Chapter 3

METHODOLOGY OF THE PRICE INDEX FOR MICROCOMPUTERS AND PRINTERS IN FRANCE

by

Antoine Moreau
INSEE, France

Introduction

Calculating a price index entails defining a basket of products, to be monitored over time. Once this basket is defined, a Laspeyres price index at t may be determined by simply observing the quantities q_{i0} sold in period 0 and prices in periods 0 and t and plugging them into the following formula:

$$I_{t/0} = \frac{\sum_{i=1}^{n} p_{it} q_{i0}}{\sum_{i=1}^{n} p_{i0} q_{i0}} \qquad [1]$$

In practice, a basket of products does not remain stable over time; some products are withdrawn from the market, and others are changed. As a result, new products have to be introduced in order to keep the basket representative. The problem is particularly acute in the realm of computing, but it arises in many other sectors as well, be it for cars, refrigerators, washing machines, and so on.

How can we take account of these products that spring up or are modified – a phenomenon we shall refer to by the generic term "quality factor"? One option is to resort to econometric methods. Such methods have been used at INSEE to calculate selling price indices for microcomputers and printers indices that have been published quarterly since autumn 1991 for microcomputers (going back to 1988) and since autumn 1992 for printers (going back to 1990).

The purpose of this chapter is to assess the experience that has been gained as these indices were being developed. We shall first review the methodological principle involved, and the various ways to apply it. Then, after describing the data available to INSEE for calculating the two indices, we shall outline the methods used to take the quality factor into account.

Harnessing econometrics to take the quality factor into account

Disregarding the quality factor

To fully grasp the reasoning behind the econometric method, let us first show how new products can be incorporated into a price index without taking the quality factor into account. This involves the classical, so-called matched models method, which brings in econometrics naturally.

The matched models method consists in calculating the t/t-1 price index using only those products present in both t and t-1:

$$I_{t/t-1} = \frac{\sum_{i=1}^{n} p_{it} q_{it-1}}{\sum_{i=1}^{n} p_{it-1} q_{it-1}} \qquad [2]$$

If any products were introduced in t, they are not used to calculate $I_{t/t-1}$, the assumption being that the change in the index that is due to the introduction of new products is equal to the average change in the prices of products present in both t and t-1.

The simplified example below illustrates the bias that the matched models method can introduce into the calculation of a price index. Here we take an industry in which each period sees the introduction of a new product of higher quality. The prices of the various products are constant, all equal to 100, and each product has a lifetime of only two periods. The situation can be summarised in the following table:

Periods	1	2	3		11
Quality 1	100	100	100		
Quality 2		100	100		
....				....	
Quality 10				100	100

When calculated using the paired models method, the index remains constant at 100, since the method splices together indices of products that are each present in two consecutive periods. And yet, in period 11, for the same price as in period 1, it is possible to buy a product that is ten times more effective; the price index ought therefore to have dropped.

Over a long period, the measurement of price changes can become considerably biased. Gordon (1990) recalculated a price index for durable goods in the United States – one that spanned the entire post-war period. The contrast between his index and the official one of the Bureau of Labor Statistics (BLS), which used the matched models method for most of the period, is striking. Using 1982 as the base figure of 100, the official BLS index would be 27 in 1947, whereas Gordon's (1990) recalculated index was 78.1. A price index that properly incorporates the quality factor can therefore make a substantial difference in the profile of national accounts.

The above example clearly illustrates what is needed in order to calculate an accurate price index: new products have to be linked to goods already on the market. It is therefore necessary to imagine at what price product 2 would have sold for in period 1, and to calculate the 2/1 index using not just the change in the price of product 1 between periods 1 and 2, but the corresponding hypothetical change in the price of product 2 as well.

The price of a product for the period before its introduction can be estimated “manually” by comparing the new product with existing ones. In the above example, since quality is determined by a single feature, the price of product 2 in period 1 could be estimated to be 200.

Clearly, as soon as product quality is determined by a number of different characteristics, the manual adjustment method becomes tricky to use. What is needed is a method whereby reasoning can be carried out on an “all else being equal” basis, and this is the aim of econometrics.[1]

The basic econometric method

Here the starting point is to pinpoint the characteristics, *X*, that determine the quality of a product. Next, a relationship is posited between a product's price and its characteristics:

$$p = f(X, \beta) \quad [3]$$

where β is a set of coefficients to be estimated. Eventually, those coefficients are used to estimate the price of products for the period prior to their introduction.

To facilitate calculation, it is generally assumed that f is a linear function in the parameters β. In point of fact, the form of the function f can be deduced from economic theory by constructing a complete supply and demand model; it hinges on difficult-to-test hypotheses about consumer preferences and producers' cost functions (see, for example, Moreau, 1993).

Here we shall restrict ourselves to a linear function with the parameters β, which is generally more than adequate for calculating a price index. Consequently, assuming that a microcomputer's only characteristics are its memory capacity (cap) and internal clock (hor), we posit:

$$\ln(p) = \beta_0 \beta_1 cap + \beta_2 hor \quad [4]$$

In equation [4], the logarithm of p has been introduced because experience has shown that, in the case of microcomputers, it yields a far better adjustment than prices in level form.

Given N microcomputers in period t, we estimate the equation:

$$\ln(p_{it}) = \beta_0 + \beta_1 cap + \beta_2 hor_i + u_{it} \qquad i = 1, \ldots.., N$$

using the ordinary least squares method. If a microcomputer with a capacity of 100 and a clock of 25 comes out in period t+l, its price in t will be estimated as:

$$\exp\left(\hat{\beta}_0 + 100\hat{\beta}_1 + 25\hat{\beta}_2\right) \quad [5]$$

where $\hat{\beta}$ is the estimator of the ordinary least squares of β. This estimated price will be used to calculate the index for t/t+1, and thus a microcomputer introduced in t+1 in fact comes into the index in period t.

Note that [5] is not an unbiased estimator of p_{it}. To obtain an unbiased estimator, it would be necessary to introduce a correction factor depending on the variance of the estimated residuals. The bias in all of our regression calculations being very slight, and to simplify the illustration, we shall go no further than formula [5].

The above method has a number of drawbacks and in practice it is often necessary to make additional assumptions. This will be the topic of the following section.

Determining the role of econometrics

Econometrics has frequently been used to incorporate the quality factor into the calculation of a price index: for cars (Griliches, 1971), refrigerators (Triplett and McDonald, 1977), cranes (Early and Sinclair, 1979) and, of course, computers (Cole, 1986). The beginnings of information technology indices in the United States are described in Cartwright (1986) and Cartwright and Smith (1988) (see also Moreau, 1990). Econometrics is also used by official statisticians in the United States to calculate an index of the cost of constructing single-family dwellings. Gordon (1990) uses econometric methods as well, but they are only part of his panoply of tools.

INSEE's experience in this area can be summed up as follows: econometrics provides a convenient way of dealing with quality factors but brings nothing new to the actual concept of a price index. Once the characteristics that determine a product's quality have been defined, econometrics must yield the same result as the "manual" method, which compares products two at a time. While the latter does not always produce an exact estimate of the price of a new product, it generally offers a range in which the econometrically estimated price will fall. If it does not, it is probably because the wrong functional form of f – see [3] – has been chosen to relate price to characteristics. In calculating the index, every possible effort was made to corroborate econometric results by inspecting each firm's catalogues very closely.

Applying the econometric method

Drawbacks of the basic method

The basic method described above has one major drawback: inherently, new products, at least new microcomputers, often lie outside the cluster of points from which the equation can be estimated. It is well-known that this causes maximum distortion when projections are made using the ordinary least squares method, the inaccuracy of which increases with distance from the point cluster on which an estimation is based. This observation was made repeatedly as the French index for microcomputers was being calculated, whenever it was attempted to estimate prices for t-1 of products introduced in t, using an equation estimated for period t-1 only.

In this case, there are two possible solutions:

- To estimate an equation over a number of periods, imposing a certain form on the distortion over time of the relationship between prices and characteristics. This is what was done for microcomputers. But which form should be imposed on the temporal distortion of the price/characteristics relationship and how can its validity be tested? In most cases, the indicators method is used.
- To limit the estimation to a single period, but not use [5] to estimate prices. Later we shall spell out how to make an estimate that, empirically, seems more precise. We shall refer to this process as the "characteristics method".

Moreover, it often happens that there are not enough observations in any one period to come up with a good estimation. The only way around this problem is to estimate the relationship [3] over several periods.

The principle behind the indicators method

Let us go back to our example, in which a microcomputer's memory capacity and internal clock are the only variables. We have N_t observations for period t and N_{t-1} for t-1. Some microcomputers from period t-1 were gone in t, but others were introduced. In its simplest form, the indicators method consists in estimating the following model:

$$\ln(p_{it-1}) = \beta_0 + \beta_1 cap_i + \beta_2 hor_i + u_{it-1} \qquad i = 1, \ldots.., N_{t-1}$$

$$\ln(p_{it}) = \beta_0 + \beta_1 cap_i + \beta_2 hor_i + \beta_3 + u_{it} \qquad i = 1, \ldots.., N_t$$

The same coefficients are therefore imposed for the characteristics in both periods, the change in price being summarised by β_3. It would be possible to weight the equations by the quantities sold of product i. This would make the t/t-1 price index:

$$I_{t/t-1} = \exp(\hat{\beta}_3) \qquad [6]$$

There is no reason to limit an estimation to two consecutive periods. The above regression can be carried out over T periods, forcing characteristics coefficients to remain constant throughout, and setting a new indicator for each period. Clearly, the longer the period over which these coefficients have to be equal, the greater the constraints on the data.

The advantage of this method is that calculations are easy and fast; the considerable drawback is that the index can change even if no new products are introduced and all prices stay the same. For example, to return to the example above, if we assume that no new microcomputer came on the market in t+1, that on average those that were withdrawn in t+1 were more expensive than the others (owing to characteristics not included in the equation) and that no prices were changed, then the estimator of β_3 can be significantly negative.

As we shall show in the Annex, this can stem from the fact that the variable corresponding to the β_3 coefficient – *i.e.* the indicator of whether a product has been withdrawn yet – is correlated with characteristics that are not included in the equation. In econometric terms, the problem is one of endogeneity.

This property is obviously unacceptable. A trial run on microcomputers in France showed that the indices obtained via this method for the second quarter of 1992 were only a third to half as much as the indices that were published.

This problem of endogeneity could be dealt with conventionally by applying an instrumental variable to the corresponding indicator, but it is difficult to find a proper instrument. Moreover, it is important to keep a method of estimation that is as simple as possible.

We did not try to resolve this problem of endogeneity while computing our indices, since in practice it is not the simple version of the indicators method, as presented above, that is generally used. In point of fact, the changes that are actually introduced probably lessen the extent of the bias induced by any endogeneity.

The basic indicators method can be changed in two ways:

- First, by actually incorporating the prices of products present in t-1 and t in the index $I_{t/t-1}$. This illustrates a point already mentioned: prices contained in a database should be exploited as fully as possible. Accordingly, rather than compute the index directly using estimated coefficients, it is better to proceed as follows: let J_0 be the set of all products present in t and t-1 and J_1 the set of all products introduced in t. The index t/t-1 will be:

$$I_{t/t-1} = \frac{\sum_{j_0} p_{it} q_{it-1} + \sum_{j_1} p_{it} q_{it-1}}{\sum_{j_0} p_{it-1} q_{it-1} + \sum_{j_1} \hat{p}_{it-1} q_{it-1}} \qquad [7]$$

where $\hat{P}_{it-1}$ are the estimated prices for t-1 of products introduced in t.[2]

- Second, by making the price/characteristics relationship in the estimated equation more flexible. One of the shortcomings of that equation is, in particular, that it forces all new products to have the same estimated change in price, which in turn is equal to that of products present in both periods.

Refinements in the indicators method

We are now going to review a number of ways in which the relationship between prices and characteristics can be made more flexible. These are obviously not exhaustive; we present only those that have been used for the indices calculated at INSEE.

It may be noted at the outset that the problem of estimation facing us here is anything but conventional: as was pointed out in the first section, we generally have a fairly accurate idea of the end result. Before applying classic statistical tests in order to choose any particular specifications, we may reject any estimate that would yield results that fall outside the range estimated by looking at catalogues.

The basic idea for enhancing the flexibility of the indicators method is very simple: there must not be just one β_3 coefficient for all products present in t, but different coefficients to be applied to large well-defined families of products. Traditionally, a distinction is made for period t between products that were present in t-1 and those that came on the market in t (see, for example, Triplett and McDonald, 1977; or Griliches, 1971). For t, the equation becomes:

$$\ln(p_{it}) = \beta_0 + \beta_1 cap_i + \beta_2 hor_i + \hat{\beta}_3 + u_{it} \qquad (products\ present\ in\ t-1)$$

$$\ln(p_{it}) = \beta_0 + \beta_1 cap_i + \beta_2 hor_i + \beta_4 + u_{it} \qquad (new\ products)$$

In this case, the estimated price for t-1 of a microcomputer introduced in t at a price of p_t would be:

$$\hat{P}_{t-1} = p_t \exp\left(-\hat{\beta}_4\right) \qquad [8]$$

$\hat{\beta}_3$ is not used to compute the index, but it can be compared with the price change actually observed in order to check that the specification is not too inaccurate.

It may be noted, following the reasoning of the ordinary least squares method, that the estimator of p_{t-1}, should be:

$$\hat{P}_t = \exp\left(-\hat{\beta}_4\right) \qquad [9]$$

Intuitively (we have not demonstrated it), formula [8], which uses the observed p_t, yields a more accurate estimator than [9]. This too corresponds to an empirical observation.

This method has been used systematically to calculate the microcomputer index. Let us note that it is not very logical to use the method continuously: in the t-1/t regression products introduced in t are treated differently from those already present in t-1. In contrast, all these products are treated similarly in the t+l/t regression, in which only those products introduced in t+1 are treated separately. This imposes a particular price change profile that is based on no explicit economic assumption. Berndt and Griliches (1990) proposed adding generation indicators to equations, having each product in the regression pinpointed by a coefficient indicating its period of introduction. With very few exceptions, we did not use this idea – again, for the sake of simplicity.

Another option is to differentiate price changes by product family (type of processor) and by manufacturer. Both distinctions were used extensively in calculating the index; we shall return to the subject when we explain exactly what was done for microcomputers.

The characteristics method

The first use for the characteristics method is when an existing product is changed. To illustrate, let us take the example of a microcomputer whose price is monitored for the index, and whose clock rate is increased from 20 MhZ in period t-1 to 25 MhZ in period t, with all other characteristics staying the same. Let us assume that the price remains constant at FF 10 000. We estimated equation [4] and found $\hat{\beta}_2$ to equal 0.02. In calculating the t/t-1 index, for t-1 the FF I0 000 price must be replaced by:

10 000 exp[0.02(25-20)] = 11 052

Thus, rather than estimate the price for the preceding period by using formula [5], we use only the coefficient of the characteristic that was altered. The underlying logic is the same as our reasons for preferring [8] to [9] – estimation is more accurate.

The same principle can be applied to new products. Whenever a manufacturer adds a new product to its catalogue in t, we look for the product sold in t-1 that is most similar and we estimate the missing price in the same way as if a product had simply been changed. In practice, a new product must not be linked to one that is too dissimilar. It is necessary to find the one nearest in terms of the characteristics plugged into the equation, *i.e.* to ensure that the above equation has the smallest possible number of terms in the exponential. But it is also necessary to take account of the type of characteristics differentiating the two models being linked. Accordingly, we feel it would be dangerous to link two microcomputers that do not share the same type of processor. Finally, it is necessary to incorporate characteristics that are not included in the equation (because they cannot be quantified).

There is obviously an element of subjectivity involved in choosing the model to be linked to a new product. The BLS makes its choices using a breakdown-into-main-components method. In our view, the drawback of such a method is that it is too automatic. In particular, there is no guarantee that the two products being linked are made by the same company and one can never be certain that prices charged by two different firms are comparable.

Which method to use?

Four methods have been used to introduce a new product into price indices for microcomputers and printers:

- Close examination of catalogues to deduce missing prices by studying the characteristics and prices of products sold in t-1. This is the first method to apply, since it generally yields at least a range for the estimated price.
- The indicators method, with sufficient differentiation of indicators by families of products.
- The characteristics method.
- The matched models method, which was used only as a last resort.

Once the characteristics that affect product quality have been determined, the first three methods ought to produce the same result. If not, the model has been poorly constructed.

Which econometric method should be used? The indicators method is quick: it allows the same change in price to be applied to an entire family of new products on the market, however, for the same reason it is more approximate. The characteristics method entails finding the best matches between new products and old, and calculating price changes per product, and is therefore more cumbersome.

In fact, everything depends on the type of data available. If there are many observations and relatively few characteristics, it would be possible to make sufficient distinctions between indicators by product family to obtain good results; such is the case for microcomputers that are not portable. If the reverse is the case, it would be better to use the characteristics method; this is the case for portable computers and especially printers. For the latter, there are many characteristics that are hard to quantify, and few observations sharing the same characteristics. The indicators method would not yield a sufficiently accurate result.

The data used

The firms taking part in the price indices for microcomputers and printers are the main firms present on the French market: Apple, Bull-Zenith, Compaq, Epson, Hewlett-Packard, IBM, NEC, Olivetti and Toshiba. Not all these firms manufacture in France. It was clear from the start that it would be impossible to ask these companies to distinguish between what is made in France and what is not: the concept of domestic production and imports means little to these large multinational firms, and correspondents for a survey such as the one on industrial selling prices do not have at their disposal the statistics they would need to distinguish one from the other.

We must now specify the exact scope of the indices calculated: the terms “microcomputers” and “printers” which have been used so far are imprecise. With respect to microcomputers, the index covers all monoprocessors equipped with Intel (from the 8086 to the Pentium) or Motorola (from the 68000 to the 68400) microprocessors and listed in the catalogues of participating manufacturers. We do not monitor what are commonly known as “workstations”. The printers covered by the index are models that are intended to be connected to a computer and that use needle, laser or ink-jet technologies.

All products listed in the catalogues of participating manufacturers are incorporated in the index. This provides more observations for regressions, and also enables new products to be integrated as soon as they come out without having to rely on the firms questioned for product renewal. As a result the number of products monitored changes quarterly. In the third quarter of 1993, we monitored 440 microcomputers and 119 printers.

Prices

Until recently, virtually all of the microcomputers bought in France by businesses were sold through distributors. The purpose of the industrial selling price index is to track changes in the prices charged to the first buyers on French soil – in this case, distributors.

Until the end of 1991, all microcomputer firms published recommended list prices. These prices provided a basis for negotiation with distributors. Actual street prices were much lower than list prices because distributors' discounts varied between 45 and 50 per cent. At the end of 1991, and over the course of 1992, one business after another decided not to release list prices. Nevertheless, they maintain a confidential internal price scale intended for distributors. Along with this change came a steep drop in discounts.

The prices gathered by INSEE are recommended list prices and their successors. The cuts in discounts in the winter of 1991 were reflected in the index to the extent that firms informed us thereof. In early 1994, it became clearly apparent that two big players in the microcomputer market had not been reporting their discounts correctly. Publication of the index was suspended, and the pattern of past discounts was reconstituted. Publication resumed in autumn 1994, with corrections to the indices that had been published in 1992 and 1993.

In the absolute, recommended list prices do not correspond to the prices of any transactions. In contrast, all the people we talked to in these firms felt that, for the purposes of our survey, it was sufficient to monitor those prices. Clearly this does not enable us to detect all changes in price (in particular those at the end of a product’s life), but it does reveal the most important changes.

A number of observations lead us to believe that monitoring such prices was satisfactory between 1988 and 1992:

- ♦ We had access to highly specific data from one of the participating companies, breaking down 1990 turnover by product. Those data showed clearly that it was justified to assume that the change in invoiced prices was approximately equal to the change in list prices.

- ♦ Recommended list prices change very frequently, sometimes several times per quarter, and the magnitude of the price changes is often substantial.

To avoid the problems encountered with regard to 1992-93, a specialised pollster now calls on firms every other year to take stock of the situation.

Weights

At the most disaggregated level, microcomputers are divided into families according to manufacturer, type of processor and portability; for example, IBM 486 desktop; Compaq 386 portable. At present, the index covers the following processor types: 86, 286, 386, 486, Pentium and 68000, 68020, 68030 and 68040. No distinction is made between, say, a 386SX and a 386DX. In the case of printers, the characteristics crossed in order to constrict these basic families are print technology (impact, laser, bubble-jet) and manufacturer.

The families are the first level of aggregation. Within a given family, all microcomputers (or printers) have the same weight in terms of sales figures, leading to the assumption that the quantity sold in the base period (q_{io} in formula [1]) is inversely proportional to the price in the base period.

The families are weighted in the aggregate indices by their sales on the French market. Each firm consulted concerning its computers is asked for a breakdown of its turnover according to processor type and whether or not it is portable. Likewise, firms consulted on printers are asked to provide a breakdown according to the type of technology used.

Sales figures are requested annually and each year the index base period changes, so that the indices are of the Laspeyres chain type. Until publication was temporarily suspended in 1994, indices for year n were calculated with the weights for year n-1. Since then, the same method has been used to calculate provisional indices. These are then revised once the weights for year n are known, so that the final indices for year n are calculated with weights for year n.

Product features

This section describes the features selected in order to take account of the quality factor. They are taken from the manufacturers' catalogues. For printer features, we also consulted *L'ordinateur individual*, a monthly computer magazine giving technical data on products available on the market.

Microcomputers

The features of the desktop models are confined to the central processing unit. The monitor is generally sold separately and has not been taken into account for the purpose of dealing with the quality factor. For the purpose of calculating the index, however, each CPU has been connected to a standard monitor (usually a 14" VGA colour display), so the index accurately reflects price changes for the two units together. Apple is the only manufacturer in the sample to have sold microcomputers with monitors that could not be separated from the CPU. Moreau (1991*b*) explains how this particular case was dealt with.

The features taken into account are:

- ♦ type of processor: 86, 286, 386SX, 386DX, 386SL, 486SX, 486DX, 486DX2, Pentium, 68000, 68020, 68030, 68040;

- clock speed (in MhZ);
- standard RAM (random access memory) and maximum RAM installable (in MB);
- hard disk memory capacity (in MB);
- number and type of floppy disk drives;
- the capacity of the CPU to manage more than one hard disk (below we use the term "mini-tower" borrowed from Hewlett-Packard);
- the number and type (8, 16, 32 bytes) of expansion slots.

The following additional features are taken into account for portables:

- type of monitor (Plasma, backlit LCD, etc.);
- display resolution and size;
- stand-alone capability;
- weight of stand-alone portables;
- presence of expansion slots, or possibility of installing them.

Printers

Printer features are more difficult to define than features of microcomputers, principally because they vary according to the technology used (dot-matrix, laser or bubble-jet), but also because some features, particularly with respect to paper handling, are not easy to quantify or codify synthetically. The features we take into account are:

- Printer technology: dot-matrix (9, 18, 24, 48 pin), ink-jet (with the number of print-heads), laser;
- Print speed. In the case of dot-matrix printers, print speed depends on image density. The speed taken into account is maximum print speed, in characters-per-second (cps) for dot-matrix printers, in pages-per-minute (ppm) for laser printers, and in either one of these units for ink-jet printers.
- Standard memory and maximum memory installable, in kB.
- Resolution, in dots-per-inch (dpi). In the case of dot-matrix and ink-jet printers, there are generally two numbers given (one for vertical resolution, the other for horizontal resolution).
- The number of columns (80 or 136), the standard number of fonts, and colour print capability. These features apply only to dot-matrix and ink-jet princers.
- Noise level, in decibels.
- The standard number of trays.
- The type of interface (PC or Appletalk, or both).
- For laser printers, the availability of Postscript, and curve-smoothing capability.

Lastly, for some printers *L'ordinateur individuel* gives details concerning average monthly capacity (laser) and head life (ink-jet).

How the quality factor was taken into account

Periodicity of the estimations

It has already been pointed out that products are incorporated in the index as soon as they come on the market. If a product comes on the market in quarter t, its price in a previous period has to be estimated in order for it to be incorporated in the index as from the calculation of period t. But how do we choose the estimation period?

One way, which seems to be the method adopted by the BLS in the United States (see Sinclair and Catron, 1990), would be to estimate an equation in a base period and to use the equation to take account of the quality factor over several quarters, or even years. This method has few drawbacks in theory, but is tricky to apply in practice. For example, when should the equation be re-estimated, and according to what criteria? We have chosen to calculate the prices of new products in the period immediately before they come on the market, which involves estimating an equation in each quarter.

Dealing with the brand effect

In addition to all the detailed technical features set out above, another factor exerting a powerful influence on the market for microcomputers (at least until 1992) is the "brand effect", which enables some firms to sell their products at a higher price than other firms offering the same technical features simply because they have a better image or are more fashionable, or because some features (*e.g.* after-sales service) are not taken into account by statisticians.

Using a straightforward supply and demand model, it is easy to show that there is no reason why the functional form [3] linking prices to features need be the same for all firms operating on this market (see, for example, Moreau, 1993). That being the case, how can we build a model for this brand effect?

The most straightforward method is to estimate an equation for each firm. One reason why we opted to use the entire catalogue of each firm to calculate the index was that this method, used until summer 1993 for desktop microcomputers, is only possible if enough observations are available for each firm. A less flexible way of dealing with the brand effect is to introduce an indicator for each firm in the estimated equation (or to cross certain features with these indicators). This is the method used for printers and portable microcomputers, and for desktop microcomputers from the third quarter of 1993 onwards.

The following section looks in detail at the methods used for each type of product.

Desktop microcomputers

Two methods

The principle applied until the end of 1993 was as follows: the indicators method was used, with a separate regression for each firm over two consecutive periods, and a distinction was made between indicators according to the type of processor (in particular between 386SX/386DX and 486SX/486DX/486DX2) and whether or not the product was new on the market. Apple, with few references in its catalogue, was the only exception to this principle. In the case of Apple, the indicators method was used over more than two periods.

Moreau (1991*b*) provides an assessment of the indicators method; he presents a number of regressions and reviews the various econometric problems encountered with the method. We shall not comment here on the results.

The econometric method was used for two years and then modified. It was probably too costly and time-consuming. Estimating a regression for each manufacturer is very laborious, in addition to which the fact that some manufacturers' catalogues contain relatively few references can casts doubt on the reliability of certain estimations. This is not really a problem for the index, though, given that the estimations are not being drawn up in the normal way. The econometric results are used only insofar as they seem to be compatible with a detailed survey of the manufacturers' catalogues.

The purpose of the change was to simplify the method used, by incorporating all the firms in the same regression, which was estimated over more than two consecutive periods. The index is computed using the indicators method, whereby in each period products are differentiated:

- By firm, to take account of the brand effect, and also the fact that prices are not exactly comparable between two different firms, even though the discount rates applied by the firms in the sample are actually all very similar.
- By processor: 286/386/486. No further distinction is made (386SX, 386DX, etc.) in the indicators used to calculate the index.
- According to whether a product is new or not. The aforementioned inconsistency still applies, inasmuch as a product is only treated differently to its predecessors in the period in which it comes on the market.

Concrete examples of indicators are the IBM 386 in the first quarter of 1989, already present in the previous period, and the Compaq 286 in the second quarter of 1990, marketed for the first time in that same quarter, etc.

Finally, account must be taken of the fact that Apple is the only manufacturer in the sample to use Motorola processors. To incorporate Apple products in the same regression as other firms, the 68000 has been likened to a 286, the 68020 to a 386SX, the 68030 to a 386DX, and the 68040 to a 486DX.

The indices calculated using the two different methods have been meticulously compared, although statistical secrecy unfortunately prevents us from presenting the results of the comparison in this chapter. Suffice it to say that the differences between the two indices were deemed sufficiently minor for the initial method (undoubtedly more cumbersome but probably also more accurate) to be abandoned.

Results of the regressions for desktop models

The results of two different regressions are given below. One uses prices collected since 1988, while the other does not start until 1991. It is important, when using the indicators method over several consecutive periods, to make sure that the period of time chosen does not affect the result. We shall see that there are slight differences in the coefficients of the features, but once again the indices calculated by means of the two regressions (again, not presented here for reasons of confidentiality) are very similar.

The explanatory variables used are the constant 11 variables describing the features of the microcomputer and the indicators.

The regression beginning in the first quarter of 1988 and ending in the second quarter of 1992 is based on 2 375 observations and there are 273 indicators. For the regression starting in the first quarter of 1991, the figures are 1 415 observations and 123 indicators.

The following table gives the results of both regressions for the variables other than the indicators used to calculate the index.

Dependent variable: logarithm of price	88Q1 - 92Q2	91Q1 - 92Q2
Constant	7.755 (0.105)	8.014 (0.133)
386SX processor	0.400 (0.169)	-0.266 (0.226)
386SL processor	-0.095 (0.048)	-0.125 (0.053)
486SX processor	-0.019 (0.209)	-0.054 (0.024)
486DX2 processor	-0.284 (0.039)	-0.258 (0.043)
Hard disk space in MB	0.0008 (0-00002)	0.0007 (0.00003)
No hard disk	-0.224 (0.009)	0.210 (0.013)
Logarithm of internal clock (286, 386SX, 68000, 68020)	0.486 (0.054)	0.615 (0.072)
Logarithm of internal clock (other processors)	0.713 (0.028)	0.627 (0.07)
Logarithm of standard RAM	0.162 (0.010)	0.187 (0.015)
Logarithm of maximum RAM installable on mother board	0.025 (0.007)	0.032 (0.009)
Vertical model	0.213 (0.010)	0.248 (0.014)
Number of observations	2 375	1 414
Number of variables	285	135
R^2	0.9573	0.9598

The signs and values of the coefficients are wholly acceptable. When constructing the index, we were always careful not to use a regression in which a coefficient had an unexpected sign (*e.g.* a negative sign for random access memory). The only significant differences between the two regressions are the position change of the 386SX indicator in relation to the 486SX and 486DX processors, and the respective values of the internal clock coefficients. There is justification for splitting this variable in two in the first regression, but in the second the coefficients are not significantly different.

Calculation of the index

The following example shows how the regression results are used to calculate the index.

Let us suppose we want the index to include Firm A's 286 processors which came on the market in the first quarter of 1989 (89Ql). The values for the relevant indicators are as follows (on the basis of the 88-92 regression).

88Q4: Firm A's 286s already present in 88Q3	1.256 (0.188)
89QI: Firm A's 286s come on the market	1.080 (0.188)

The 88Q4 price of the models marketed in 89Q1 is obtained by multiplying their 89Q1 price by exp(1.256-1.080). Models present in both periods are incorporated in the index in the usual way.

Portable microcomputers

Portable microcomputers cannot be dealt with in the same way as desktop models. There are few references available and it is not possible to differentiate the indicators according to firm and product family. Using all the available periods, we would obtain a regression with 230 indicators for slightly over 600 observations. Several trial runs have been carried out using the indicators method without differentiating the indicators by firm or processor type. The brand effect has been taken into account in these regressions either by means of indicators or by differentiating one of the features by firm. Moreau (1991*b*) presents a regression in which hard disk space interacts with the firm indicators.

In the end, the method which gives the most acceptable results is the features method, with regression over a single period and all firms taken together. When a new model comes out, there are two possible options: either it can be likened to a model in the previous period, in which case the equation coefficients are used to estimate the price of the new product based on the price of the model to which the new product is likened, or else the new product cannot be likened to a model in the previous period, in which case the new model is added using the matched models method.

The question concerning the choice of model to which the new product is to be likened has already been addressed. Above all, it would seem dangerous to link two portables which do not have the same processor, monitor or extension facility.

The following table shows an estimated equation for portable microcomputers in the second quarter of 1992. The equation has been estimated solely for stand-alone portables. The formulation has been kept as simple as possible, the details of certain variables having been regrouped following a Fisher test. Having said that, we chose:

- not to include indicators for the type of processor, the power of the portables being represented only by their internal clock.
- to incorporate firm indicators directly, without having them interact with other variables.

These basic choices were made in order to keep the number of variables to a minimum, bearing in mind the limited number of observations.

Dependent variable: Logarithm of price	
Constant	8.196 (0.532)
Hard disk space in MB	0.004 (0.001)
Logarithm of internal clock	0.316 (0200)
Logarithm of standard RAM	0.270 (0.117)
Logarithm of maximum RAM installable	0.170 (0.088)
Weight	-0.002 (0.032)
Colour monitors	-0.800 (0.178)
Other monitors	-0.415 -(0.137)
Extension possible or already installed	0.248 (0.081)
Firm A	0.581 (0.108)
Firm B	-0.261 (0.068)
Firm C	0.049 (0.173)
Firm D	0.361 (0.123)
Firm E	0.082 (0.099)
Number of observations	52
R^2	0.9259

Printers

As with portable microcomputers, there are few references for printers. Since the indicators method would be far too imprecise, we use the features method.

The results obtained for printers are less conclusive than those obtained for computers. We have not managed to estimate a satisfactory equation for ink-jet printers, owing to a lack of statistically significant features, and wrong-sign coefficients. Econometrics is therefore not used for these products.

By contrast, the estimated equations for dot-matrix and laser printers are used to take account of the quality factor, even though there are few significant variables:

- for dot-matrix printers, the significant variables are speed, the number of needles, the number of columns and colour capability;
- for laser printers, the significant variables are speed, the number of trays and the availability of Postscript as a standard feature.

Annex

ENDOGENOUS BIAS IN THE INDICATORS METHOD

Taking the econometric equation in the basic indicators method, we estimate the equation:

$$\ln(p_{it-1}) = \beta_0 + \beta_1 cap_i + \beta_2 hor_i + u_{it-1} \qquad i = 1,\ldots..,N_{t-1}$$

$$\ln(p_{it}) = \beta_0 + \beta_1 cap_i + \beta_2 hor_i + \beta_3 + u_{it} \qquad i = 1,\ldots..,N_t$$

This may be expressed as:

$$Y = X\beta + \beta_3 Z + u$$

where Z is the indicator which has a value of 1 in period t and of 0 in period t-1. The variables X are the observed product features.

To use notations used in panel data analysis, let us suppose that error term u_i is the sum of a specific individual effect η_i, which denotes unobserved features of the computer, and a residual effect v_{it}. In order for the ordinary least squares to converge, it is sufficient that:

$$E(v|X) = E(v|Z) = E(\eta|X) = E(\eta|Z) = 0 \qquad \text{[A1]}$$

This means that the variables omitted from the equation are not correlated with either the features of the microcomputers or the indicator Z.

Still assuming that the observable features of the microcomputers are not correlated with the individual effect, let us now suppose that the microcomputers which are withdrawn from the market in period t-1 are actually more expensive on average than those that remain on the market. Therefore, if D is a random variable which has a value of 1 if the microcomputer remains on the market in t and of 0 if it is withdrawn in t-1, it is assumed that:

$$E(\eta|X) = 0 \qquad E(\eta|D=1) < 0 \qquad \text{[A2]}$$

Thus, the estimators of β and β_3 by the ordinary least squares method are not convergent. In fact, $\hat{\beta}_3$ is negatively biased. Where M_x is the orthogonal projection matrix on the orthogonal of X and N, the number of observations:

$$\hat{\beta}_3 = \beta_3 + \left(Z' M_x Z\right)^{-1}\left(Z' M_x u\right)$$

and therefore:

$$p\lim\left(\hat{\beta}_3 - \beta_3\right) = p\lim\left(\left\{\frac{Z' M_x Z}{N}\right\}^{-1}\right) p\lim\left\{\frac{Z' M_x Z}{N}\right\}$$

Thus, if the microcomputers withdrawn from the market are more expensive on average than those remaining on the market, assuming the same observable features, the bias of the relevant indicator coefficient is negative. In a context of falling prices, the result is an exaggerated drop in the prices in the index, which is exactly what happens using this method.

The changes which can be made to the indicators method (and which are set out in the third section of this paper) do not eliminate the bias, since indicator Z above is always equal to the sum of the indicators plugged into the equation. However, the bias is probably not as great and its impact on the index is also reduced if we use the prices of products present in both periods (formula [7]).

NOTES

1. Here we have no intention of getting into the debate over the use of data on changes in product costs to deal with the quality factor. On this, see Triplett (1979).
2. Note that this chapter does not deal with the treatment of products that are withdrawn. It would be conceivable to project changes in their prices that would differ from the average change in the prices of remaining products.

REFERENCES

BERNDT, E.R. and Z. GRILICHES (1990), "Price Indexes for Microcomputers: An Exploratory Study", NBER Working Paper No. 3378, Cambridge, MA.

CARTWRIGHT, D.W. (1986), "Improved Deflation Purchases of Computers", *Survey of Current Business*, March.

CARTWRIGHT, D.W. and S.D. SMITH (1988), "Deflators for Purchases of Computers in GNP: Revised and Extended Estimates 1983-88", *Survey of Current Business*, November.

COLE, R. (1986) "Quality-adjusted Price Indexes for Computer Processors and Selected Peripheral Equipment", *Survey of Current Business*, January.

EARLY, J.F. and J.H. SINCLAIR (1979), "Quality Adjustment in the Producer Price Indexes", in M.F. Foss (ed.), *The US National Income and Product Accounts*, Studies in Income and Wealth, No. 47, NBER.

GORDON, R.J. (1990), "The Measurement of Durable Good Prices", NBER Monograph.

GRILICHES, Z. (1971), "Price Indexes and Quality Changes", Harvard University Press.

MOREAU, A. (1990), "Changements de qualité et indices de prix : l'exemple du prix des ordinateurs", Note interne INSEE, division PV1 47 1/E 140.

MOREAU, A. (1991*a*), "A Price Index for Microcomputers in France, Document de Travail de la Direction des Statistiques Economiques 9109.

MOREAU, A. (1991*b*), "Un indice de prix pour les micro-ordinateurs", *Courrier des Statistiques*, pp. 58-59, 52-54.

MOREAU, A. (1993), "Market Power of Personal Computer Producers in France", University College London, D.P. 9307.

SINCLAIR, J. and B. CATRON (1990), "An Experimental Price Index for the Computer Industry", *Monthly Labor Review*, October.

TRIPLETT, J.E. (1979), "Concepts of Quality in Input and Output Price Measures: A Resolution of the User-Value Resource-Cost Debate", in M.F. Foss (ed.), *US National Income and Product Accounts*, Studies in Income and Wealth, No. 47, NBER.

TRIPLETT, J.E. and R.J. McDONALD (1977), "Assessing the Quality Error in Output Measures: The Case of Refrigerators", *Review of Income and Wealth*, pp. 137-176.