The Introduction of Hedonic Regression Techniques for the quality adjustment of computing equipment in the Producer Prices Index (PPI) and Harmonised Index of Consumer Prices (HICP)

Adrian Ball, Consumer Prices and General Inflation Division, ONS
E-mail: adrian.ball@ons.gov.uk

Andrew Allen, Business Prices and Sales Division, ONS
E-mail: andrew.allen@ons.gov.uk

National Statistics customer contact centre: +44 (0)845 601 3034
e-mail: info@statistics.gov.uk

Introduction

One of the most challenging issues facing the compilers of price indices is the separate identification and valuation of changes in the quality of goods and services. This needs to be done so that quality changes can be excluded from the index to ensure that the movement in the latter reflects only the underlying price change.

This challenge is most pronounced in high technology goods, particularly computing equipment, where large quality improvements frequently take place as models are upgraded, replaced by manufacturers or cease to be available in shops and are replaced with ones with different specifications. In these circumstances the precise method of quality adjustment is particularly important in ensuring that the true underlying price change is measured. In consequence the focus of ONS’s development work has been in the production of quality adjusted price indices for these goods.

This article explains, by way of background, the producer cost and option cost methods of quality adjustment currently employed for computing equipment in the Producer Prices Index (PPI), the UK Harmonised Index of Consumer Prices (HICP) and the RPI. It then goes on to describe the method of quality adjustment using hedonic regression to be used in the PPI and HICP to adjust for quality changes in computing equipment indices from the February 2003 index, published in March. For the PPI this move to hedonic techniques covers both desktop personal computers (PCs) and notebook computers. For the HICP it covers just desktop PCs, but in principle will include notebook computers once these have been shown to have sufficient weight to enter the index. Current methods of quality adjustment will continue to be employed for other goods and services.

It should be noted that at present there is no plan for the introduction of hedonics into the RPI. The option cost method will continue to be used. We consider it to be appropriate to be more reserved about introducing methodological changes into the RPI given its widespread use in many important contexts and the fact that uniquely amongst National Statistics it cannot be revised. Because we do not revise the RPI, we will use our experiences with the PPI and HICP to assess whether that may prevent us from using hedonic measures in that index. The position will be kept under review.

1 The plans for the Producer prices index also apply to the Export and Import prices indices which are part of the PPI family.
**Current Methods of Quality Adjustment**

Quality adjustment is the process by which prices are adjusted to account for changes in the quality of a good or service. For example, consider the case where the price of a PC with 256Mb of memory is collected in January, but because this machine is not available in February, the price is collected of a replacement machine with 512Mb of memory. Under this scenario a direct comparison of prices will be inappropriate because it will include the additional cost of the quality improvement, i.e., the extra memory. In order to derive a price change that solely represents inflation, it is necessary to standardise the prices to a common specification for memory. Up to January 2003 this was done using either producer or option cost in the PPI and option cost in the HICP.

*Production cost [currently used in the PPI]*

Until recently the generally preferred method of quality adjustment for any manufactured good in the PPI is to adjust the base month price to reflect the additional cost of the change in specification, from costs supplied by the manufacturer. The use of production costs ensures that allowance is made for the cost of specification changes even where these may not be apparent or of direct relevance to the consumer. For example, the cost of better circuitry in a computer to make full use of change in the size of memory would not appear in the retail specification but is relevant for the costs of the producer.

When a manufacturer gives an assessment of the cost of adding the improvement, the full cost is used to adjust the base month price.

In practice it can be difficult to obtain fully reliable production costs from the manufacturer, and where the manufacturer is unable to supply the production cost of the specification change a 50 per cent option cost methodology is employed as a fallback using the estimated retail cost. This is discussed in detail in the next section. In this case only 50 per cent of the cost of the option is applied to account for economies of scale in production and variable margins.

*Option Cost [currently used in the HICP and RPI]*

In option costing the retail cost associated with a change in specification is obtained from the cost of purchasing the change separately or as an added option. Fifty per cent of this is added to the price of the original model to give a price comparison that is independent of any changes in quality.

Fifty per cent of the cost is applied for a number of reasons. In part this is because 100 per cent option cost could lead to an over adjustment for quality change, due to the fact that the cost of buying features separately is generally greater than buying them as a package. In addition it was considered prudent, when personal computers were introduced into the HICP in 1996, to take a proportion of the option cost given that it was the first time that this method had been used in a consumer price index.
The following example shows how option cost is currently applied in practice.

In January a PC is selected costing £1,000 with 256mb of memory. In March a PC with similar specifications is unavailable so a new model with 512Mb of memory is selected, and costs £1,030. A direct comparison of prices in the index is inappropriate because of the change in specification. Therefore the base price needs to be adjusted as follows:

\[\text{Adjusted January price} = \frac{\text{January price} \times \text{March price}}{\text{March price} - 50\% \text{ option cost}}\]

Research indicates that 256Mb of memory (the difference in the specifications) costs £70. Using the above formula the adjusted base price is calculated as follows:

\[\frac{1000 \times 1030}{1030 - 1/2 \times 70} = 1035\]

Thus the quality-adjusted version of this index is:

\[100 \times \frac{1030}{1035} = 99.5\]

compared with the unadjusted index:

\[100 \times \frac{1030}{1000} = 103.0\]

It can be observed that the application of option cost necessarily involves an element of judgement. This includes the determination of the appropriate proportion of the option cost that should be taken, the identification of the relevant subset of options to price from the complete set of options available to producers and consumers and unravelling their individual basic cost.

**Hedonic Regression methods**

The application of hedonic methods dates back to the first half of the last Century, when it was applied in the USA to calculate rental values, taking into account the number of rooms and other amenities. Since then the technique has been developed and, as can be seen from Table 1, is now used in several countries to calculate the cost associated with quality change in price indices. The name has its origins in the presumption that hedonic indices measure consumer utility, ie the value and pleasure to the user.
Table 1 Use of hedonics in price indices: an international comparison

<table>
<thead>
<tr>
<th>Country</th>
<th>Hedonics used now</th>
<th>Hedonics planned to be used</th>
<th>Under development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td></td>
<td></td>
<td>PCs</td>
</tr>
<tr>
<td>Canada</td>
<td>PCs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>Used cars, owner-occupied dwellings</td>
<td>Consumer durables, mobile telephones</td>
<td>PC software, Clothing</td>
</tr>
<tr>
<td>France</td>
<td>Dishwashers, TVs, Books, Men’s long sleeved shirts, PCs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>PCs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Japan</td>
<td></td>
<td></td>
<td>Rail fares</td>
</tr>
<tr>
<td>Korea</td>
<td></td>
<td></td>
<td>PCs (and other electrical goods)</td>
</tr>
<tr>
<td>Netherlands</td>
<td></td>
<td></td>
<td>New cars</td>
</tr>
<tr>
<td>New Zealand</td>
<td></td>
<td></td>
<td>Fridge-freezers</td>
</tr>
<tr>
<td>Sweden</td>
<td></td>
<td></td>
<td>New cars</td>
</tr>
<tr>
<td>USA</td>
<td>PCs, TVs, DVDs, VCRs, Audio Systems, Microwave Ovens, Clothing</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The estimation method used in hedonic regressions is a set of ordinary least square regressions that relate the price of an item to its measurable characteristics. For computers, these characteristics may include the speed of the processor, the size of the hard disk drive and the amount of memory in a computer. In the ONS regressions are calculated separately for desktop and notebook computers, and the results applied to the PPI and HICP.

The method has two main modes of application: the indirect method and the time dummy approach.

The time dummy method pools all available data on a good over several time periods, fits a regression model and uses the coefficients on the variables relating to time as a direct measure of the price index. This is illustrated in equation 1.

Equation 1: \( \text{Price} = \alpha + \Sigma \gamma \beta_1 + \delta_1 t_1 + \delta_2 t_2 + \epsilon \)

In this the \( \beta \) coefficients represent the cost of the features of the good (in the case of PCs this would include processor chip, memory and hard disc space). The \( \delta \) coefficients show the price index for time periods 1 and 2, and these would be the figures put into any published index.

This time-dummy approach is favoured by many academics because it has some good statistical properties when employed on a single dataset. But it is not preferred by National Statistical Institutes (NSIs) because in practice indices constructed by this method tend to lack stability when employed over several datasets. In addition the use of this approach raises a potential revisions issue – when results are estimated for a new month, new estimates are available for previous months.
Indirect method

This is the approach preferred by NSIs and the one being adopted by ONS. The particular application of the indirect method that is being used for desktop PCs and notebooks in the PPI and desktop PCs in the HICP is based on the use of predicted prices. Predicted retail prices, derived from single reference period hedonic regressions, are applied to actual retail prices to derive an indirect estimate of the effect of quality change. The latter is then used to adjust, post-hoc, the observed price difference between the outgoing and the replacement item.

The predicted price approach is preferred for two main reasons:

- for models where attributes are related, the omission of one can cause missing variable bias which can be problematic for individual coefficients, though not for the regression equation as a whole.
- where a log price model is used (as is the case with PCs), bias is more problematic for the individual coefficients than for the model as a whole (even after a correction factor of half the standard error is applied).

In a traditional matched model index each good is represented by a series of price quotes, where price collectors select a particular variety to be priced for the whole of the period of the index – in the HICP this is January to January. Each of these is followed for as long as it is available, and a price index is calculated as a ratio of current price to the base price. If the particular variety disappears from the shops, a replacement is sought, and if this is of a different quality then the base price is adjusted to reflect the change. The results of the hedonic regressions are used to calculate these adjusted base prices using the approach described below.

The PPI is also based on a matched model approach and the sampled items are reviewed each year. When a producer no longer manufactures a particular model a replacement will be sought and the base price will be adjusted to reflect a change in quality using the same hedonic function. The arguments for applying an adjustment that has been calculated on the basis of retail prices to the PPI, and the issues that arise are elaborated in the next section. The remainder of this section describes the process of hedonic estimation and quality adjustment in more detail.

The hedonic regressions themselves are calculated on the basis of a single month's data, using unweighted regressions based on list price data from computer magazines for desktops and websites for notebooks. These magazines are searched to obtain information on the price and attributes for as many PCs as possible with the prices collected for at least two hundred computers each month. An electronic template is used which eases the burden of data collection and assists data validation. The advantages are significant, enabling hundreds of observations to be collected over a few days, yielding a substantial set of data for the calculation of hedonic regressions. Timelags are reduced to a minimum and the regression estimates can be produced very soon after the publication of price lists. This compares favourably with the option cost method which is resource intensive and because of this is more restrictive in terms of sample size and timeliness.

Data for hedonic modelling of laptops in the PPI is drawn from information extracted from the websites of manufacturers and dealers rather than from magazines. This is because background research has indicated that for this market, websites cover a greater
A variety of models, give more detailed specifications and also tend to be generally more up-to-date. Information is collected on about 400 laptops each month.

A wide range of attribute data is collected for both PCs and laptops. In the case of PCs, although much of the price is accounted for in “core” attributes such a processor speed and memory size, changes in technology have led to attributes such as graphic and sound cards also having a significant influence. Much work is undertaken to ensure that the hedonic regressions do not suffer from missing-variable bias, so such things as on and off site warranty are also included. Below is a list of the attributes currently collected:

- Processor Type & Speed (CPU Score)
- Memory Size & Type
- Hard Drive
- Monitor size (Flatscreen)
- DVD, CD-RW, DVD-RW, Combo
- Graphic Card Type
- Sound Card Type
- Operating System
- Others (ie Printer, Scanner, Firewire, Sub-Woofer, Digital Camera, On and Off-site warranty)

Unsurprisingly, many of the attributes listed above are also collected for laptops. In addition, several attributes, which relate specifically to portability and computing on the move, are collected. These include:

- Ethernet cards
- Infra red connections
- Battery type
- Physical Dimensions
- Weight

An iterative approach is used to derive the hedonic regressions. This procedure, which includes an element of statistical judgement and product/market knowledge, has been extensively tested. It is based on similar methods employed by the Bureau of Labour Statistics in the USA (BLS). It is preferred over the more traditional automatic stepwise regression technique because it is better able to cope with the potential relationships between independent variables in the regressions. For instance, printers and scanners are often inter-correlated because companies who provide a printer as part of a PC package often bundle in a scanner as well. These relationships can cause the automatic methods of regression estimation to produce either sub-optimal regressions, or in some circumstances ones in which the relationships revealed are counter-intuitive.

Regressions are based on log of price rather than price. There are two main reasons for this. Firstly, a log-linear model produces a multiplicative relationship between price and attributes which is a better reflection of pricing in the retail market. This is because the cost of adding a new feature tends to be related to the underlying quality and price of a machine. For example, the addition of a DVD drive to an expensive PC costs more than for a cheaper PC, because a higher quality drive will be included in the more expensive PC. Secondly, multiplicative relationships are more robust to general changes in price, and so have a longer life span.
The regressions are then used to predict prices, and price adjustments are made based on these predicted prices.

The following is an illustrative example of how hedonic based quality adjustment can be applied in a situation where an individual model was priced in January, but could not be found in February. The replacement is close in quality, but has a single change in specification – an increase in processor speed.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Coefficient</th>
<th>January Model</th>
<th>February Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Level</td>
<td>Level</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Effect on Price</td>
<td>Effect on Price</td>
</tr>
<tr>
<td>Brand</td>
<td></td>
<td>PC Company</td>
<td>PC Company</td>
</tr>
<tr>
<td>Intercept</td>
<td>5.6337</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Monitor</td>
<td>-0.0069</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>Processor Speed</td>
<td>0.0004</td>
<td>2000</td>
<td>2800</td>
</tr>
<tr>
<td>Hard Drive</td>
<td>0.0050</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>Memory</td>
<td>0.0003</td>
<td>256</td>
<td>256</td>
</tr>
<tr>
<td>Video Card</td>
<td>-0.0039</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td>Predicted Price</td>
<td></td>
<td>£588.30</td>
<td>£802.50</td>
</tr>
<tr>
<td>Actual</td>
<td></td>
<td>£625.00</td>
<td>£825.00</td>
</tr>
</tbody>
</table>

(Only change is increase in processor speed)

Note: Predicted price = Intercept x effect of monitor x effect of processor speed x effect of hard drive x ………

**Step 3: Adjust base price to reflect new attributes**

Change to January due to changes in quality = \( \frac{\text{Predicted price new model}}{\text{Predicted price old model}} \)

= £802.50 / £588.30

= 1.364

New base price = Base price old model x quality change

= £625 x 1.364

= £852.56

**Step 4: Compare current price with new**

PC Index = (\( \frac{\text{£825}}{\text{£852.56}} \)) x 100

= 96.8

Unadjusted index = (\( \frac{\text{£825}}{\text{£625}} \)) * 100

= 132
Applying hedonic models on retail prices to the PPI

As mentioned earlier the hedonic regressions ONS use are based on retail information – magazine advertisement prices for desktop PCs, and websites for notebooks. This is obviously appropriate for use in a consumer price index, such as the HICP. However we need to address the question of whether this is appropriate for the PPI.

Clearly, we would prefer to use factory-gate price/specifications transactions data to construct hedonic models for the PPI. But hedonic modelling demands an extensive, rich source of data, and in practice it is not feasible to assemble a sufficient volume of producer price quotes to build a rigorous model. It is for this reason that hedonic models based on retail prices are being applied to the PPI.

Two main issues arise in using the retail-based models to quality adjust producer prices:

- When applying hedonics based on retail prices to producer prices in the PPI, a central assumption is that a specification change that the hedonic model values as an x per cent change in quality at the retail stage, translates to an x per cent change at the factory gate. In practice there will be fluctuations in retailers’ and other intermediaries’ margins. For example, it is often assumed that one of the determinants of margins is the stage within a product’s life cycle, with higher margins being generally associated with newly launched products. The ONS PPI sample replenishment strategy gives a mix of models at different stages of their respective life cycles, which makes this less problematic. It is also worth noting that the BLS, which has accumulated considerable experience in hedonic modelling, also uses a common hedonic model based on website data to quality adjust in both the PPI and CPI series (Holdway 1999).

- It is generally accepted that, in principle, for consumption oriented indices such as the HICP, a user value approach to measuring quality change is appropriate, whereas for production based indices, such as PPI output series, a resource cost based approach is deemed appropriate. As has been pointed out by Diewert (Diewert 2002) and others, this can potentially lead to inconsistency problems in the compilation of the National Accounts. Rosen (Rosen, 1974) has demonstrated that in a perfectly competitive market a hedonic function can provide a single measure that is indicative of both user value and resource cost. For real markets, where few of the conditions that underpin the perfectly competitive assumption are fully satisfied, quality adjustments derived from hedonic models cannot be assumed to accurately reflect either user value or resource cost changes. Nevertheless, the use of hedonics can still be justified on the grounds that it yields better estimates of both user value and resource costs than are obtainable via other approaches.

Finally, it should be noted that there might be situations where it is not possible to use the predicted price approach within the PPI, due to manufacturers being unable to provide information on all attributes. In this case the use of individual coefficients from the models will be adopted as a fall back option. In addition, we will revert to the traditional methods of quality adjustment in individual instances where hedonics does not provide sensible results. This can happen when very large changes in specification take place. It should also be noted that that the traditional production cost measure of quality change will continue to be used for palmtops and handhelds.
Impact on the PPI and HICP

During the course of our comprehensive testing of hedonic methods of quality adjustment in 2002, the fall in PC and laptop prices tended to be steeper than under the existing methodology. However the difference was not large and would not have been discernible in the overall PPI.

For the HICP, the downward impact on the computer price index was more marked though because the weight associated with PCs in the HICP is small the change still would have had little impact on the overall published index.

Strategy for updating Regression models

The plan is to update the models only when necessary subject to a minimum of three times a year. Earlier trials have shown that relying solely on market information to detect emerging changes in specifications and features, to trigger an update in an hedonic model is not totally reliable. The hedonic model will be updated more frequently if a monthly analysis of predicted prices against actual prices shows that the model no longer accurately predicts PC prices. This procedure will be followed within the tight time-scales dictated by the monthly compilation round.

Other Goods

Clearly computing equipment is not the only area in which technological improvements have meant rapid shifts in the market. For this reason we have also been looking at the possible effect of a move to hedonic regression techniques for other goods including televisions, dishwashers, washing machines, vacuum cleaners and cameras.

The quality adjustment technique currently used for these goods in the HICP is an implicit method called bridged overlap (within PPI other goods are quality adjusted using either manufacturers costs or option costs). This is a method in common use amongst NSIs. It assumes that the pure price change associated with the move to a replacement item is equal to the corresponding average price movement experienced by other similar goods that have not been replaced. Any remaining price change observed for the good in question is then attributed to changes in quality. Whilst this method may be criticised for being simplistic, experience suggests it is adequate in markets where there is limited turnover and technological innovation. In addition, it has the advantage of being relatively cheap to apply. Hedonic techniques will need to deliver significant technical improvements in index methodology in other parts of the PPI and HICP to make their wider implementation good value for money.

We have no immediate plans to introduce hedonic quality adjustment methods for the goods listed above. Research has indicated that in current market conditions a move to hedonic techniques would have little impact on the index. The situation will be kept under review through our ongoing research programme to evaluate the use of hedonics elsewhere in the PPI and HICP and the robustness of the method.

Implications of the changes

The change of quality adjustment method will be implemented in the PPI, Export and Import price indices and the HICP from the February indices published in March. No backward revisions will be made to previously published time series since the
construction of historic models is at best both resource intensive and very difficult and in many instances is precluded by lack of past data. The full impact of the change will be spread over twelve months.

For National Accounts, it is generally believed that hedonic quality adjustment methods yield lower price trends and deflators than the previous ONS methods. However the effect of lower deflators is not simply to increase GDP growth rates, as other effects come into play via external trade deflators. Previous work on the sensitivity of the national accounts to hedonics on computers describe this more complex effect (Vaze 2001).

References


