmatrix}

(20)

References


Craven, P. and G. Wahba (1979), Smoothing Noisy Data with Spline Functions: estimating the Correct Degree of Smoothing by the Method of Generalized Cross-Validation, Numerische Mathematik 31, 377-403.


