7 ADJUSTING FOR QUALITY CHANGE

Introduction

7.1 The measurement of changes in the level of consumer prices is complicated by the appearance and disappearance of new and old goods and services, as well as changes in the quality of existing ones. If there were no such complications, then a representative sample could be taken of the items households consume in period 0, their prices recorded and compared with the prices of the same matched items in subsequent periods, say *t*. In this way the prices of like would be compared with like. However, such complications do exist. For example, an item may no longer be produced in period $t + 1$, so its price comparison cannot be undertaken between periods 0 and $t + 1$.

7.2 A number of methods are available to remedy this. A replacement item may exist in period $t + 1$. If it is of the same quality, its price can be compared with the “old” item’s price in period *t*. But the replacement item may well be of a different quality. One option is to ignore the quality difference and continue comparing the price of the “new” replacement item in $t + 1$ with that of the old one in period *t* to continue the series. An adjustment for the difference in quality is still being made; it is just that it is a very poor adjustment, because the change in quality has no effect on the price. A second option is to exclude from the index those items for which quality changes, and to compile the index link between *t* and $t + 1$ only for matched items having characteristics that are the same. This exclusion amounts to an implicit quality adjustment, one that assumes the overall price change of existing matched items will be the same as the quality-adjusted price change between the missing old and replacement new items. In reality, however, price changes generally vary over the stages of a product’s life cycle. Price changes at the time of, say, a model’s upgrade – when an item is missing and replaced – may be quite different from those at other stages. The implicit assumption may therefore be inappropriate. Third, the price change of a new replacement item may be spliced onto the index if the prices of the disappearing and replacement items are available in a common overlap period, say period *t*. The old item’s price change between periods 0 and *t* is multiplied by the replacement item’s price change between periods *t* and $t + 1$. Yet again, there is an implicit quality adjustment, one that requires the price difference between the old item and its replacement in period *t* to reflect the effect of the quality difference on price. Such differences may also be in part the result of strategic price-setting behaviour related to the period in the item’s life cycle.

7.3 There are other methods of adjusting the prices of non-comparable replacements for quality differences, including ones that use explicit estimates of the effect of the quality change on price. There are a number of methods of deriving such explicit estimates, and the suitability of explicit quality adjustments depends as much on the method used as on the availability of appropriate data to implement the method. In each case, whatever procedure a statistical office follows, a quality adjustment to prices *is* made in every period when an item is not available. The purpose of this chapter is to help ensure that these quality adjustments are the appropriate ones.

7.4 There are three main reasons for considering how to adjust for quality change. First, the scale and pace of methodological innovations are substantial. Second, there is a lack of consistency in the methods chosen by statistical offices for dealing with quality changes; thus comparisons of consumer price indices between product areas, across countries, and over time may be misleading. Finally, a number of empirical studies on the effects of using different methods found that choice of method does indeed matter (Dulberger, 1989; Armknecht and Weyback, 1989; Moulton and Moses, 1997; Lowe, 1996).

7.5 Against these concerns, it must be recognized that statistical agencies do guard against quality changes by using the matched models method. Price collectors record the features of selected items and collect prices for the very same models in subsequent periods in order to compare like with like. If a product group exists in which there are no items whose quality changes and no new or disappearing goods and services, then the matched models method based on representative items works. More generally, three potential sources of error arise from the matched models approach: missing items, sample space change, and new products.

Why the matched models method may fail

7.6 The long-run price change for an item is measured by comparing the price of the item in the current period with that in the price reference period, the period in which it, along with most other items, entered the sample.

Missing items

7.7 The first source of error, and the focus of this chapter, is when an item is no longer available in the outlet. It may be discontinued or it may not be available to the same specification – its quality has changed – and it is effectively missing in the current period. The item's price may be missing for other reasons. It may be a seasonal item or one whose price does not need to be recorded so frequently, or it may be that the item is a custom-made product or service, supplied each time to the customer's specification.

7.8 It is necessary to distinguish between items that are permanently and temporarily missing. Items that are *temporarily* missing are items not available and not priced in the month in question, but that are priced in subsequent months. The items may be missing because, for example, demand is seasonal, as is the case with some fruits and vegetables, or there are shortages. Some commodities are priced on a less frequent basis, maybe quarterly or biannually, because their price changes are irregular. They are therefore missing when they are "off cycle".

7.9 The concern with seasonal items is to impute their missing prices until the item reappears. The imputation methods used are similar in some cases to those used for quality adjustment. The temporary nature of the imputation, however, requires that they be separately identified by the respondent as "temporarily missing" or "seasonal". Principles and methods for such imputations are outlined by Armknecht and Maitland-Smith (1999) and Feenstra and Diewert (2001), and in Chapter 22. The concern in this chapter is with permanently missing items and with making imputations on a continuing basis or using a replacement item.

7.10 A number of approaches are available for dealing with missing items:

- The item may be dropped on the assumption that the aggregate price change of a group of other items reflects change in the missing item – an implicit quality adjustment to price.
- A replacement item may be selected and the replacement item's price may be used for the comparison because the replacement is deemed to be comparable in quality to the missing item.
- The replacement may be deemed to be non-comparable with the missing item, but

prices on both the missing and replacement items may be available in an overlap period before the former item was missing. The price difference in this overlap period may be used as an estimate of the quality difference to quality-adjust the replacement item's price.

- The replacement price of a non-comparable replacement may be used, with an explicit estimate of the adjustment for the quality difference to extricate the "pure" price and quality change.

7.11 In many cases, therefore, there is a need to make a quality adjustment to the replacement item's price. A quality adjustment in this instance is an adjustment to the price (price change) of the replacement item (compared with the missing item) to remove that part of the price change that results from quality differences. A quality adjustment can be taken to be a coefficient that multiplies the price of, say, the replacement item to make it commensurate, from the consumer's point of view, with the price of the original.

7.12 To take a simple example, suppose that the size (or quantity) an item is sold in is a quality feature. Suppose that the size of the missing item and its replacement differ. Assume that a quantity k of the replacement is sold for the same price as a quantity j of the original. Whether the consumer buys one unit of the original or j/k units of the replacement makes no difference – they are worth the same. In order to make the price of one unit of the replacement commensurate with the price of one unit of the original, the replacement must be multiplied by k/j. This is the required quality adjustment. For example, if 2 units of the replacement item were equivalent to 3 of the original, the required quality adjustment to be applied to the price of the replacement item is 2/3. Suppose one unit of the replacement actually sells at the same price as one unit of the original, then the price of the replacement, after adjusting for the change in quality, is only 2/3 that of the price of the original. If one unit of the replacement sells for twice the price of the original, then the quality-adjusted price is 4/3 that of the original: the price increase is 33 per cent, not 100 per cent. The consumer price index seeks to record the change between the price of the original and the quality-adjusted price of the replacement.

7.13 The approaches listed in paragraph 7.10 will be discussed later in some detail, along with the assumptions implied by them. By definition, the prices of the unavailable items cannot be determined. The veracity of some of the assumptions about their price changes, had they been available, is therefore difficult to establish. What is stressed here is that the matching of prices of items allows for the measurement of price changes untainted by quality changes. When items are replaced with new ones of a different quality, then a quality-adjusted price is required. If the adjustment is inappropriate, there is an error, and if it is inappropriate in a systematic direction, there is a bias. Careful quality adjustment practices are required to avoid error and bias. Such adjustments are the subject of this chapter.

Sampling concerns

7.14 There are four main concerns with regard to sampling. First, the matching of prices of identical items over time, by its nature, is likely to lead to the monitoring of a sample of items increasingly unrepresentative of the population of transactions. 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. Ignoring such “unmatched” models in measuring a consumer price index will bias the index downwards (see paragraphs 7.150 to 7.152 below). Therefore, in a curious way, the very method of matching, used to ensure constant quality, may itself lead to bias by omitting items whose price changes are unusual (see also Koskimäki and Vartia (2001) for an example). Chapter 8 suggests that the strategy for quality adjustment of prices should be linked to one of item selection and chaining. The strategy is particularly pertinent to sectors with dynamic technological innovations (see also the discussion of hedonic price indices, below).

7.15 Second, because of the additional resources required for quality adjustments to prices, it may be in the interests of the price collectors and desk statisticians, and indeed fall within their guidelines, to avoid making non-comparable replacements and quality adjustments. Thus items continue to be monitored until they are no longer produced. This means that old items with limited sales are monitored. Such items may exhibit unusual price changes as they near the end of their life cycle, because of the marketing strategies of firms. Firms typically identify gains to be made from different pricing strategies at different times in the life cycle of products, particularly at the introduction and end of the cycle (Parker, 1992). The (implicit or otherwise) weight of end-of-cycle items in the index would thus remain relatively high, being based on their sales share when they were sampled. Furthermore, new unmatched items with possibly relatively large sales would be ignored. As a consequence, undue weight would be given to the unusual price changes of matched items at the end of their life cycle.

7.16 A third sampling concern relates to the timing of item substitution: when a replacement item is chosen to substitute for an old one. Instructions to pick a comparable replacement to avoid messy quality adjustments to prices compound the problem. Obsolete items are by their nature at the end of their cycles and comparable replacements, to be comparable, must also be near or at the end of their cycles. Obsolete items with unusual price changes at the end of their cycles are thus replaced by obsolete items with, again, unusual price changes. This compounds the problem of unrepresentative samples and continues to bias the index against technically superior items delivering cheaper service flows.

7.17 The final sampling problem with the matching procedure is when the price collector continues to report prices of items until replacements are forced, that is, until the items are no longer available, and has instructions to replace those items with typically consumed or popular items. This improves the coverage and representativity of the sample. But it also makes reliable quality adjustments of prices between the old obsolete and new popular items more difficult. The differences in quality are likely to be beyond those that can be attributed to price differences in some overlap period, as one item is in the last stages of its life cycle and the other in its first. Furthermore, the technical differences between the items are likely to be of an order that makes it more difficult to provide reliable, explicit estimates of the effect of quality differences on prices. Finally, the (quality-adjusted) price changes of very old and very new items are unlikely to meet assumptions of “similar price changes to existing items or classes of items”, as required by the imputation methods. Many of the methods of dealing with quality adjustment for unavailable items may be better served if the switch to a replacement item is made earlier rather than later. Sampling concerns can be seen to be inextricably linked to quality adjustment methods. This will be taken up in Chapter 8 on item selection and the need for an integrated approach to dealing with both representativity and quality-adjusted prices.

New products

7.18 A third potential source of error arises when something new is introduced into the marketplace. It is difficult to distinguish between new items and quality changes in old ones; this difficulty will be discussed in Chapter 8. When a really new item is introduced, there is an immediate gain in welfare or utility as demand switches from the previous technology and other goods. For example, the introduction of the zip fastener for clothing, instead of buttons, was a completely new good that led to an initial gain in utility or welfare to consumers as they switched from the old to the new technology. This gain from its introduction would not be properly brought into the index by waiting until the index was rebased, or by waiting for at least two successive periods of prices for zip fasteners and linking the new price comparison to the old index. Subsequent prices might be constant or even fall. The initial welfare gain would be calculated from a comparison between the price in the period of introduction and the hypothetical price in the *preceding* period, during which supply would be zero. The practical tools for estimating such a hypothetical price are not well developed, though this subject is discussed in more detail in Chapter 21. For a consumer price index built on the concept of a base period and a fixed basket, there is, strictly speaking, no problem. The new good was not in the old basket and should be excluded. Although an index properly measuring an old fixed basket would be appropriate in a definitional sense, it would not be representative of what we buy. Such an index would thus be inappropriate. For a cost of living index concerned with measuring the change in expenditure necessary to maintain a constant level of utility (see Chapter 17), there is no doubt that it would be conceptually appropriate to include the new good.

The nature of quality change

7.19 This section considers what is meant by quality change and then outlines the methods available for dealing with unavailable price quotes. To understand the “meaning” of quality change requires a conceptual and theoretical platform, so that adjustments to prices for quality differences are made against a well-considered framework.

7.20 A starting point is to appreciate that over time the quality of what is produced changes. The example of new cars is used here. Bode and van Dalen (2001) undertook an extensive study of the price measurement of new cars in the Netherlands between 1990 and 1999. The average nominal price increase over this period was found to be around 20 per cent, but the mix of average quality characteristics changed over this period. For example, the horsepower (HP) increased on average from 79 to 92 HP; the average efficiency of fuel consumption improved from 9.3 to 8.4 litres/100 km; the share of cars with fuel injection rose from 51 per cent to 91 per cent; the proportion of cars with power steering increased from 27 per cent to 94 per cent; airbags from 6 per cent to 91 per cent, and similarly for central locking, tinted glass and much more. This churning in the quality mix of what is purchased is one aspect of quality change. In matching the prices of a sample of models in, for example, January with the self-same models in subsequent months, the quality mix is kept constant in an attempt to avoid contaminating the price measurement through quality differences. As will be seen later, however, the resulting sample of models is one that gives less emphasis to models subsequently introduced which may have benefited from more recent technological change and have different price changes given the quality of services they provide. One approach, which corrects for such quality changes but uses the whole sample, is that of the dummy variable hedonic regressions (see below). Bode and van Dalen (2001), using a variety of formulations of hedonic regressions, found quality-corrected prices of these new automobiles to be about constant over this period, while their average nominal price increase was around 20 per cent.

7.21 It will be argued in Chapter 21 that observed changes in prices arise in theory from a number of sources, including quality changes, changes in tastes and preferences, and changes in the technology of producers. More formally, the observed data on prices are the locus of intersection of the demand curves of different consumers with varying tastes and the supply curves of different producers with possibly varying technologies of production. The separation of the effects of changes in tastes and preferences from quality changes is only possible in highly restrictive circumstances. Chapter 8 suggests chaining or regular rebasing, so that weights – which reflect tastes and preferences – are not unduly out of date.

7.22 The changing mix of the observed characteristics of items is not the only concern. There is also the practical problem of not always being able to observe or quantify characteristics such as the style, reliability, ease of use and safety of what is produced. Chapter 16 of the *System of National Accounts, 1993* (*SNA 1993*) on price and volume measurement notes factors other than changes in physical characteristics that give rise to improved quality. These include "transporting a good to a location in which it is in greater demand is a process of production in its own right in which the good is transformed into a higher quality good". The same good provided at a different and more convenient location may command a higher price and be of a higher quality. Furthermore, different times of the day or periods of the year may also give rise to quality differences: "For example, electricity or transport provided at peak times must be treated as being of higher quality than the same amount of electricity or transport provided at off-peak times. The fact that peaks exist shows that purchasers or users attach greater utility to the services at these times, while the marginal costs of production are usually higher at peak times … ." Other differences, including the conditions of sale and circumstances or environment in which the goods or services are supplied or delivered, can make an important contribution to differences in quality. A retailer, for example, may attract customers by providing free delivery, credit opportunity or better variety, by being more accessible, by offering shorter order times, smaller tailor-made orders, clearer labelling, better support and advice, more convenient car parking or a wider range of brands, or simply by operating in a more pleasant or fashionable environment. These sorts of benefits are not always specified in the item description because, first, the services are provided without specific charge – they are incorporated in the prices of the goods sold. Second, by matching the prices of models in specific outlets the level of such services is assumed to remain constant. This does not mean, however, that conceptually such quality improvements should be outside the scope of the index. If any such benefits change, a price adjustment for the estimated value of the benefits should be made.

7.23 To ask how to adjust prices for quality changes, it is first necessary to ask what is meant by quality. While there may be an intuition as to whether an item consumed in one period is better than its counterpart in the next, a theoretical framework will help in establishing the basis for such comparisons. For example, an item of clothing is sampled and, after a few months, it is missing. One option is to replace it with a similar item. The nearest comparable option may have more cloth in it, or have a lining, be a different colour, have different buttons, have better stitching or be considered to be better styled in some fashionable sense. There is a need to put a price estimate on the difference in quality between the old and new items so that like can be compared with like. To propose or criticize a quality adjustment procedure requires some concept of what is ideally required and how the procedure stands up to this. Although such a discussion takes us away from the practicalities of the procedures for a while, its use will become apparent in subsequent sections.

A utility-based approach

7.24 In Chapter 17 a cost of living index (COLI) is defined as the ratio of the minimum expenditures in the base and current period required to achieve a given standard of living or "utility". Quality adjustments to prices involve trying to measure the price change for a product which has exhibited some change in its characteristics from an earlier period that provides a different level of utility to the consumer. The equating of the value of a quality change with the change in utility derived by the consumer, while falling naturally under a COLI framework, is not exclusive to it. A cost of a fixed basket of goods index (COGI) can also benefit from regarding quality in this way. While a COGI requires the pricing of a fixed basket of products, some items will become unavailable and the replacement items selected to maintain the sample may not be of the same quality. The aim is to determine what proportion of the total price change results from a change in quality and what results from pure price change. The concept of utility will be used to help with the former.

7.25 Note that the definition of a quality change is based on equating some change in characteristics to a different level of utility provided. Consider an example in which a new, improved quality item is substituted for an old one in period *t*, the consumer having to choose between the two. Suppose that after the new quality item appeared, both qualities were offered to a consumer at the same price, say p^t = 100. The consumer was then asked to choose between them and naturally preferred the new quality. Say the price of the old quality was then progressively reduced until it reached a point, say p^{t*} = 75, at which the consumer was indifferent as regards the choice between purchasing the old quality at p^{t*} = 75 and the new quality at p^t = 100. The consumer might then select the old quality at 75 or the new one at 100. Either way, the consumer would obtain the same utility, because of being indifferent as to which to choose. Any further decrease below p^{t*} = 75 would cause the consumer to switch back to the old quality.

7.26 The difference between p^t and p^{t*} would be a measure of the additional utility that the consumer placed on the new quality as compared with the old quality. It would measure the maximum amount that the consumer was prepared to pay for the new quality over and above the price of the old quality. In economic theory, as will be outlined in Chapter 21, if consumers (or households) are indifferent between two purchases, the utility derived from them is the same. The difference between 75 and 100 must therefore arise from the consumers' valuation of the utility they derive from the two items: their quality difference. The definition is sensible as a conceptual framework. It naturally has problems relating to implementation, but this is not our concern here. Our initial concern is with the provision of an analytical framework on which to ground our thinking and analysis.

7.27 The utility-based framework is concerned with the question of how consumers choose between items of different qualities. The answer, in part, is because more utility is derived from an item of higher quality than from an item of lower quality, and thus consumers prefer it. But this does not explain why one item is bought rather than the other. For this it is also necessary to know the relative price of one item with respect to the other, since if the lower-quality item is cheaper, it may still be purchased. The above thought experiment to determine the price below which the old quality would be purchased, $p^{t*} \leq 75$, serves this purpose.

7.28 Defining quality change in terms of its effect on utility is of obvious benefit to the economic approach to index numbers (Chapter 21). Fixler and Zieschang (1992), Feenstra

(1995), Triplett (1987) and Diewert (2003a) have developed theoretical frameworks for COLIs akin to those defined in Chapter 21, but which also incorporate goods and services whose quality changes. Silver and Heravi (2001a and 2003) and Kokoski et al. (1999) have undertaken empirical studies based on these frameworks for comparisons over time and between geographical areas, respectively. The use of utility as a guide towards understanding quality adjustments to prices is not, however, confined to the economic theory of cost of living indices (Chapter 21). Consumer price indices based on a fixed basket concept have the pragmatic need to adjust for quality differences when an item is unavailable, and there is nothing in the definition of a fixed basket index that precludes differences in utility being used as a guideline. If item A is better than its old version, item B, it is because it delivers something more to the consumer who is willing to pay more. That “thing” is called utility.

7.29 It is as well to distinguish between two concepts of value used in the analysis of quality adjustment: *resource cost* and *user value.* The value users derive from their consumption is their utility. Triplett (1990, pp. 222-223) considers how a consumer price index differs from a producer price index:

> Fisher and Shell (1972) were the first to show that different index number measurements (they considered output price indexes and consumer price indexes) imply alternative treatments of quality change, and that the theoretically appropriate treatments of quality change for these two indexes correspond respectively, to “resource-cost” and “user-value” measures. Triplett (1983) derives this same result for cases where “quality change” is identified with characteristics of goods – and therefore with empirical hedonic methods; the conclusions are that the resource cost of a characteristic is the appropriate quality adjustment for the output price index, and its user value is the quality adjustment for the COLI index or input index.

7.30 This position is not without difficulties. Diewert (2002d) has advocated a user cost approach for the producer price output index. This in part arises from the need to consolidate the inputs and outputs at constant prices in national accounts. If different quality adjustments are used for the same items in the producer price *input* index and the producer price *output* index, then the deflated constant price value added series, as their difference, will not balance. The issue arises generally in the field of producer price indices, since it concerns the question of whether the producer price output index should use a user-value concept. It does not dispute the use of this concept in consumer price indices.

Conditional indices

7.31 The domain of a COGI is its fixed basket of goods and services. The use of a COLI framework requires consideration of wider issues concerning our quality of life. There are changes in the social, physical and economic environment that require more or less expenditure to maintain a given level of utility. Many factors affect our welfare, and in practice not all can be included in a consumer price index. It is thus appropriate to consider indices that are *conditional* on excluded factors remaining constant. These generally include health status, the environment and the quantity and quality of government-provided goods and services. The minimum expenditure necessary for achieving a given level of utility will increase as, for example, the police become less effective. Expenditure would then be necessary for better household security. It would cost more to maintain a given level of utility than in the previous period. Similarly, an outbreak of illness would lead to increased expenditure on medicines to maintain a given level of utility. Bad winter weather increases heating bills to maintain the same utility as before. In each case there is a very real sense in which the cost of living will have changed. Yet it is not generally accepted that the consumer price index should directly reflect such changes. What should be reflected are changes in the prices of locks, medicines and fuel that arise because the demand for such items changes. In addition, as more or less is spent on such items, the index should eventually incorporate such changes in the weighting as and when the weights are updated – and the more frequent the update, the better are such effects incorporated. But the index should not normally reflect short-run changes in the *quantities* used of security, medicine, heat and the like as a result of such external factors. Gordon and Griliches (1997, p. 87) comment in a similar vein:

> It is not clear, moreover, whether events such as a colder winter, the appearance of AIDS, or a rise in the crime rate should be included in the definition of a *price* index. A change in expenditures due to an unanticipated change in the weather should raise the price index only to the extent that energy prices go up, not quantities consumed. If the event persists, ultimately it will affect the commodity weights in the index, but that is a different matter. (Authors' emphasis)

7.32 It may be inappropriate to disregard environmental factors if they seriously affect a given group of people. In such cases, indexing for special factors sometimes takes place outside the index. For example, a government may provide cold-weather payments to pensioners if the temperature falls below a threshold condition. If a specific factor has a substantial effect for a significant group of households, an additional index might be compiled which includes the effect.

An overview of methods of quality adjustment when matched items are unavailable

7.33 It is apparent from the above that quality adjustments to prices are not a simple matter of applying routine methods to prices in specified product areas. A number of alternative approaches are suggested below. Some will be more appropriate than others for specific product areas. An understanding of the consumer market, technological features of the producing industry, and alternative data sources will all be required for the successful implementation of quality adjustments. Specific attention will need to be devoted to product areas with relatively high weights, where large proportions of items are turned over. Some of the methods are not straightforward and require a level of expertise. Quality adjustment needs to be implemented by developing a gradual approach on a product-by-product basis. Such concerns should not be used as excuses for failing to attempt to estimate quality-adjusted prices. The practice of statistical agencies in dealing with missing items, even if it is to ignore them, implicitly involves a quality adjustment. Such an implicit approach may not be the most appropriate method, and may even be misleading. The extent of quality changes and the pace of technological change require that appropriate methods be used.

7.34 To measure aggregate price changes, a representative sample of items is selected from a sample of outlets, along with a host of details that define each price. The items are repriced each month. The detailed specifications are included on the repricing form each month as a prompt to help ensure that the same items are being priced. Merkel (2000) has proposed that detailed checklists of item descriptions should be used, as any lack of clarity in the specifications may lead to errors. It should be borne in mind that price collectors may have no incentive to report changes in specifications, since this will invariably involve additional work. Attention should also be devoted to ensuring that the specifications used contain all pertinent, price-determining elements, otherwise there may be cases in which the quality change would become invisible in the price measurement process.

7.35 When an item is missing in a month for reasons other than being off season or off cycle, the replacement may be of a different quality – like may no longer be compared with like. A number of approaches exist for dealing with such situations and are well documented for the consumer price index (CPI), as outlined in Turvey et al. (1989), Moulton and Moses (1997), Armknecht et al. (1997), Moulton et al. (1999) and Triplett (2002). Though the terminology differs between authors and statistical agencies, they include:

- imputation – where no information is available to allow reasonable estimates to be made of the effect on price of a quality change. The price changes of all items, or of more or less similar items, are assumed to be the same as that for the missing item;
- overlap – used where no information is available to allow reasonable estimates to be made of the effect on price of a quality change, but where a replacement item exists in the same period as the old item. The price difference between the old item and its replacement in the overlap period is then used as a measure of the quality difference;
- direct comparison – if another item is directly comparable, that is, it is so similar it can be assumed to have had more or less the same quality characteristics as the missing one, its price replaces the unavailable price. Any difference in price level between the new and old is assumed to arise from price changes and not quality differences;
- explicit quality adjustment – where there is a substantial difference between the quality of the old and replacement items, estimates of the effect of quality differences on prices are made to enable quality-adjusted price comparisons to be made.

7.36 Before outlining and evaluating these methods it is as well to say something about the extent of the problem. This arises when the item is unavailable. It is not just a problem when *comparable* items are unavailable, for the judgement as to what is and what is not comparable itself requires an estimate of quality differences. Part of a statistical meta-information system for statistical offices (outlined in Chapter 8) is to identify and monitor sectors that are prone to such replacements and whether the replacements used really are comparable. Seminal studies in Canada and the United States throw some light on the extent of such replacements. Moulton et al. (1999) examined the extent to which items became unavailable for televisions in the compilation of the United States CPI. Between 1993 and 1997, a total of 10,553 prices on televisions were used, of which 1,614 (15 per cent) were replacements, of which, in turn, 934 (57 per cent) were judged to be directly comparable. Thus a typical television remained in the sample less than a year. The Canadian experience for televisions over an almost identical period (1993 to November 1997) found 750 of the 10,050 prices (7.5 per cent) to be replacements. Of these, 178 (24 per cent) were directly comparable, 162 (22 per cent) required a judgement and 410 (55 per cent) were “spliced” – the price difference between the replacement and the unavailable model in the two periods being attributed to quality differences (Lowe, 1999). Thus, there was wide variation in the frequency of total replacements, although the frequency of non-comparable replacements was roughly similar (6.4 per cent in the United States sample and 5.7 per cent in Canada). Liegey (2000) found that of the 215 average (August 1999 to April 2000) monthly prices collected for major appliances for the United States CPI, 22 item replacements were required because of missing prices, of which comparable replacements were found for 16 and non-comparable replacements for the remaining six.

7.37 Information across a wider range of items is available for the United States. Armknecht (1996) found that, over the three years 1993 to 1995, the annual average number of price observations collected for the United States CPI was 835,443, of which 59,385 (7.1 per cent) were substitutions (as opposed to imputations for missing values). Of these substitutes, about half were carried out using comparable replacements, under a quarter using overall mean imputation, about 12 per cent using direct quality adjustment, and 10 per cent using class mean imputation. It should be borne in mind that these figures ignore the implicit quality adjustments that take place when the Bureau of Labor Statistics rotates its sample between rebasing. The *overlap* method is effectively applied on sample rotation, the outlet and item samples being reselected for about one-fifth of the geographical areas, with prices of old and new items sampled in the same month. All price-level differences between the old and new items are treated as quality differences as the new sample is spliced onto the old.

7.38 Methods of quality adjustment for prices are generally classified into implicit/imputed (or indirect) quality adjustment methods – the differences in terminology are notorious in this area – and explicit (or direct) methods. Implicit and explicit methods are discussed below. Both decompose the price change between the old item and its replacement into quality and pure price changes. For explicit adjustments, however, an explicit estimate is made of the quality difference, usually on the basis of external information, and the pure price effect is identified as a remainder. For implicit adjustments, a measurement technique is used to compare the old item to the replacement item, in which the extent of the quality and pure price change is implicitly determined by the assumptions of the method. The accuracy of the method relies on the veracity of the assumptions as opposed to the quality of the explicit estimate. Explicit adjustments make use of separate estimates of the portion of prices ascribed to quality differences, so that the price of the original item can be compared with that of a replacement of the same quality. The suitability of the explicit methods thus depends to a large extent on how good such estimates are, on average. Implicit adjustments involve assumptions about price movements, and for these informed intuition or theory is relied upon – though in some cases national statistical offices may make use of more specific empirical market knowledge.

Additive versus multiplicative adjustment

7.39 The quality adjustments to prices may be undertaken either by adding a fixed amount or multiplication by a ratio. For example, where *m* is the old item and *n* its replacement for a comparison over periods t, $t + 1$, $t + 2$, the use of the overlap method in period $t + 1$ requires the ratio p_n^{t+1} / p_m^{t+1} to be used as a measure of the relative quality difference between the old item and its replacement. This ratio could then be *multiplied* by the price of the old item in period t, p_m^t to obtain the quality-adjusted prices p_m^{*t} as follows:

	t	$t + 1$	$t + 2$
old item *m*		p_m^{t+1}	
replacement *n*	p_m^{*t}	p_n^{t+1}	p_n^{t+2}

7.40 Such multiplicative formulations are generally advised, as the adjustment is invariant to the absolute value of the price. It would otherwise be possible for the absolute value of the change in specifications to exceed the value of the item in some earlier or (with technological advances) later period. Yet there may be some items for which the worth of the constituent parts is not considered to be in proportion to the price. In other words, the constituent parts have their own, intrinsic, absolute, additive worth, which remains constant over time. Producers selling over the World Wide Web may, for example, include postage, which in some instances may remain the same irrespective of what is happening to price. If postage is subsequently excluded from the price, this fall in quality should be valued as a fixed sum.

Base versus current period adjustment

7.41 Two variants of the approaches to quality adjustment are to make the adjustment either to the price in the base period or to the price in the current period. For example, in the overlap method, described above, the implicit quality adjustment coefficient was used to adjust p_m^t. An alternative procedure would have been to multiply the ratio p_m^{t+1} / p_n^{t+1} by the price of the replacement item p_n^{t+2} to obtain the quality-adjusted price p_n^{*t+2}, etc. The first approach is easier since, once the base period price has been adjusted, no subsequent adjustments are required. Each new replacement price can be compared with the adjusted base period price. For multiplicative adjustments, the end result is the same whichever approach is used. For additive adjustments, the results differ and it is more appropriate to make the adjustment to prices near to the overlap period.

Long-run versus short-run comparisons

7.42 Much of the analysis of quality adjustments in this manual has been undertaken by comparing prices between two periods, say, period 0 prices with those in a subsequent period 1. For long-run comparisons the base period is taken as, say, period *t* and the index is compiled by comparing prices in period *t* first with *t* + 1; then *t* with *t* + 2; then *t* with *t* + 3, etc. The short-run framework allows long-run comparisons, say, between periods *t* and *t* + 3, to be built up as a sequence of links joined together by successive multiplication, say period *t* with *t* + 2 and period *t* + 2 with *t* + 3; or with chaining, period *t* with *t* + 1, *t* + 1 with *t* + 2 and *t* + 2 with *t* + 3. The advantages of the short-run framework for imputations are discussed in paragraphs 7.165 to 7.173.

7.43 Following a discussion of implicit and explicit methods of quality adjustment, issues relating to choice of method are considered. The implicit and explicit adjustment methods are outlined under a standard long-run Laspeyres framework, in which prices in a base (or reference) period are compared with those in each subsequent period. Where products are experiencing rapid technological change, however, these methods may be unsuitable. The matching and re-pricing of like items, and “patching in” of quality-adjusted replacement prices when the matching fails, are appropriate when failures are the exception. But in high-technology product markets likely to experience rapid turnover of models, they are the rule. Alternative methods using chained or hedonic frameworks are therefore also considered. These are quite radical approaches to meet the needs of rapidly changing production portfolios. Finally, the use of short-run comparisons as an alternative to long-run ones is considered as an intermediary – and for imputation a more appropriate – approach. Chapter 22 discusses issues relating to seasonal items in more detail.

Implicit methods of quality adjustment

7.44 This section discusses the following implicit methods of quality adjustment: overlap; overall mean or targeted mean imputation; class mean imputation; comparable replacement; linked to show no price change; and carry-forward.

Overlap

7.45 Consider for illustration the case where the items are sampled in, say, January and prices are compared over the remaining months of the year. Matched comparisons are undertaken between the January prices and their counterparts in successive months. Five items are assumed to exist in January sold by two outlet types with prices p^{11}, p^{21}, p^{51}, p^{61} and p^{81} (Table 7.1(a)). At this level of aggregation, the weights can be ignored assuming only one quote is taken on each item. A price index for February compared with January = 100.0 is straightforward, in that prices of items 1, 2, 5, 6 and 8 only are used and compared by way of the geometric mean of price ratios, the Jevons index (which is equivalent to the ratio of the geometric mean in February over the geometric mean in January – see Chapter 20). In March the prices for items 2 and 6 are missing, one from specialized chain stores and one from department stores.

Table 7.1 Example of the overlap method of quality adjustment

(a) General illustration

Outlet	Item	January	February	March	April
Specialized chain stores	1	p^{11}	p^{12}	p^{13}	p^{14}
	2	p^{21}	p^{22}		
	3			p^{33}	p^{34}
	4		p^{42}	p^{43}	p^{44}
Department stores	5	p^{51}	p^{52}	p^{53}	p^{54}
	6	p^{61}	p^{62}		
	7			p^{73}	p^{74}
	8	p^{81}	p^{82}	p^{83}	p^{84}

(b) Numerical illustration

Outlet	Item	January	February	March
Specialized chain stores	1	4	5	6
	2	5	6	
	2. overlap			**6.9**
	–imputation			**6.56**
	–targeted imputation			**7.2**
	–comparable replacement			**6.5**
	3			6.5
	4		7.5	8
Department stores	5	10	11	12
	6	12	12	
	–imputation			**13.13**
	–targeted imputation			**12.533**
	7			14
	8	10	10	10

7.46 Table 7.1(b) is a numerical counterpart to Table 7.1(a) to further illustrate the calculations. The overlap method requires prices of the old and replacement items to be available in the same period. In Table 7.1(a), in March item 2 has no price quote. Its new replacement is, say, item 4. The overlap method simply measures the ratio, in a common overlap period (February), of the prices of the old and replacement items (items 2 and 4, respectively). This is taken to be an indicator of their quality differences. The two approaches outlined above are apparent: either to insert a quality-adjusted price in January for item 4 and continue to use the replacement item 4 series, or to continue the item 2 series by patching in quality-adjusted item 4 prices. Both yield the same answer. Consider the former. For a Jevons geometric mean for January to March *for specialized chain stores only*, assuming equal weights of unity:

$$P_J(p^1, p^3) = \left[p^{13}/p^{11} \times p^{43}/\left(p^{42}/p^{22}\right) \times p^{21} \right]^{1/2}$$
$$= \left[6/4 \times 8/((7.5/6) \times 5)\right]^{1/2} = 1.386 \qquad (7.1)$$

Note that the comparisons are long-run ones. That is, they are between January and the month in question. The short-run modified Laspeyres framework provides a basis for short-run changes based on data in each current month and the immediately preceding one. In Tables 7.1(a) and (b) the comparison for specialized chain stores only would first be undertaken between January and February using items 1 and 2, and this would be multiplied by the comparison between February and March using items 1 and 4. This still implicitly uses the differences in prices in the overlap in February between items 2 and 4 as a measure of this quality difference. It yields the same result as before:

$$\left[\frac{5}{4} \times \frac{6}{5}\right]^{\frac{1}{2}} \times \left[\frac{6}{5} \times \frac{8}{7.5}\right]^{\frac{1}{2}} = 1.386$$

The advantage of recording price changes for, say, January to October in terms of January to September, and September to October, is that it allows the compiler to compare immediate month-on-month price changes for data editing purposes. Moreover, it has quite specific advantages for the use of imputations (as discussed in paragraphs 7.53 to 7.68 below) for which different results arise for the long- and short-run methods. The long-run and short-run frameworks are discussed more fully in paragraphs 7.159 to 7.173.

7.47 The method is only as good as the validity of its underlying assumptions. Consider $i = 1 \ldots m$ items where p_m^t is the price of item *m* in period *t*, p_n^{t+1} is the price of a replacement item *n* in period *t* + 1, and there are overlap prices for both items in period *t*. Now item *n* replaces *m*, but is of a different quality. So let *A*(*z*) be the quality adjustment to p_n^{t+1} which equates its quality to p_m^{t+1} such that the quality-adjusted price $p_m^{*t+1} = A\left(z^{t+1}\right) p_n^{t+1}$. Very simply, the index for the item in question over the period *t*-1 to *t* + 1 is:

$$I^{t-1,t+1} = \left(p_m^t / p_m^{t-1}\right) \times \left(p_n^{t+1} / p_n^t\right) = \frac{p_n^{t+1}}{p_m^{t-1}} \times \frac{p_m^t}{p_n^t} \qquad (7.2)$$

7.48 Now the quality adjustment to prices in period *t* + 1 is defined as previously, $p_m^{*t+1} = A\left(z^{t+1}\right) p_n^{t+1}$, which is the adjustment to p_n in period *t* + 1 which equates its utility to p_m in period *t* + 1 (had it existed then). A desired measure of price changes between periods *t*-1 and *t* + 1 is thus:

$$\left(p_m^{*t+1} / p_m^{t-1}\right) \qquad (7.3)$$

The overlap formulation equals this when:

$$\frac{p^{*t+1}}{p_m^{t-1}} = A\left(z^{t+1}\right) \frac{p_n^{t+1}}{p_m^{t-1}} = \frac{p_n^{t+1}}{p_n^t} \times \frac{p_m^t}{p_m^{t-1}}$$

$$A\left(z^{t+1}\right) = \frac{p_m^t}{p_n^t}$$ and similarly for future periods of the series

$$A\left(z^{t+i}\right) = \frac{p_m^t}{p_n^t} \text{ for } \frac{p_m^{*t+i}}{p_m^{t-1}} \quad \text{for } i = 2, \ldots T \qquad (7.4)$$

The assumption is that the quality difference in any period equates to the price difference at the *time of the splice*. The *timing* of the switch from *m* to *n* is thus crucial. Unfortunately, price collectors usually hang onto an item so that the switch may take place at an unusual period of pricing, near the end of item *m*'s life cycle and the start of item *n*'s life cycle.

7.49 But what if the assumption does not hold? What if the relative prices in period *t*, $R^t = p_m^t / p_n^t$, do not equal *A(z)* in some future period, say $A(z^{t+i}) = \alpha_i R^t$? If $\alpha_i = \alpha$, the comparisons of prices between future successive periods, say between *t* + 3 and *t* + 4, are unaffected, as would be expected, since item *n* is effectively being compared with itself,

$$\frac{p_m^{*t+4}}{p_m^{t-1}} / \frac{p_m^{*t+3}}{p_m^{t-1}} = \frac{\alpha R^t}{\alpha R^t} \frac{p_n^{t+4}}{p_n^{t+3}} = \frac{p_n^{t+4}}{p_n^{t+3}} \qquad (7.5)$$

However, if differences in the relative prices of the old and replacement items vary over time, then:

$$\frac{p_m^{*t+4}}{p_m^{t-1}} / \frac{p_m^{*t+3}}{p_m^{t-1}} = \frac{\alpha_4}{\alpha_3} \frac{p_n^{t+4}}{p_n^{t+3}} \qquad (7.6)$$

Note that the quality difference here is not related to the technical specifications or resource costs, but to the relative prices consumers pay.

7.50 Relative prices may also reflect unusual pricing policies aimed at minority segments of the market. In the example of pharmaceutical drugs (Berndt et al., 2003), the overlap in prices of a generic with a branded product was argued to reflect the needs of two different market segments. The overlap method can be used with a judicious choice of the overlap period. It should if possible be a period before the use of the replacement since in such periods the pricing may reflect a strategy to dump the old model to make way for the new one.

7.51 The overlap method is implicitly employed when samples of items are rotated. That is, the old sample of items is used to compute the category index price change between periods *t*-1 and *t,* and the new sample is used between *t* and *t* + 1. The "splicing" together of these index movements is justified by the assumption that – on a group-to-group rather than item-to-item level – differences in price levels at a common point in time accurately reflect differences in qualities.

7.52 The overlap method has at its roots a basis in the law of one price: that when a price difference is observed it must arise from some difference in physical quality or some such factors for which consumers are willing to pay a premium, such as the timing of the sale, location, convenience or conditions. Economic theory would dictate that such price differences would not persist, given markets made up of rational producers and consumers. However, Chapter 16 of *SNA 1993* notes three reasons why this might fail:

> First, purchasers may not be properly informed about existing price differences and may therefore inadvertently buy at higher prices. While they may be expected to search out for the lowest prices, costs are incurred in the process.
> Secondly, purchasers may not be free to choose the price at which they purchase because the seller may be in a position to charge different prices to different categories of purchasers for identical goods and services sold under exactly the same circumstances – in other words, to practise price discrimination.
> Thirdly, buyers may be unable to buy as much as they would like at a lower price because there is insufficient supply available at that price. This situation typically occurs when there are two parallel markets. There may be a primary, or official, market in which the quantities sold, and the prices at which they are sold, are subject to government or official control, while there may be a secondary

market – a free market or unofficial market – whose existence may or may not be recognized officially.

Overall mean or targeted mean imputation

7.53 This method uses the price changes of other items as estimates of the price changes of the missing items. Consider a Jevons elementary price index, i.e., a geometric mean of price relatives (Chapter 20). The prices of the missing items in the current period, say $t + 1$, are imputed by multiplying their prices in the immediately preceding period t by the geometric mean of the price relatives of the remaining matched items between these two periods. The comparison is then linked by multiplication to the price changes for previous periods. It is the computationally most straightforward of methods since the estimate can be undertaken by simply dropping the items that are missing from both periods from the calculation. In practice, the series is continued by including in the database the imputed prices. It is based on the assumption of similar price movements. A targeted form of the method would use similar price movements of a cell or elementary aggregate of similar items, or be based on price changes at a higher level of aggregation if either the lower level had an insufficient sample size or price changes at the higher level were judged to be more representative of the price changes of the missing item.

7.54 In the example in Table 7.1, the January to February comparison for both outlet types is based on items 1, 2, 5, 6 and 8. For March compared with January – weights all equal to unity – the item 2 and item 6 prices are imputed using the short-run price change for February (p^2) compared with March (p^3) based on items 1, 5 and 8. Since different formulae are used for elementary aggregation, the calculations for the main three formulae are illustrated here (but see Chapter 20 on choice of formulae). The geometric mean of the price ratios – the Jevons index – is:

$$\begin{aligned} P_J(p^2, p^3) &= \prod_{i=1}^{N} \left[p_i^3 / p_i^2 \right]^{1/N} \\ &= \left[\left(p^{13}/p^{12}\right) \times \left(p^{53}/p^{52}\right) \times \left(p^{83}/p^{82}\right) \right]^{1/3} \\ &= \left[(6/5) \times (12/11) \times (10/10) \right]^{1/3} \end{aligned} \qquad (7.7)$$

= 1.0939, or a 9.39 per cent increase.

The ratio of the average (mean) prices – the Dutot index – is:

$$\begin{aligned} P_D(p^2, p^3) &= \left(\sum_{i=1}^{N} p_i^3 / N\right) / \sum_{i=1}^{N} p_i^2 / N) \\ &= \left[\left(p^{13} + p^{53} + p^{83}\right)/3 \div \left(p^{12} + p^{52} + p^{82}\right)/3 \right] \end{aligned} \qquad (7.8)$$

= (6+12+10)/(5+11+10) = 1.0769, or a 7.69 per cent increase.

The average (mean) of the price ratios – the Carli index – is:

$$\begin{aligned} P_C(p^2, p^3) &= \sum_{n=1}^{N} (p_n^3 / p_n^2) / N \\ &= \left[\left(p^{13}/p^{12}\right) + \left(p^{53}/p^{52}\right) + \left(p^{83}/p^{82}\right) \right] / 3 \end{aligned} \qquad (7.9)$$

= [(6/5+12/11+10/10)]/3 = 1.09697, or a 9.697 per cent increase.

In practice, the imputed figure would be entered on the data sheet. In Table 7.1(b) the overall mean imputations in March for items 2 and 6, using the Jevons index, are $1.0939 \times 6 = 6.563$ and $1.0939 \times 12 = 13.127$, respectively: these are shown in bold. It should be noted that the Dutot index is in this instance lower that the Jevons index, a result not expected from the relationship established in Chapter 20. The relationship in Chapter 20 assumed that the variance of prices would increase over time, while in Table 7.1(b) it decreases for the three items. The arithmetic mean of relatives, the Carli index, equally weights each price change while the ratio of arithmetic means, the Dutot index, weights price changes according to the prices of the item in the base period relative to the sum of the base period prices. Item 1 has a relatively low price (4), and thus weight, in the base period, but this item has the highest price increase (6/5). The Dutot index is thus lower than the Carli index.

7.55 As noted above, it is also possible to refine the imputation method by "targeting" the imputation: by including the weight for the unavailable items in groupings likely to experience similar price changes, say by outlet type, specific product area or geographical region. Any stratification system used in the selection of outlets would facilitate this. For example, in Table 7.1 assume that the price change of the missing item 2 in March is more likely to follow price changes of item 1 in specialized chain stores, and item 6 is more likely to experience similar price changes to those of items 5 and 8 in department stores. For March compared with February, and weights all equal to unity, the geometric mean of the price ratios – the Jevons index – is:

$$\begin{aligned} P_J(p^2, p^3) &= \prod_{i=1}^{N} (p_i^3 / p_i^2)^{1/N} \\ &= \left[\left(p^{13} / p^{12}\right)^2 \times \left(p^{53} / p^{52} \times p^{83} / p^{82}\right)^{3/2} \right]^{1/5} \\ &= \left[(6/5)^2 \times (12/11 \times 10/10)^{3/2} \right]^{1/5} = 1.1041. \end{aligned} \tag{7.10}$$

Note the weights used: for specialized chain stores the one price represents two prices, while for department stores the two prices represent three prices, or 3/2 = 1.5 each.

The ratio of the average (mean) prices – the Dutot index – is:

$$\begin{aligned} P_D(p^2, p^3) &= \left(\sum_{i=1}^{N} p_i^3 / N\right) / \sum_{i=1}^{N} p_i^2 / N) \\ &= \left[\left(2p^{13} + 1.5p^{53} + 1.5p^{83}\right)/5\right] \div \left[\left(2p^{12} + 1.5p^{52} + 1.5p^{82}\right)/5\right] \\ &= \left[(2 \times 6 + 1.5 \times 12 + 1.5 \times 10)\right] \div \left[(2 \times 5 + 1.5 \times 11 + 1.5 \times 10)\right] = 1.0843 \end{aligned} \tag{7.11}$$

The average (mean) of the price ratios – the Carli index – is:

$$\begin{aligned} P_C(p^2, p^3) &= \sum_{i=1}^{N} (p_i^3 / p_i^2) / N \\ &= \frac{2}{5}\left(p^{13} / p^{12}\right) + \frac{3}{5}\left[\left(p^{53} / p^{52} + p^{83} / p^{82}\right)/2\right] \\ &= \frac{2}{5}(6/5) + \frac{3}{5}\left[(12/11 + 10/10)/2\right] = 1.1073 \end{aligned} \tag{7.12}$$

7.56 Alternatively, and more simply, imputed figures could be entered in Table 7.1(b) for items 2 and 6 in March, just using specialized chain stores and department store price movements for items 2 and 6 respectively, and indices calculated accordingly. Using a Jevons index, for item 2 the imputed value in March would be $6/5 \times 6 = 7.2$ and for item 6 it would be $\left[(12/11) \times (10/10)\right]^{1/2} = 12.533$. It is thus apparent that not only does the choice of formula matter, as discussed in Chapter 20, but so too may the targeting of the imputation. In practice, the sample of items in a targeted subgroup may be too small. An appropriate stratum is required with a sufficiently large sample size, but there may be a trade-off between the efficiency gains from the larger sample and the representativity of price changes achieved by that sample. Stratification by product area and region may be preferred to stratification just by product area, if regional differences in price changes are expected, but the resulting sample size may be too small. In general, the stratum used for the target should be based on the analyst's knowledge of the market, as well as an understanding of similarities of price changes between and within strata, and the reliability of the sample available to be representative of price changes.

7.57 The underlying assumptions of these methods require some analysis since, as discussed by Triplett (1999 and 2002), they are often misunderstood. Consider $i=1...m$ items where, as before, p_m^t is the price of item *m* in period *t*, p_n^{t+1} is the price of a replacement item *n* in period *t* + 1. Now *n* replaces *m*, but is of a different quality. So, as before, let *A*(*z*) be the quality adjustment to p_n^{t+1} which equates its quality services or utility to p_m^{t+1} such that the quality-adjusted price $p_m^{*t+1}=A(z)p_n^{t+1}$. For the imputation method to work, the average price changes of the *i*=1….*m* items, including the quality-adjusted price p_m^{*t+1}, given on the left-hand side of equation (7.13), must equal the average price change from just using the overall mean of the rest of the *i*=1….*m*-1 items, on the right-hand side of equation (7.13). The discrepancy or bias from the method is the balancing term *Q*. It is the implicit adjustment that allows the method to work. The arithmetic formulation is given here, though a similar geometric one can be readily formulated. The equation for one unavailable item is given by:

$$\frac{1}{m}\left[\frac{p_m^{*t+1}}{p_m^t}+\sum_{i=1}^{m-1}\frac{p_i^{t+1}}{p_i^t}\right]=\left[\frac{1}{(m-1)}\sum_{i=1}^{m-1}\frac{p_i^{t+1}}{p_i^t}\right]+Q \qquad (7.13)$$

$$Q=\frac{1}{m}\frac{p_m^{*t+1}}{p_m^t}-\frac{1}{m(m-1)}\sum_{i=1}^{m-1}\frac{p_i^{t+1}}{p_i^t} \qquad (7.14)$$

and for *x* unavailable items by:

$$Q=\frac{1}{m}\sum_{i=m-x+1}^{m}\frac{p_m^{*t+1}}{p_m^t}-\frac{x}{m(m-x)}\sum_{i=1}^{m-x}\frac{p_i^{t+1}}{p_i^t} \qquad (7.15)$$

7.58 The relationships are readily visualized if r_1 is defined as the arithmetic mean of price changes of items that continue to be recorded and r_2 of quality-adjusted unavailable items. For the arithmetic case,

where $r_1=\left[\sum_{i=1}^{m-x}p_i^{t+1}/p_i^t\right]\div(m-x)$ and $r_2=\left[\sum_{i=m-x+1}^{m}p_i^{*t+1}/p_i^t\right]\div x$ (7.16)

then the bias of arithmetic mean of ratios from substituting equation (7.16) in (7.15) is:

$$Q=\frac{x}{m}(r_2-r_1) \qquad (7.17)$$

which equals zero when $r_1=r_2$. The bias depends on the ratio of unavailable values and the difference between the mean of price changes for existing items and the mean of quality-adjusted replacement price changes. The bias decreases as either (x/m) or the difference between r_1 and r_2 decreases. Furthermore, the method is reliant on a comparison between price changes for existing items and quality-adjusted price changes for the replacement or unavailable comparison. This is more likely to be justified than a comparison without the quality adjustment to prices. For example, suppose there were m = 3 items, each with a price of 100 in period t. Let the t + 1 prices be 120 for two items, but assume the third is unavailable, i.e., x = 1 and is replaced by an item with a price of 140, of which 20 is attributable to quality differences. Then the arithmetic bias as given in equations (7.16) and (7.17), where $x=1$ and m = 3, is

$$\frac{1}{3}\left[(-20+140)/100 \;-\; \left(\frac{120}{100}+\frac{120}{100}\right)/2=0\right]$$

Had the bias depended on the unadjusted price of 140 compared with 100, the imputation would be prone to serious error. In this calculation, the direction of the bias is given by $(r_2 - r_1)$ and does not depend on whether quality is improving or deteriorating, in other words whether $A(z)>1$ or $A(z)<1$. If $A(z)>1$, a quality improvement, it is still possible that $r_2 < r_1$ and for the bias to be negative, a point stressed by Triplett (2002).

7.59 The analysis here is framed in terms of a short-run price change framework. That is, the short-run price changes between the prices in a period and those in the preceding period are used for the imputation. This is different from the long-run imputation where a base period price is compared with prices in subsequent months, and where the implicit assumptions are more restrictive.

7.60 Table 7.2 provides an illustration in which the (mean) price change of items that continue to exist, r_1, is allowed to vary for values between 1.00 and 1.5 – corresponding to a variation between no price change and a 50 per cent increase. The (mean) price change of the quality-adjusted new items compared with the items they are replacing is assumed not to change, i.e., r_2 = 1.00. The bias is given for ratios of missing values of 0.01, 0.05, 0.1, 0.25 and 0.5, both for arithmetic means and geometric means. For example, if 50 per cent of price quotes are missing and the missing quality-adjusted prices do not change, but the prices of existing items increase by 5 per cent (r_1=1.05), then the bias for the geometric mean is represented by the proportional factor 0.9759; i.e., instead of 1.05, the index should be 0.9759 x 1.05 = 1.0247. For an arithmetic mean, the bias is –0.025; instead of 1.05 it should be 1.025.

7.61 Equation (7.17) shows that the ratio x/m and the difference between r_1 and r_2 determine the bias. Table 7.2 shows that the bias can be quite substantial when x/m is relatively large. For example, for x/m = 0.25, an inflation rate of 5 per cent for existing items translates to an index change of 3.73 per cent and 3.75 per cent for the geometric and arithmetic formulations, respectively, when r_2 = 1.00, i.e., when quality-adjusted prices of unavailable items are constant. Instead of being 1.0373 or 1.0375, ignoring the unavailable items would give a result of 1.05. Even with 10 per cent missing (x/m = 0.1), an inflation rate of 5 per cent for existing items translates to 4.45 per cent and 4.5 per cent for the geometric and arithmetic formulations, respectively, when r_2 = 1.00. Considering a fairly low ratio of x/m, say 0.05, then even when r_2 = 1.00 and r_1 = 1.20, Table 7.2 shows that the corrected rates of inflation should be 18.9 per cent and 19 per cent for the geometric and arithmetic formulations, respectively. In competitive markets, r_1 and r_2 are unlikely to differ by substantial amounts since r_2 is a price comparison between the new item and the old item after adjusting for quality differences. If r_1 and r_2 are the same, then there would be no bias from the method even if x/m = 0.9. There may, however, be more sampling error. It should be borne in mind that it is not appropriate to compare bias between the arithmetic and geometric means, at least in the form they take in Table 7.2. The latter would have a lower mean, rendering comparisons of bias meaningless.

Table 7.2. Example of the bias from implicit quality adjustment when the (mean) price change of quality-adjusted new items compared with the items they are replacing is assumed not to change (r_2=1.00)

	Geometric mean					Arithmetic mean				
	Ratio of missing items, *x/m*					Ratio of missing items, *x/m*				
	0.01	0.05	0.1	0.25	0.5	0.01	0.05	0.1	0.25	0.5
$\boldsymbol{r_1}$										
1	1	1	1	1	1	0	0	0	0	0
1.01	0.999901	0.999503	0.999005	0.997516	0.995037	-0.0001	-0.0005	-0.001	-0.0025	-0.005
1.02	0.999802	0.99901	0.998022	0.995062	0.990148	-0.0002	-0.001	-0.002	-0.005	-0.01
1.03	0.999704	0.998523	0.997048	0.992638	0.985329	-0.0003	-0.0015	-0.003	-0.0075	-0.015
1.04	0.999608	0.998041	0.996086	0.990243	0.980581	-0.0004	-0.002	-0.004	-0.01	-0.02
1.05	0.999512	0.997563	0.995133	0.987877	0.9759	-0.0005	-0.0025	-0.005	-0.0125	-0.025
1.1	0.999047	0.995246	0.990514	0.976454	0.953463	-0.001	-0.005	-0.01	-0.025	-0.05
1.15	0.998603	0.993036	0.986121	0.965663	0.932505	-0.0015	-0.0075	-0.015	-0.0375	-0.075
1.2	0.998178	0.990925	0.981933	0.955443	0.912871	-0.002	-0.01	-0.02	-0.05	-0.1
1.3	0.99738	0.986967	0.974105	0.936514	0.877058	-0.003	-0.015	-0.03	-0.075	-0.15
1.5	0.995954	0.979931	0.960265	0.903602	0.816497	-0.005	-0.025	-0.05	-0.125	-0.25

r_1 =(mean) price change for items that continue to exist.

7.62 An awareness of the market conditions relating to the commodities concerned is instructive in understanding likely differences between r_1 and r_2. The concern here is when prices vary over the life cycle of the items. Thus, for example, at the introduction of a new model, the price change may be quite different from price changes of other existing items. Thus assumptions of similar price changes, even with quality adjustment, might be inappropriate. Greenlees (2000) gives the example of personal computers: new computers enter the market at prices equal to, or lower than, prices of previous models, but with greater speed and capability. An assumption that $r_1 = r_2$ could not be justified. He continues with the example of apparel, in which new clothing enters the market at relatively high quality-adjusted prices, while old, end-of-season or out-of-style clothes are being discounted. Again there will be bias, as r_1 differs from r_2.

7.63 Some of these differences arise because markets are composed of different segments of consumers. Indeed, the very training of consumer marketers involves consideration of developing different market segments and ascribing to each appropriate pricing, product quality, promotion and place (method of distribution) – the 4Ps of the marketing mix (Kotler, 1991). In addition, consumer marketers are taught to plan the marketing mix for the life cycle of items. Such planning allows for different inputs of each of these marketing mix variables at different points in the life cycle. This includes “price skimming” during the period of introduction, when higher prices are charged to skim off the surplus from segments of consumers willing to pay more. The economic theory of price discrimination would also predict such behaviour. Thus the quality-adjusted price change of an old item compared with a new replacement item may be higher than price changes of other items in the product group. After the introduction of the new item its prices may fall relative to others in the group. There may be no law of one price change for differentiated items within a market. Berndt et al. (2003) clearly show how, after patents expire, the price of branded prescription pharmaceuticals can increase with the entry of new generic pharmaceuticals at a lower price, as particularly loyal, less price-sensitive customers maintain their al-legiance to the branded pharmaceuticals.

7.64 There is thus little in economic or marketing theory to support any expectation of similar (quality-adjusted) price changes for new and replacement items, as compared to other items in the product group. Some knowledge of the realities of the particular market under study would be helpful when considering the suitability of this approach. Two aspects need to be considered in any decision to use the imputation approach. The first is the proportion of replacements; Table 7.2 provides guidance here. The second is the expected difference between r_1 and r_2. It is clear from the above discussion that there are markets in which they are unlikely to be similar. This is not to say the method should not be used. It is a simple and expedient approach. What arguably should not happen is that it is used by default, without any prior evaluation of expected price changes and the timing of the switch. Furthermore, its use should be targeted, by selecting items expected to have similar price changes. The selection of such items, however, should take account of the need to include a sufficiently large sample so that the estimate is not subject to undue sampling error.

7.65 The manner in which these calculations are undertaken is also worth considering. In its simplest form, the pro forma setting for the calculations, say on a spreadsheet, would usually have each item description and its prices recorded on a monthly basis. The imputed prices of the missing items are inserted into the spreadsheet, and are highlighted to show that they are imputed. The need to highlight such prices is, first, because they should not be used in subsequent imputations as if they were actual prices. Second, the inclusion of imputed values may give a false impression of a larger sample size than actually exists. Care should be taken in any audit of the number of prices used in the compilation of the index to code such observations as “imputed”.

7.66 The method described above is an illustration of a short-run imputation. As is discussed in paragraphs 7.165 to 7.173 below, there is a strong case for using short-run imputations as against long-run ones.

Class mean imputation

7.67 The class mean (or substitution relative) method of implicit quality adjustment to prices as used in the United States CPI is discussed by Schultz (1996), Reinsdorf, Liegey and Stewart (1996), Armknecht, Lane and Stewart (1997), and Armknecht and Maitland-Smith (1999). It arose from concerns similar to those considered in the previous section, that unusual price changes were found in the early introductory period, when new models were being introduced, particularly for consumer durables. Moulton and Moses (1997), using United States CPI data for 1995 in their study of selected products, found the average pure price change to be only 0.12 per cent for identical items being repriced (on a monthly or bimonthly basis), compared to an average 2.51 per cent for comparable substitutes – items judged equivalent to the items they replaced. The corresponding average price change for directly substituted quality-adjusted price changes was 2.66 per cent. Thus, the price movement of continuing items appears to be a flawed proxy for the pure price component of the difference between old and replacement items.

7.68 The class mean method was adopted in the United States CPI for cars in 1989 and was phased in for most other non-food commodities, beginning in 1992. It differed from the overall mean imputation method only in the source for the imputed rate of price change for the old item in period $t + 1$. Rather than using the category index change, obtained using all the non-missing items in the category, the imputed rate of price change was based on constant quality replacement items – those items that were judged comparable or that were quality-adjusted directly. The class mean approach was seen as an improvement on the overall mean imputation approach because the imputed price changes were based on items that had not just been replaced, but whose replacement price had benefited from a quality adjustment or the new replacement item had been judged to be directly comparable. It may be the case, however, that sufficiently large samples of comparable substitutes or directly quality-adjusted items are unavailable. Or it may be that the quality adjustments and selection of comparable items are not deemed sufficiently reliable. In that case, a targeted imputation might be considered. The targeted mean is less ambitious in that it seeks only to capture price changes of similar items, irrespective of their point in the life cycle. Yet it is an improvement on the overall mean imputation, as long as sufficiently large sample sizes are used.

Comparable replacement

7.69 The comparable replacement method requires the respondent to make a judgement that the replacement is of a similar quality to the old item and any price changes are untainted by quality changes. For specialized chain stores in Table 7.1(b), item 3 might be judged to be comparable to item 2 and its prices in subsequent months might be used to continue the series. The price of item 3 (6.5) in March would be used as the price in March of item 2, whose January to March price change would be 6.5/6 x 100=1.0833 or 8.33 per cent. Lowe (1999) notes the common practice of manufacturers of television sets to change model numbers with a new production run, though nothing physically has changed, or when small changes take place in specifications, such as the type of remote controls, or the number or placement of jacks. The method of comparable replacement relies on the efficacy of the price collectors and, in turn, on the adequacy of the specifications used as a description of the items. Statistical agencies may rightly be wary of sample sizes being reduced by dropping items for which prices need to be imputed, and also wary of the intensive use of resources to make explicit estimates as outlined below. The use of repriced items of a comparable specification has much to commend it. If the quality of items is improving, however, the preceding item will be inferior to the current one. Continually ignoring small changes in the quality of replacements can lead to an upward bias in the index. The extent of the problem will depend on the proportion of such occurrences, the extent to which comparable items are accepted as being so despite quality differences, and the weight attached to those items. Proposals in Chapter 8 to monitor types of quality adjustment methods by product area provide a basis for a strategy for applying explicit adjustments where they are most needed.

Linked to show no price change

7.70 Linking attributes any price change between the replacement item in the current period and the old item in the preceding period to the change in quality. For example, in Table 7.1(b), a replacement item 7 is selected from a department store for the missing March item 6. Items 6 and 7 may be of different quality, the price difference being quite large. The change in price is assumed to be attributable to a change in quality. An estimate is made for p^{72} by equating it to p^{73}, to show no change, i.e., the assumed price of item 7 in February is 14 in Table 7.1(b). There is thus assumed to be no price change over the period February to March for item 7. The January to March result for item 6 is $(12/12)\times(14/14)$ = 1.00, indicating no change. For the period March to April, however, the price of item 7 in March can be compared with the imputed p^{72} for February and linked to the preceding results. So the

January to April comparison is composed of the January to February comparison for item 6, linked to (multiplied by) the February to April comparison for item 7. This linking is analogous to the procedures used for the chained and short-run framework discussed in paragraphs 7.153 to 7.158 and 7.171 to 7.173 below. The method is born out of circumstances where comparable replacement items are not available and there are relatively large price differences between the old and replacement items, these being from different price bases and of different qualities. It is not possible to separate out how much of this difference is attributable to price changes and how much to quality changes, so the method attributes it all to quality and holds price constant. The method introduces a degree of undue price stability into the index. It may well be the case that the period of replacement is when substantial price changes are taking place and that these are wrongly attributed to quality changes by this method. Article 5 of the European Commission (EC) Regulation No. 1749/96 requires Member States to avoid "automatic linking". Such linking is equivalent to the assumption that the difference in price between two successive models is wholly attributable to a difference in quality (Eurostat, 2001a, p. 125).

Carry-forward

7.71 With the carry-forward method, when an item becomes unavailable, say in period *t*, the price change calculation uses the old *t*-1 price, simply carried forward as if there were no change. Thus from Table 7.1(a) for specialized chain stores for the period January to March, the Jevons and Dutot indices (Chapter 20) are:

$$P_J(p^1, p^3) = \left[\left(p^{13}/p^{11} \times p^{22}/p^{21}\right)\right]^{1/2} \text{ and}$$
$$P_D(p^1, p^3) = [(p^{13} + p^{22})/(p^{11} + p^{21})] \qquad (7.18)$$

with p^{22} filling in for the missing p^{23}. This introduces undue stability into the index, which is aggravated if the old price, p^{22}, continues to be used to fill in the unobserved prices in subsequent periods. It induces an inappropriate amount of stability into the index and may give a misleading impression of the active sample size. The practice of the carry-forward method is banned under Article (6) of the EC Regulation No. 1749/96 for Harmonized Indices of Consumer Prices (Eurostat, 2001a, p. 126). To use this method an assumption is made that the price from this outlet would not change. This method should only be used if it is fairly certain that there would be no price change.

Explicit methods of quality adjustment

7.72 The aforementioned methods do not rely on explicit information on the value of the change in quality, $A(z)$. This section discusses the following methods that rely on obtaining an explicit valuation of the quality difference: expert judgement; quantity adjustment; differences in production or option costs; and the hedonic approach.

Expert judgement

7.73 Hoven (1999) describes comparable replacement as a special case of subjective quality adjustment, because the determination of product equivalence is based on the judgement of the commodity specialist. One objection to subjective methods is the inability to provide results that can be independently replicated. Yet in comparable replacement, and for the selection of representative items, a subjective element is part of normal procedure. This is not, of course, an argument for expanding the use of subjective methods.

7.74 Hoffman (1999) describes a possibly unique alternative for quality adjustment of replacement items in the German CPI. When a new product is more expensive than the item it replaces, a flexible adjustment factor can be employed, attributing none, some, or all of the price difference to improved quality. In particular, when no precise information is available on which to make a quality determination, it is permissible for an adjustment to be made of 50 per cent of the price difference. The guidelines used in Germany since 1997 replaced flawed procedures in which the particular methods chosen for individual quality adjustments depended on the difference in price alone. As Hoffmann notes, however, even in the current approach no quality adjustment is made if the new item is less expensive than the old. Consequently, problems could arise if an increase in quality were accompanied by a decrease in price (or *vice versa*). The methods used in the German CPI are needed because quality adjustments for most goods are made not in the central CPI office but by price collectors in the field. Wide use of the hedonic and production cost approaches is precluded under these conditions. Thus, the organizational structure of the statistical agency, as well as its funding level, will necessarily influence its choice of quality adjustment methods.

7.75 Reports by consumer associations and product evaluations in consumer magazines are not advised by Turvey (1998), who cites a study which correlated quality ratings and prices for 135 products categories using *Consumer Reports*. The average correlation was 0.26, with over half having a positive association, just over a third no association and the rest a negative one. He also argues against “best buy” estimates, which are expert views as to what a sensible consumer should pay, as opposed to what the market price will be (see also Combris, Lecocqs and Visser, 1997).

7.76 The use of expert views as to consumer calculations may be appropriate for highly complex items where alternative methods are not feasible. Experts should be guided with regard to the nature of the estimate required. More than one expert should be chosen and, where possible, the experts should be from different backgrounds. It is also advisable to give the experts some indication of the interval in which their estimate should lie. The well-known Delphi method (for example, see Czinkota and Ronkainen, 1997) may be applicable. In this approach a panel of experts never meet, to avoid any “bandwagon” effect regarding their estimates. They are asked to provide an estimate of the average response and the range of likely responses. The median is taken of these estimates and any estimate that is considered extreme is sent back to the expert concerned, who is asked to account for possible reasons behind the difference. It may be that the particular expert has a useful perspective on the problem, which the other experts had not considered. If the expert argues the case convincingly, the response is fed back to the panel who are asked if they wish to change their views. A new median is taken, and further iterations are possible. The Delphi method is time-consuming and expensive, but it reflects the care needed in such matters. If an adjustment is required for a product area with a large weighting in the CPI, and no other techniques are available, it is a possible alternative.

Quantity adjustment

7.77 Quantity adjustment is one of the most straightforward explicit adjustments to undertake. It is applicable when the size of the replacement item differs from that of the available item. In some situations there is a readily available quantity metric that can be used to compare the items. Examples are the number of units in a package (e.g., paper plates or vitamin pills), the size or weight of a container (e.g., kilogram of flour, litre of cooking oil), or the size of sheets or towels. Quantity adjustment to prices can be accomplished by scaling the price of the old or new item by the ratio of quantities. The index production system may do this scaling adjustment automatically, by converting all prices in the category to a price per unit of size, weight or number. Scaling is important. For example, if cooking oil is now sold in 5 litre containers instead of 2.5 litre ones, it should not be the case that prices have doubled.

7.78 There is, however, a second aspect. In the pharmaceutical context, for example, prices of bottles of pills of different sizes differ. A bottle of 100 pills, each having 50 milligrams of a drug, is not the same as a bottle of 50 pills of 100 milligrams, even though both bottles contain 5,000 milligrams of the same drug. If there is a change, say, to a larger size container, and a *unit* price decrease of 2 per cent accompanies this change, then it should not be regarded as a price fall of 2 per cent if consumers gain less utility from the larger and more inconvenient containers. In practice it will be difficult to determine what proportion of the price fall is attributable to quality and what proportion to price. A general policy is not to automatically interpret unit price changes arising from packaging size changes as pure price changes, if contrary information is available.

7.79 Consider a further example: a branded bag of flour previously available in a 0.5 kilogram bag priced at 1.5 is replaced by a 0.75 kilogram bag priced at 2.25. The main concern here is with rescaling the quantities. The method would use the relative quantities of flour in each bag for the adjustment. The prices may have increased by [(2.25/1.5) ×100 = 150] 50 per cent but the quality-adjusted prices (i.e. prices adjusted by size) have remained constant [(2.25/1.5)×(0.5/0.75)×100=100]. The approach can be outlined in a more elaborate manner by recourse to Figure 7.1. The concern here is with the part of the unbroken line between the (price, quantity) coordinates (1.5, 0.5) and (2.25, 0.75), both of which have *unit* prices of 3 (price = 1.5/0.5 and 2.25/0.75). There should be no change in quality-adjusted prices. The symbol Δ denotes a change. The slope of the line is β which is Δprice/Δsize = (2.25-1.5)/(0.75-0.50)=3, i.e., the change in price arising from a unit (kilogram) change in size. The quality- (size-)adjusted price in period *t*-1 of the old *m* bag is:

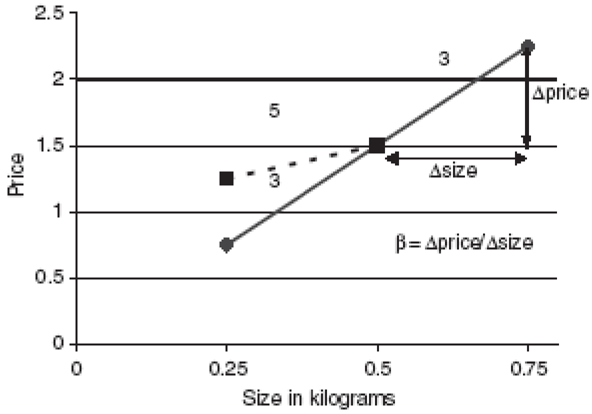
$$\hat{p}_m^{t-1} = p_m^{t-1} + \beta \Delta \text{size} = 1.5 + 3\,(0.75 - 0.5) = 2.25 \qquad (7.19)$$

The quality-adjusted price change shows no change, as before:

$$p_n^t / \hat{p}_m^{t-1} = 2.25 / 2.25 = 1.00$$

The approach is outlined in this form so that it can be seen as a special case of the hedonic approach (discussed below), where price is related to a number of quality characteristics of which size may be only one.

Figure 7.1 Quality adjustment for different sized items



7.80 The method can be seen to be successful on intuitive grounds as long as the unit price of different-sized bags remains constant. If the switch was from the replacement of the 0.5 kilogram bag to a 0.25 kilogram one priced at 0.75, as shown by the continuation to coordinate (0.75, 0.25) of the unbroken line in Figure 7.1, the quality-adjusted prices would again not change, assuming, however, that the unit (kilogram) prices were 5, 3 and 3 for the 0.25, 0.5 and 0.75 kilogram bags, respectively, as shown in Table 7.3 and in Figure 7.1 (including the broken line). Then the measure of quality-adjusted price change would depend on whether the 0.5 kilogram bag was replaced by the 0.25 kilogram one (a 67 per cent increase) or the 0.75 kilogram one (no change). This is not satisfactory because the choice of replacement size is arbitrary. The rationale behind the quality adjustment process is to ask: does the difference in unit price in each case reflect different levels of utility? If so, adjustments should be made to the unit prices to bring them into line. If not, adjustments should be made to the unit prices for that proportion attributable to differences in utility gained from, say, more convenient packaging or the availability of smaller lots. It may be obvious from the nature of the product that an item packaged in a very small size with a disproportionately high unit price carries an unusually high profit margin, and that an appropriate replacement for a large-sized item would not be this very small one.

Table 7.3 Example of size, price and unit price of bags of flour

Size (kilograms)	**First price**	**First unit price**	**Second price**	**Second unit price**
0.25	0.75	3	1.25	5
0.5	1.50	3	1.50	3
0.75	2.25	3	2.25	3

Differences in production or option costs

7.81 A natural approach to quality adjustment is to adjust the price of an old item by an amount equal to the resource costs of the additional features of the new item; i.e., to compare relative prices using:

$$p_n^t / \hat{p}_m^{t-1} \quad \text{where} \quad \hat{p}_m^{t-1} = p_m^{t-1} + x \qquad (7.20)$$

and x is the value of the additional features in period t-1 prices. This value should be a consumer’s valuation, reflecting the additional flow of services or utility. One source of data is the manufacturers. They would be asked to provide data on production costs, to which retail mark-ups and associated indirect taxes would be added. This approach is most practicable in markets where there is a relatively small number of manufacturers, and where updates of models are infrequent and predictable. It only works if there is good communication between manufacturers and the statistical agency staff. It is particularly suitable when the quality adjustments are also being undertaken to calculate the producer price index (PPI) or other price programmes. Greenlees (2000) provides an example for new trucks and motor vehicles in the United States in 1999. Just prior to the introduction of the annual models, BLS staff visit selected manufacturers to collect cost information. The data are used in the PPI and International Comparison Programmes as well as in the CPI, and the information-gathering activity is a joint operation of the three programmes. Allowable product changes for the purpose of quality adjustments include occupant safety enhancements, mechanical and electrical improvements to overall vehicle operation or efficiency, changes that affect length of service or need for repair, and changes affecting comfort or convenience.

7.82 Bearing in mind the caveat in paragraph 7.30, the producer orientation of the PPI implies that resource cost is the appropriate criterion for quality adjustment to prices

(Triplett, 1983). One distinction, then, between the use of producer cost estimates in the CPI and PPI is that only the former programme will add retail mark-ups and indirect taxes. Another important difference may occur in situations where product improvements are mandated by government. Some of these mandated improvements provide no direct benefit to the purchaser. In these cases it is appropriate to make a quality adjustment to prices for the associated resource cost in the PPI, but not in the CPI, where the appropriate criterion is user value. If only production cost data are available, then estimates of the retail mark-up must take into account the (average) age of the models under consideration. Mark-ups will decrease as models come to the end of their life cycles. Therefore, mark-ups based on models at the end of their life cycle should not be applied to the production costs of models at the start of their life cycle.

7.83 Because of these difficulties in using the production cost approach, the option cost method is generally preferred. Often it is the retail price of an option that is available and this of course includes the mark-up for profit. Consider an example of the *price* of an option being used to adjust for quality. Let the prices for an item in periods *t*-1 and *t* be 10,000 and 10,500, respectively, but assume the price in period *t* is for the item with a new feature or "option", and let the price of the additional feature in period *t* be known to be 300. Then the price change would be 10,200/10,000=1.02 or 2.0 per cent. The adjustment may take a multiplicative form (see paragraphs 7.39-7.40 above): the additional option is worth 300/10,500=0.028571 of the period *t* price. The adjusted price in period *t*-1 is therefore 10000×1.028571=10,285.71 and the price change 10,500/10,285.71=1.020833 or about 2.08 per cent. If in subsequent periods either of these elements changes, then so too must $\hat{p}_{n,t-1}$ for those comparisons. The option cost method is thus a method for use in stable markets with stable technologies. Alternatively, it may be preferable to estimate a one-off adjustment to the preceding base period price and then compare all subsequent prices with the new option to this estimate; i.e. 10,500/10,300=1.019417 or approximately 2 per cent.

7.84 Option costs are thus useful in situations in which the old and new items differ by quantifiable characteristics that can be valued in monetary terms by reference to market prices. For example, nuts may be available roasted or unroasted, and food items may be available cooked or uncooked. Consider the addition of a feature to a car model. The feature may have been available as an option either in the prior period or currently for other models, providing an absolute or proportional consumer valuation. Armknecht and Maitland-Smith (1999) note that when radial tyres became a standard feature on new cars, the price of adding optional radial tyres was used to determine the quality adjustments in the United States CPI. The valuation of a quantifiable product feature may be readily available from the comparison of different product prices. Turvey et al. (1989) give the example of whiskies of different proofs (percentage alcohol content). The quality adjustment for a change in the alcohol content of one product may be inferred from the market relationship between proof and price.

7.85 Consider the addition of a feature to a product – say an installed automatic ice-maker in a refrigerator (Shepler, 2000). Refrigerators can be sold as standard or with an installed automatic ice-maker. The price collector may always have collected prices on the standard model, but this may no longer be in production, being replaced by a model with an installed automatic ice-maker. The cost of the option is thus known from before and a continuing series can be developed by using equation (7.20) and simply adjusting the old price in the base period for the option cost. Even this process may have its problems. First, the cost of producing something as standard may be lower than when it was an option, all new refrigerators now having the installed automatic ice-maker. This saving may be passed on, at least in part, to the consumer. The option cost method would thus understate a price increase. Triplett (2002) cites a study by Levy et al. (1999) in which a car theft system was installed as standard but disabled when the option was not required. It was seemingly cheaper to produce this way. Second, by including something as standard the consumer's valuation of the option may fall since buyers cannot refuse it. Some consumers may attribute little value to the option. The overall effect would be that the estimate of the option cost, priced for those who choose it, is likely to be higher than the implicit average price consumers would pay for it as standard. Estimates of the effect on price of this discrepancy should in principle be made, though in practice are quite difficult.

7.86 Option cost adjustments can be seen to be similar to quantity adjustments, except that instead of size being the additional quality feature of the replacement, the added quality can be any other individual feature. The comparison is: $p_n^t / \hat{p}_m^{t-1}$ where $\hat{p}_m^{t-1} = p_m^{t-1} + \beta \Delta z$ for an individual z characteristic where $\Delta z = \left(z_n^t - z_m^{t-1}\right)$. The characteristics may be the size of the random access memory (RAM) of a personal computer (PC) when a specific model of PC is replaced by a model that is identical except for the amount of RAM it possesses. If the relationship between price and RAM is linear, the above formulation is appropriate. Many web pages give the price of additional RAM as being independent of other features of PCs, and a linear adjustment is appropriate. Bear in mind that a linear formulation values the worth of a fixed additional amount of RAM to be the same, irrespective of the amount of RAM the machine possesses.

7.87 The relationship may, of course, be non-linear. Say, for example, for every additional 1 unit of *x*, *y* increases by 1.5 unit $(\beta = 1.015)$. In this case,

$$\hat{p}_m^{t-1} = p_m^{t-1} \beta^z \qquad (7.21)$$

for $p_n^t / \hat{p}_m^{t-1}$ as a measure of quality-adjusted price changes. Again the *z* change may reflect the service flow, but the non-linearity in the price–*z* relationship may reflect the increasing or decreasing utility to the scale of the provision. Possession of the characteristic in up-market models of the item may be priced at a higher rate than in a lower-priced one, i.e. $\beta \geq 1$ in equation (7.21).

7.88 Consider Figure 7.1 with the *z* characteristic being the option on the horizontal axis. The similarity between the quantity adjustment and the option cost approaches is apparent since both relate price to some dimension of quality: the size or the option. The option cost approach can be extended to more than one quality dimension. Both approaches rely on the acquisition of estimates of the change in price resulting from a change in the option or size: the β slope estimates. In the case of the quantity adjustment, this was taken from an item identical to the one being replaced, aside from the fact that it was of a different size. The β slope estimate in this case was perfectly identified from the two pieces of information. It is as if the nature of the experiment controlled for changes in the other quality factors by comparing prices of what is essentially the same thing except for the quantity (size) change.

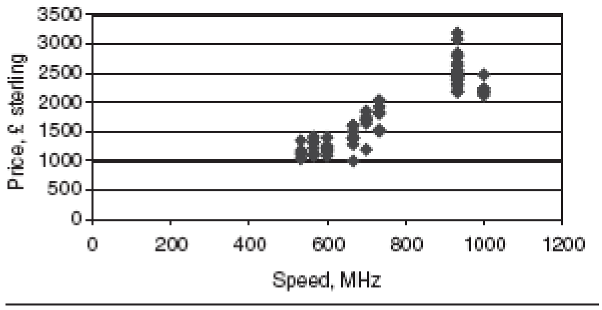
7.89 The same reasoning applies to option costs. There may be, for example, two items, identical but for the possession of a feature. This allows the value of the feature to be determined. Yet sometimes the value of a feature or option has to be extracted from a much larger data set. This may be because the quality dimension takes a relatively large range of possible numerical values without an immediately obvious consistent valuation. Consider the simple example of only one feature varying for a product, the speed of processing of a PC. It is not a straightforward matter to determine the value of an additional unit of speed. To complicate matters, there may be several quality dimensions to the items and not all combinations of these may exist as items in the market in any one period. Furthermore, the combinations existing in the second period being compared may be quite different to those in the first. Considering these aspects leads to a more general framework, known as the hedonic approach.

Hedonic approach

7.90 The hedonic approach is an extension of the two preceding approaches in that, first, the change in price arising from a unit change in quality – the slope of the line in Figure 7.1 – is now estimated from a data set comprising prices and quality characteristic values of a larger number of varieties. Second, the quality characteristic set is extended to cover, in principle, all major characteristics that might influence price, rather than just the quantity or option adjustment. The theoretical basis for hedonic regressions will be covered in Chapter 21 and is briefly reviewed below, following an example based on personal computers.

7.91 It should be noted that the method requires an extension of the data set to include values of the price-determining quality characteristics for each item. Under the matched models method each price collector needed only to collect sufficient data on each model to allow the model to be identified for subsequent repricing. The extension required in the hedonic approach is that all price-determining characteristics should be collected for each model. Checklists for the characteristics of a product have been found by Merkel (2000) to improve the quality of data collected, as well as serving the needs of hedonic adjustments (see also Chapter 6 on price collection and Liegey, 1994). If an item goes missing, any difference in the characteristics of its replacement can be identified and, as will be shown, a valuation can be ascribed to such differences using the hedonic approach.

Figure 7.2 Scatter diagram showing prices and processing speeds of personal computers



7.93 Appendix 7.1 to this chapter provides data taken from the United Kingdom Compaq and Dell web sites in July 2000 on the prices and characteristics of 64 desktop personal computers (PCs). Figure 7.2 is a scatter diagram constructed from this information, relating the price (£ sterling) to the processing speed (MHz). It is apparent that PCs with higher speeds command higher prices – a positive relationship. Under the option cost framework above, a switch from a 733 MHz to a 933 MHz PC would involve a measure of the slope of the line between two unique points. The approach requires that there are 733 MHz and 933 MHz PCs that are identical except for their processing speed. From Figure 7.2 and Appendix 7.1 it is apparent that there are several PCs with the same speed but different prices, resulting from the fact that other things differ. To estimate the value given to additional units of speed, an estimate of the slope of the line that best fits the data is required. In Figure 7.1 the actual slope was used; for the data in Figure 7.2 an estimate of the slope needs to be derived from an estimate of the equation of the line that best fits the data, using ordinary least squares regression. Facilities for regression are available on standard statistical and econometric software, as well as spreadsheets. The estimated (linear) equation in this instance is:

$$\hat{\text{Price}} = -658.436 + 3.261\text{Speed} \qquad \bar{R}^2 = 0.820 \qquad (7.22)$$

The coefficient of speed is the estimated slope of the line: the change in price (£3,261) resulting from a 1 MHz change in speed. This can be used to estimate quality-adjusted price changes for PCs of different speeds. The value of $\bar{R}^2$ indicates that 82 per cent of price variation is explained by variation in processing speed. A *t*-statistic to test the null hypothesis of the coefficient being zero was found to be 18.83: recourse to standard tables on *t*-statistics found the null hypothesis was rejected at a 1 per cent level. The fact that the estimated coefficient differs from zero cannot be attributed to sampling errors at this level of significance. There is a probability of 1 per cent that the test has wrongly rejected the null hypothesis.

7.93 The range of prices for a given speed – for example for 933 MHz – can, however, be seen from Appendix 7.1 to be substantial. There is a price range of about £1,000, which suggests that other quality characteristics may be involved. Table 7.4 provides the results of a regression equation that relates price to a number of quality characteristics using the data in Appendix 7.1. Such estimates can be provided by standard statistical and econometric software, as well as spreadsheets.

7.94 The second column provides the results from a linear regression model, the dependent variable being price. The first variable is (processor) speed, with a coefficient of 2.731 – a unit MHz increase in processing speed leads to an estimated £2.731 increase (positive sign) in price. A change from 733 MHz to 933 MHz would be valued at an estimated 200 * 2.731 = £546.20. The coefficient is statistically significant – its difference from zero (no effect) not being attributable to sampling errors at a 0.1 per cent level of significance. This estimated coefficient is based on a multivariate model: it is the effect of a unit change in processing speed on price, having controlled for the effect of other variables in the equation. The preceding result of 3.261 in equation (7.22) was based on only one variable, and is different from this improved result.

7.95 The brand variables are dummy intercepts taking values of 1 for, say, a Dell computer, and zero otherwise. While brands are not in themselves quality characteristics, they may be proxy variables for other factors such as reliability or after-sale service. The inclusion of such brand dummies also goes some way towards reflecting segmented markets as communities of buyers, as discussed in Chapter 21. Similar dummy variables were used for the other makes or brands (Compaq Presario and Compaq Presignia), except for one brand (Compaq Deskpro) which, in this case, was taken to form the benchmark against which other models are compared. The coefficient of the Dell brand is an estimate of the difference between a Dell brand's worth and that of a Compaq Deskpro, other variables being constant, i.e. £1,330.78 cheaper. Similarly, an Intel Pentium III commands an estimated £282.78 premium on an AMD Athlon.

7.96 The estimate for processor speed was based on data for Dell and Compaq PCs. If the adjustment for quality is between two Dell PCs, it might be argued that data on Compaq PCs should be ignored. Separate regressions could be estimated for each make, but this would severely restrict the sample size. Alternatively, an interaction term or slope dummy can be used for variables which are believed to have a distinctive brand-interaction effect. An example of such a dummy would be, say, Dell * Speed, which takes the value of "speed" when the PC is a Dell and zero otherwise. The coefficient of this variable (see Table 7.4) is 1.714; it is an estimate of the additional (positive sign) price arising *for a Dell PC* over and above that already arising from the standard valuation of a 1 MHz increase in speed. For Dell PCs it is £2.731 + £1.714 = £4.445. Thus if the replacement Dell PC is 200 MHz faster than the unavailable PC, the price adjustment to the unavailable PC is to add $200 \times £4.465$ = £893. Interactive terms for other variables can similarly be defined and used. The estimation of regression equations is easily undertaken using econometric or statistical software, or data analysis facilities in spreadsheets. An explanation of the techniques is given in many texts, including Kennedy (1998) and Maddala (1988). In Chapter 21, econometric concerns particular to the estimation of hedonic regressions are discussed.

7.97 The value $\overline{R}^2$ is the proportion of variation in price explained by the estimated equation. More formally, it is 1 minus the ratio of the variance of the residuals, $\Sigma_{i=1}^{N}\left(p_i^t - \hat{p}_i^t\right)^2 / N$, of the equation to the variance of prices, $\Sigma_{i=1}^{N}\left(p_i^t - \overline{p}_i^t\right)^2 / N$. The bar on the term R^2 denotes that an appropriate adjustment for degrees of freedom is made to this expression, which is necessary when comparing equations with different numbers of explanatory variables. At 0.934 (see Table 7.4), the value $\overline{R}^2$ is very high. A high value of $\overline{R}^2$ can, however, be misleading for the purpose of quality adjustment. First, such values indicate that the explanatory variables account for much of the price variation. This may be over a relatively large number of varieties of goods in the period concerned. This, of course, is not the same as implying a high degree of prediction for an adjustment to a replacement item of a single brand in a subsequent time period. Predicted values depend for their accuracy not just on the fit of the equation, but also on how far the characteristics of the item whose price is to be predicted are from the means of the sample. The more unusual the item, the higher the prediction probability interval. Second, the value $\overline{R}^2$ indicates the proportion of variation in prices explained by the estimated equation. It may be that 0.90 is explained while 0.10 is not explained. If the dispersion in prices is very large, this still leaves a large absolute margin of prices unexplained. Nonetheless, a high $\overline{R}^2$ is a necessary condition for the use of hedonic adjustments.

7.98 Hedonic regressions should generally be conducted using a semi-logarithmic formulation (Chapter 21). The dependent variable is the (natural) logarithm of the price, but the variables on the right-hand side of the equation are kept in their normal units, hence the semi-logarithmic formulation. A double-logarithmic formulation would also take logarithms of the right-hand side *z* variables. However, if any of these *z* variables are dummy variables which take the value of zero in some instances, the double-logarithmic formulation would break down because logarithms of zero cannot be taken. The focus is thus on the semi-logarithmic form. This concern with linear and semi-logarithmic formulations is equivalent to the consideration of additive and multiplicative formulations discussed in paragraphs 7.39 to 7.40 above. A linear model would, for example, ascribe an extra £282.78 to a PC with an Intel Pentium III as opposed to an AMD Athlon, irrespective of the price of the PC. This is common in pricing strategies using the World Wide Web. More often than not, however, the same options are valued at a higher price for up-market goods and services. In this case, equation (7.22) for a multivariate model becomes:

$$\text{Price} = \beta_0 \beta_1^{z_1} \beta_2^{z_2} \beta_3^{z_3} \ldots \beta_n^{z_n} \varepsilon$$

$$\text{or ln Price} = \ln \beta_0 + z_1 \ln \beta_1 + z_2 \ln \beta_2 + z_3 \ln \beta_3 + \ldots z_n \ln \beta_n + \ln \varepsilon \quad (7.23)$$

Note that this is a semi-logarithmic form; logarithms are taken of only the left-hand-side variable, i.e., price. Each of the *z* characteristics enters the regression without having logarithms taken. This has the advantage of allowing dummy variables for the possession or otherwise of a feature to be included on the right-hand side. Such dummy variables take the value of one if the item possesses the feature and zero otherwise. Matters relating to the choice of functional form are discussed in more detail in Chapter 21.

7.99 The taking of logarithms of the first equation (7.23) allows it to be transformed in the second equation to a linear form. This allows the use of a conventional ordinary least squares (OLS) estimator to yield estimates of the logarithms of the coefficients. These are given in the third column of Table 7.4 and have a useful direct interpretation: if these coefficients are multiplied by 100, they are the percentage change in price arising from a 1 unit change in the explanatory variable. For (processor) speed there is an estimated 0.1364 per cent change in price for each additional MHz the replacement item has over and above the unavailable item. When dummy variables are used, the coefficients, when multiplied by 100, are estimates of the percentage change in price, given by $(e^{\beta} - 1)100$. For example, for a rewritable CD-RW compared to a read-only CD-ROM the change in price is 8.916 per cent. There is some bias in these coefficients; and in the (semi-) logarithmic equation, half the variance of each coefficient should be added to the coefficient before using it (Teekens and Koerts, 1972). For a read-only CD-ROM, the *t*-statistic is 2.88; this is equal to the coefficient divided by its standard error, the standard error being 0.08916/2.88 = 0.03096 and the variance: 0.03096^2 = 0.000958. The adjustment is to add 0.000958/2 to 0.08916, giving 0.089639 or 8.9639 per cent.

Table 7.4. Hedonic regression results for Dell and Compaq personal computers

Dependent variable	**Price**	**ln price**
Constant	-725.996 (2.71)**	6.213 (41.95)***
Speed (processor, MHz)	2.731 (9.98)***	0.001364 (9.02)***
RAM (random access memory, MB)	1.213 (5.61) ***	0.000598 (5.00) ***
HD (hard drive capacity, MB)	4.517 (1.96)*	0.003524 (2.76)**
Brand (benchmark: Compaq Deskpro)		
Compaq Presario	-199.506 (1.89)*	-0.152 (2.60)**
Compaq Prosignia	-180.512 (1.38)*	-0.167 (2.32)*
Dell	-1,330.784 (3.74)***	-0.691 (3.52)***
Processor (benchmark: AMD Athlon)		
Intel Celeron	393.325 (4.38)***	0.121 (2.43)**
Intel Pentium III	282.783 (4.28)***	0.134 (3.66)***
ROM-drive (benchmark: CD-ROM)†		
CD-RW (compact disk, re-writable)	122.478 (56.07)***	0.08916 (2.88)**
DVD (digital video drive)	85.539 (1.54)	0.06092 (1.99)*
Dell* Speed (MHz)	1.714 (4.038)***	0.000820 (3.49)***
N	63	63
$\bar{R}^2$	0.934	0.934

† Read only memory.

Figures in brackets are *t*-statistics testing a null hypothesis of the coefficient being zero.

***,** and * denote statistically significant at 0.1 per cent, 1 per cent and 5 per cent levels, respectively, tests being one-tailed.

7.100 The approach is particularly useful when the market does not reveal the price of the quality characteristics required for the adjustment. Markets reveal prices of items, not quality characteristics, so it is useful to consider items as tied bundles of characteristics. A sufficiently large data set of items with their characteristics and sufficient variability in the mix of characteristics between the items allows the hedonic regression to provide estimates of the implicit prices of the characteristics. The theory behind such estimates is discussed in Chapter 21. A number of ways of implementing the method are outlined below.

7.101 Some mention should first be made of the interpretation of the coefficients from hedonic regressions. The matter is discussed in detail in Chapter 21; only the conclusions are summarized here. There used to be an erroneous perception that the coefficients from hedonic methods represented estimates of user value as opposed to resource cost. The former is the relevant concept in constructing a consumer price index, while for PPI construction it is the latter. Rosen (1974) found that hedonic coefficients may reflect both user value and resource cost – both supply and demand influences. There is what is referred to in econometrics as an identification problem; in other words, the observed data do not permit the estimation of the underlying demand and supply parameters. Suppose that the production technology of sellers is the same, but that buyers differ. Then the hedonic function describes the prices of characteristics that the firm will supply with the given ruling technology to the current mixture of tastes. There are different tastes on the consumer side, so what appears in the market is the result of firms trying to satisfy consumer preferences for a constant technology and profit level. The structure of supply is revealed by the hedonic price function. Now suppose that sellers differ, but that buyers' tastes are the same. Here the hedonic function $p(z)$ identifies the structure of demand. Of these two possible assumptions, uniformity of tastes is unlikely, while uniformity of technologies is more likely, especially when access to technology is unrestricted in the long run. Griliches (1988, p. 120) has argued in the context of a consumer price index:

> My own view is that what the hedonic approach tries to do is to estimate aspects of the budget constraint facing consumers, allowing thereby the estimation of "missing" prices when quality changes. It is not in the business of estimating utility functions *per se*, though it can also be useful for these purposes … what is being estimated is the actual locus of intersection of the demand curves of different consumers with varying tastes and the supply curves of different producers with possible varying technologies of production. One is unlikely, therefore, to be able to recover the underlying utility and cost functions from such data alone, except in very special circumstances.

It is thus necessary to take a pragmatic stance. In many cases the implicit quality adjustment to prices outlined in paragraphs 7.44 to 7.71 may be inappropriate because the implicit assumptions are unlikely to be valid. In such instances, the practical needs of economic statistics require explicit quality adjustments. Not to do anything on the grounds that the measures are not conceptually appropriate would be to ignore quality change and provide wrong results. Hedonic techniques provide an important tool, making effective use of data on the price–quality relationship derived from other items in the market to adjust for changes in one or more characteristics.

7.102 The proper use of hedonic regression requires an examination of the coefficients of the estimated equations to see if they make sense. It might be argued that the very multitude of distributions of tastes and technologies, along with the interplay of supply and demand, that determine the estimated coefficients (Chapter 21) make it unlikely that "reasonable" estimates will arise from such regressions. A firm may, for example, cut a profit margin relating to a characteristic for reasons related to long-run strategic plans; this may yield a coefficient on a desirable characteristic that may even be negative (Pakes, 2001). This does not negate the usefulness of examining hedonic coefficients as part of a strategy for evaluating estimated hedonic equations. First, there has been extensive empirical work in this field and the results for individual coefficients are, for the most part, quite reasonable. Even over time individual coefficients can show quite sensible patterns of decline (van Mulligen, 2003). Unreasonable coefficients on estimated equations are the exception and should be treated with some caution. Second, one can have more faith in an estimated equation whose coefficients make sense and which predicts well, than one which may also predict well but whose coefficients do not make sense. Third, if a coefficient for a characteristic does not make sense, it may be due to multicollinearity, a data problem, and should be examined to see if this is the case (see Appendix 21.1 to Chapter 21).

Item/period	t	t+2
l	p_l^t	p_l^{t+2}
m	p_m^t	$\hat{p}_{\boldsymbol{m}}^{\boldsymbol{t}+2}$
n	$\hat{p}_n^t$	p_n^{t+2}

7.103 The implementation of hedonic methods to estimate quality adjustments for matched items which are no longer available is considered below. Consider items *l*, *m* and *n* where item *l* is available in periods *t* and *t*+2, the "old" item *m* is only available in period *t* and the replacement item *n* only in period *t*+2. The items are defined by their *z* quality characteristics, item *m* for example being z_m^t and the price of item *m* in period *t* is p_m^t as depicted below. There is no problem with comparing the prices p_l^t and p_l^{t+2} of matched items with characteristics z_l^t with z_l^{t+2}, for they have the same *l* quality characteristics. But there is a problem with item *m*. A hedonic *imputation* approach would predict the price of item *m*'s characteristics in period *t*+2 at the characteristic prices taken from a hedonic regression estimated in period *t*+2, i.e. $\hat{p}_m^{t+2}$.

In this case, item *m*'s characteristics are held constant in the comparison $\hat{p}_m^{t+2} / p_m^t$. A similar exercise can be conducted for the replacement item *n* using $p_n^{t+2} / \hat{p}_n^t$. In this comparison, item *n*'s characteristics are held constant and compared at period *t*+2 and period *t* prices. These imputation approaches are outlined below. Yet there is a second approach, an *adjustment* one. Here the characteristics of the replacement item *n* are identified and compared with those of the old item *m*, $\left(z_n^{t+2} - z_m^t\right)$, and estimated coefficients from hedonic equations used to estimate the value of the changes. These two approaches, hedonic imputations and hedonic adjustments, are considered below in further detail. This "patching" of missing prices is quite different from the use of hedonic price indices discussed in paragraphs 7.132 to 7.149 and Chapter 21. These use hedonic regressions to provide hedonic price indices of overall quality-adjusted prices using a sample of all of the data in each period with no patching. The "patching" of missing prices is a partial application of the hedonic approach, used in imputations for missing items or on non-comparable replacements for missing items when the matched models approach is being used and an item's price is missing.

7.104 *Hedonic imputation: Predicted vs. actual* In this approach a hedonic regression of the natural logarithm of the price of model *i* in period *t* on its characteristics set z_{ki}^{t} is estimated for each month, using the equation:

$$\ln\ p_i^t = \beta_0^t + \sum_{k=1}^{K} \beta_k^t z_{ki}^t + \varepsilon_i^t \qquad (7.24)$$

Say the price of an item *m* available in January (period *t*) is unavailable in March (period *t*+2). The price of item *m* can be predicted for March by inserting the characteristics of the old unavailable item *m* into the estimated regression equation for March, and similarly for successive months. The predicted price for this old item in March and price comparison with January (period *t*) are given, respectively, by:

$$\hat{p}_m^{t+2} = \exp\left[\beta_0^{t+2} + \sum_k \beta_k^{t+2} z_{k,m}^t\right] \text{ and } \hat{p}_m^{t+2} / p_m^t \qquad (7.25a)$$

That is, the old model's price is predicted for period *t*+2 and patched in. In the example in Table 7.1(a), $\hat{p}^{23}, \hat{p}^{24}$, etc. and $\hat{p}^{63}$, $\hat{p}^{64}$, etc. would be estimated and compared with p^{21} and p^{61} respectively. The blanks for items 2 and 6 in Table 7.1(a) would be effectively filled in by the estimated price from the regression equation.

7.105 An alternative procedure is to select for each unavailable *m* item a replacement item *n*. In this case the price of *n* in period *t*+2 is known, and a predicted price for *n* in period *t* is required. The predicted price for the new item and the required price comparison are:

$$\hat{p}_n^t = \exp\left[\beta_0^t + \sum_k \beta_k^t \ z_{k,m}^{t+2}\right] \text{ and } p_n^{t+2} / \hat{p}_n^t \qquad (\mathbf{7}.25b)$$

That is, the new model's price is adjusted. In this case the characteristics of item *n* are inserted into the right-hand side of an estimated regression for period *t*. The price comparisons of equation (7.25a) may be weighted by w_m^t, as would those of its replaced price comparison in equation (**7**.25b).

7.106 Another option is to take the geometric mean of the formulations in equations (**7**.25a) and (7.25b) on grounds analogous to those discussed in Chapter 15 and by Diewert (1997) with regard to similar index numbers.

7.107 *Hedonic imputation: Predicted vs. predicted*. This approach uses predicted values for, say, item *n* in both periods, e.g., $\hat{p}_n^{t+2} / \hat{p}_n^t$. Consider a misspecification problem in the hedonic equation. For example, there may be an interaction effect between a brand dummy and a characteristic – say for Dell and speed in the example in Table 7.4. Possession of both characteristics may be worth more in terms of price (in a semi-logarithmic form) than their separate individual components (for evidence of interactive effects see Curry et al., 2000).

The use of $p_n^{t+2} / \hat{p}_n^t$ would be misleading since the actual price in the numerator would incorporate the 5 per cent premium, while the one predicted from a straightforward semi-logarithmic form would not. It is stressed that, in adopting this approach, a recorded actual price is being replaced by an imputation. This is not desirable, but neither is the form of bias discussed above. Diewert (2002e) considers a similar problem and suggests an adjustment to

bring the actual price back in line with the hedonic one.
The comparisons using predicted values in both periods are given as:

$\hat{p}_n^{t+2} / \hat{p}_n^t$ for the new item

$\hat{p}_m^{t+2} / \hat{p}_m^t$ for the disappearing or old item, or

$$\left[\left(\hat{p}_n^{t+2} / \hat{p}_n^t\right)\left(\hat{p}_m^{t+2} / \hat{p}_m^t\right)\right]^{1/2} \qquad (7.26)$$

as a geometric mean of the two.

7.108 *Hedonic adjustments.* In this approach a replacement item is used and any differences between the *k* characteristics of the replacement *n* in, for example, period *t*+2 and *m* in period *t* are ascertained. A predicted price for *m* adjusted to be compatible with *n* is estimated for period *t*, i.e., $\hat{p}_m^{t+2}$ and is compared with the actual price, p_m^t where

$$\hat{p}_m^{t+2} \equiv p_n^{t+2} \exp\left[-\sum_k \beta_k^{t+2}(z_{nk}^{t+2} - z_{mk}^t)\right] \qquad (7.27a)$$

or alternatively, a predicted price for *n* adjusted to be compatible with *m* is estimated for period *t*, i.e. $\hat{p}_n^t$ is compared with the actual price, p_n^{t+2} where

$$\hat{p}_n^t \equiv p_m^t \exp\left[\sum_k \beta_k^t (z_{nk}^{t+2} - z_{mk}^t)\right] \qquad (7.27b)$$

The adjustments here are undertaken using predicted values. However, unlike the formulations in equation (7.27b) for example, $\hat{p}_n^t$ may be estimated by applying the subset of the *k* characteristics that distinguished *m* from *n* to their respective implicit prices in period *t* estimated from the hedonic regression, and adjusting the price of p_m^t. For example, if the nearest replacement for item 2 is item 3, then the characteristics that differentiated item 3 from item 2 are identified and the price in the base period, p^{31}, is estimated by adjusting p^{21} using the appropriate coefficients from the hedonic regression in that month. For example, for washing machines, if item 2 had a spin speed of 800 rpm and item 3 a spin speed of 1,100 rpm, other things being equal, the shadow price of the 300 rpm differential would be estimated from the hedonic regression and p^{21} would be adjusted for comparison with p^{33}. Note that if the *z* variables in the characteristic set are perfectly independent of each other, the results from this approach will be similar to those from equation (7.25). This is because interdependence between the variables on the right-hand side of the hedonic equation – multicollinearity – leads to imprecise estimates of the coefficients (see Chapter 21). Hedonic imputations and adjustments of the form (7.25b) and (7.27b) have an advantage over their counterparts (7.25a) and (7.27a) since the regression equation does not have to be updated in each period. However, (7.25b) and (7.27b) effectively compare a constant fixed basket of current period characteristics while (7.25a) and (7.27a) compare a fixed basket of price reference period characteristics. There is no reason to prefer one to the other and if the difference or spread between the two indices is large, this is reason for caution over the use of one against a geometric mean of the two. Regular updating of hedonic regressions would be likely to minimize spread.

7.109 *Hedonic: Indirect adjustment.* An indirect adjustment may be made for the current period, which only requires the hedonic regression to be estimated in the base period *t*, using:

$$\frac{p_n^{t+2}}{p_m^t} \div \frac{\hat{p}_n^t}{\hat{p}_m^t} \qquad (7.28)$$

The first term is the change in price between the old and replacement items in periods t and t+2 respectively. But the quality of the item has also changed, so this price change needs to be divided by a measure of the change in quality. The second term uses the hedonic regression in period t in both the numerator and denominator. The coefficients – the shadow prices of each characteristic – are held constant. The predicted prices nevertheless differ because different quantities of the characteristics are being inserted into the numerator and denominator: the characteristics of the replacement item n in the former and the old item m in the latter. The measure is the change in price after removing (by division) the change in quantity of characteristics for each item at a constant period t price. Of course, conceptually, the constant valuation by a period t+2 regression would be equally valid and a geometric mean of the two ideal. However, if hedonic regressions cannot be run in real time this is a compromise. As the spread between the current and base period results increases, its validity decreases. As such, the regression estimates should be updated regularly using old and current period estimates and results compared retrospectively as a check on the validity of the results.

Limitations of the hedonic approach

7.110 The limitations of the hedonic approach should be borne in mind. Some points are summarized below (see also Chapter 21). First, the approach requires statistical expertise for the estimation of the equations. The availability of user-friendly software with regression facilities makes this less problematic. Statistical and econometric software carries a range of diagnostic tests to help judge if the final formulation of the model is satisfactory. These include $\overline{R}^2$ as a measure of the overall explanatory power of the equation, and F-test and t-test statistics to enable tests to be conducted as to whether the differences between the coefficients of the explanatory variables are jointly and individually different from zero at specified levels of statistical significance. Most of these statistics make use of the errors from the estimated equation. The regression equation can be used to predict prices for each item by inserting the values of the characteristics of the items into the explanatory variables. The differences between the actual prices and these predicted results are the residual errors. Bias or imprecise, and thus misleading, results may arise from a range of factors including heteroscedasticity (non-constant variances in the residuals suggesting non-linearities or omission of relevant explanatory variables), a non-normal distribution for the errors, and multicollinearity, where two or more explanatory variables are related. The latter in particular has been described as the "bane of hedonic regressions" (Triplett, 1990). Such econometric issues have been discussed in the context of hedonic regressions (Berndt, 1991; Berndt et al., 1995; Triplett, 1990; Gordon, 1990; Silver, 1999; and in Chapter 21) and more generally by Kennedy (1998) and Maddala (1988). For the reasons discussed above, when multicollinearity is suspected, the use of predicted values rather than individual coefficients is advised.

7.111 Second, the estimated coefficients should be updated regularly. If the adjustment is to the old model, then the price comparison is between the price of the old model in some reference period adjusted for the quality difference between the old and new models, using coefficients from an estimated hedonic equation in the price reference period as estimates of the value of such differences, as in (7.27b). There is, at first sight, no need to update the estimated coefficients each month. Yet the valuation of a characteristic in the price reference period may be quite out of line with its valuation in the new period. For example, a feature may be worth an additional 5 per cent in the reference period instead of 10 per cent in the current period, because it might have been introduced at a discount at that point in its life cycle to encourage usage. Continuing to use the coefficients from some far-off period to make adjustments to prices in the current period is akin to using out-of-date base period weights. The comparison may be well defined, but have little meaning. If price adjustments for quality differences are being made to the old item in the price reference period using hedonic estimates from that period, then there is a need to update the estimates if they are considered to be out of date, say because of changing tastes or technology, and splice the new estimated comparisons onto the old. The regular updating of hedonic estimates when using imputations or adjustments is thus recommended, especially when there is evidence of parameter instability over time. Ideally a geometric mean of either (7.25a) and (7.25b) or of (7.27a) and (7.27b) should be used, but this requires an updating of hedonic regressions in real time.

7.112 Third, the sample of prices and characteristics used for the hedonic adjustments should be suitable for the purpose. If they are taken from a particular outlet or outlet type, trade source or web page and then used to adjust non-comparable prices for items sold in quite different outlets, then there must at least be an intuition that the marginal utilities for characteristics are similar between the outlets. A similar principle applies for the brands of items used in the sample for the hedonic regression. It should be borne in mind that high $\overline{R}^2$ statistics do not alone ensure reliable results. Such high values arise from regressions in periods prior to their application and indicate the proportion of variation in prices across many items and brands. They are not in themselves a measure of the prediction error for a particular item, sold in a specific outlet, of a given brand in a subsequent period, though they can be an important constituent of this.

7.113 Fourth, there is the issue of functional form and the choice of variables to include in the model. Simple functional forms generally work well. These include linear, semi-logarithmic (logarithm of the left-hand side) and double-logarithmic (logarithms of both sides) forms. Such issues are discussed in Chapter 21. The specification of a model should include all price-determining characteristics. Some authors advise quite simple forms with only the minimum number of variables, as long as the predictive capacity is high (Koskimäki and Vartia, 2001). Shepler (2000) included 33 variables in her hedonic regressions of refrigerators – a fairly homogeneous product. These included nine dummy variables for brand and four for colour, five types of outlets, three regions as control variables, and 11 characteristics including capacity, types of ice-maker, energy-saving control, extra drawers, sound insulation, humidifier and fil-tration device. Typically, a study would start with a large number of explanatory variables and a general econometric model of the relationship, while the final model would be more specific, having dropped a number of variables. The dropping of variables would depend on the result of experimenting with different formulations, and seeing their effects on diagnostic test statistics, including the overall fit of the model and the accordance of signs and magnitudes of coefficients with prior expectations. Reese (2000), for example, started with a hedonic regression for United States college textbooks which included about 50 explanatory variables, subsequently reduced to 14 such variables with little loss of explanatory power.

7.114 Finally, Bascher and Lacroix (1999) list several requirements for successful design and use of hedonic quality adjustment in the consumer price index, noting that these require heavy investments over a long period involving:

- intellectual competencies and sufficient time to develop and re-estimate the model, and to employ it when products are replaced;

- access to detailed, reliable information on product characteristics;
- a suitable organization of the infrastructure for collecting, checking and processing information.

7.115 Hedonic methods may also improve quality adjustment in the consumer price index by indicating which product attributes do *not* appear to have material impacts on price. That is, if a replacement item differs from the old item only in characteristics that have been rejected as price-determining variables in a hedonic study, this would support a decision to treat the items as comparable or equivalent to and include the entire price difference, if any, as pure price change. Care has to be exercised in such analysis because a feature of multicollinearity in regression estimates is that the imprecision of the parameter estimates may give rise to statistical tests that do not reject null hypotheses that are false, i.e., they do not find significant parameter estimates that are significant. The results from such regressions can nonetheless provide valuable information on the extent to which different characteristics influence price variation, and this in turn can help in the selection of replacement items. Enhanced confidence in item substitution and the quality adjustment of prices that arises from using the hedonic approach, and the parallel reduction in reliance on "linking", has been cited as a significant benefit in terms of the reliability of the measurement of price changes for apparel in the United States consumer price index (Reinsdorf, Liegey and Stewart, 1996). The results from hedonic regressions have a role to play in identifying price-determining characteristics and may be useful in the design of quality checklists in price collection (Chapter 6).

Choice between quality adjustment methods

7.116 Choice of method for quality adjustments to prices is not straightforward. The analyst must consider the technology and market for each commodity and devise appropriate methods. This is not to say the methods selected for one product area will be independent of those selected for other areas. Expertise built up using one method may encourage its use elsewhere, and intensive use of resources for one commodity may lead to less resource-intensive methods for others. The methods adopted for individual product areas may vary between countries as access to data, relationships with the outlet managers, resources, expertise and features of the production, and market for the product vary. Guidelines on choice of method arise directly from the features of the methods outlined above. A good understanding of the methods, and their implicit and explicit assumptions, is essential to the choice of an appropriate method.

7.117 Figure 7.3 provides a guide to the decision-making process. Assume that the matched models method is being used. If the item is matched for re-pricing, there being no change in the specification, no quality adjustment is required. This is the simplest of procedures. However, a caveat applies. If the item belongs to a high-technology industry where model replacement is rapid, the matched sample may become unrepresentative of the universe of transactions. Alternatively, matching may be under a chained framework, where prices of items in a period are matched to those in the preceding period to form a link. A series of successive links of matched comparisons combined by successive multiplication makes up the chained matched index. Or hedonic indices may be used which require no matching. The use of such methods is discussed in paragraphs 7.132 to 7.149. At the very least, attention should be directed to more regular item re-sampling. Continued long-run matching would deplete the sample and an alternative framework to long-run matching would be required.

7.118 Consider a change in the quality of an item and assume that a replacement item is available. The selection of a comparable item to the same specification and the use of its price as a *comparable replacement* require that none of the price difference is attributable to quality. It also requires confidence that all price-determining factors are included in the specification. The replacement item should also be representative and account for a reasonable proportion of sales. Caution is required when replacing near obsolete items with unusual pricing at the end of their life cycles with similar ones that account for relatively low sales, or with ones that have quite substantial sales but are at different points in their cycle. Strategies for ameliorating such effects are discussed below and in Chapter 8, including early substitutions before pricing strategies become dissimilar.

7.119 Figure 7.3 illustrates the case where quality differences can be quantified. *Explicit estimates* are generally considered to be more reliable, although they are also more resource intensive, at least initially. Once an appropriate methodology has been developed, they can often be easily replicated. General guidelines are more difficult here as the choice depends on the host of factors discussed above, which are likely to make the estimates more reliable in each situation. Central to all of this is the quality of the data upon which the estimates are based. If reliable data are unavailable, subjective judgements may be used. Product differences are often quite technical and very difficult to specify and quantify. The reliability of the method depends on the expertise of the experts and the variance in opinions. Estimates based on objective data are thus preferred. Good *production cost* estimates in industries with stable technologies and identifiable constant retail mark-ups and where differences between the old and replacement items are well specified and exhaustive are, by definition, reliable. Estimates of the retail mark-up are, however, prone to error and the *option cost* approach is generally preferable. This requires that the old and new items differ by easily identifiable characteristics which are or have been separately priced as options.

7.120 The use of *hedonic regressions* for partial patching is most appropriate where data on price and characteristics are available for a range of models and where the characteristics are found to predict and explain price variability well in terms of a priori reasoning and econometric terms. Their use is appropriate where the cost of an option or change in characteristics cannot be separately identified and has to be gleaned from the prices of items sold with different specifications in the market. The estimated regression coefficients are the estimate of the contribution to price of a unit change in a characteristic, having controlled for the effects of variations in the quantities of other characteristics. The estimates are particularly suited to valuing changes in the quality of an item when only a given set of characteristics changes and the valuation is required for changes in these characteristics only. The results from hedonic regressions may be used to target the salient characteristics for item selection. The synergy between the selection of prices according to characteristics defined as price determining by the hedonic regression, and their subsequent use for quality adjustment, should reap rewards. The method should be applied where there are high ratios of non-comparable replacements and where the differences between the old and new items can be well defined by a large number of characteristics.

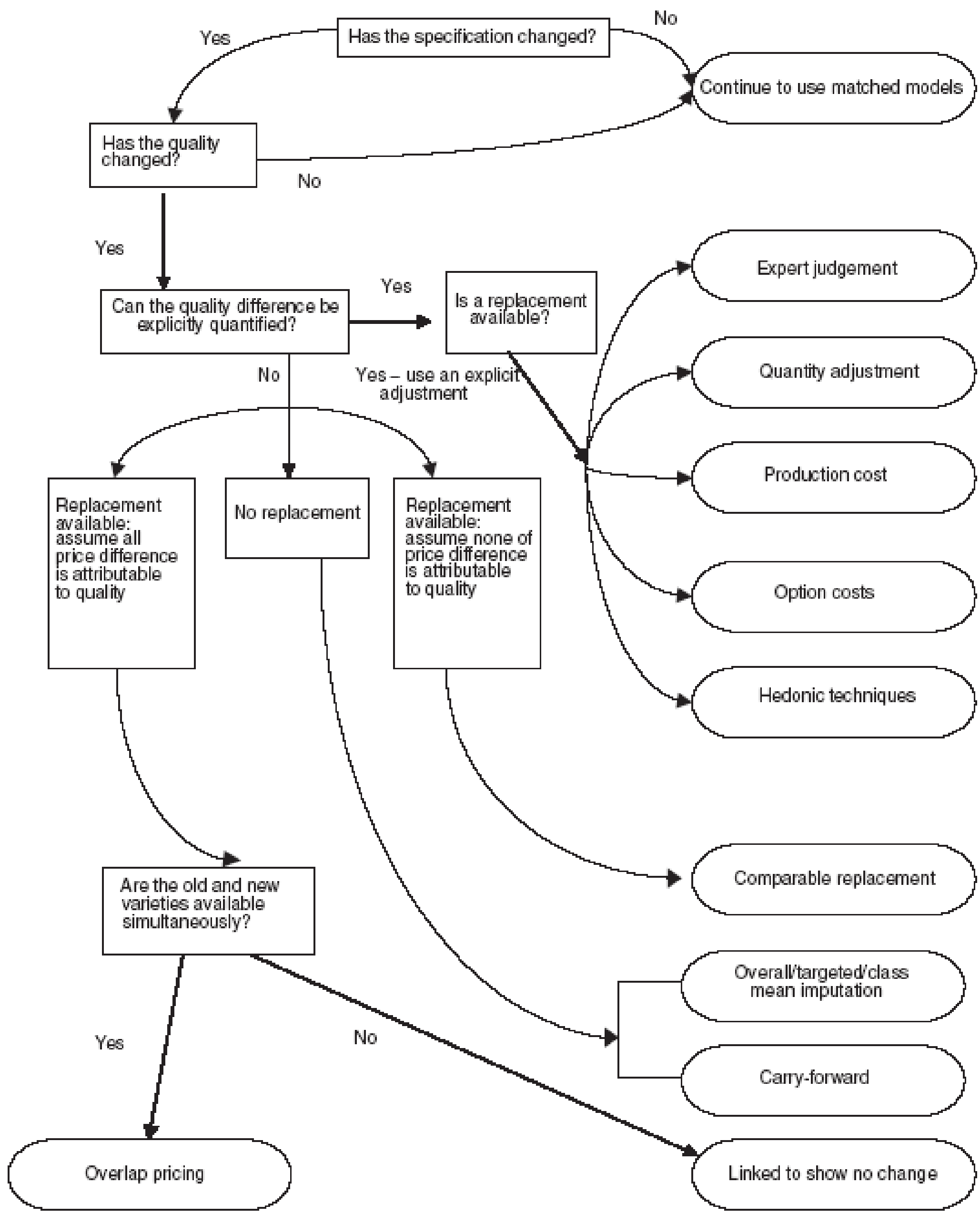
7.121 If explicit estimates of quality are unavailable, and no replacement items are deemed appropriate, then imputations may be used. The use of *imputations* has much to commend it resource-wise. It is relatively easy to employ – though some verification of the validity of the implicit assumptions might be appropriate. It requires no judgement (unless targeted) and is therefore objective. Targeted mean imputation is preferred to overall mean imputation as long as the sample size upon which the target is based is adequate. Class mean imputation is preferred when models at the start of their life cycles are replacing those around the end of their life cycle, although the approach requires faith in the adequacy of the explicit and comparable replacements being made.

7.122 Bias from using imputation is directly related to the proportion of missing items and the difference between quality-adjusted prices of available matched items and the quality-adjusted prices of unavailable ones (see Table 7.2 on page 110). The nature and extent of the bias depends on whether short-run or long-run imputations are being used (the former being preferred) and on market conditions (see paragraphs 7.159 to 7.173). Imputation, in practical terms, produces the same result as deletion of the item. The inclusion of imputed prices may give the illusion of larger sample sizes. Imputation is less likely to introduce bias where the proportion of missing prices is low. Table 7.2 can be used to estimate likely error margins arising from its use and a judgement can be made as to whether they are acceptable. The use of imputation across many products need not necessarily compound the errors since, as noted in the above discussion of this method, the direction of bias need not be systematic. It is cost-effective for product areas with a large number of missing items because of its ease of use. But the underlying assumptions required by imputation must be very carefully considered if it is widely used. Imputation should by no means be the overall catch-all strategy, and statistical agencies are advised against its use as a default device without due consideration of the nature of the markets, the possibility of targeting the imputation and the viability of estimates from the sample sizes involved if such targeting is employed.

7.123 If the old and replacement items are available simultaneously, and if the quality difference cannot be quantified, an implicit approach can be used whereby the price difference between the old and replacement items in a period in which they both exist is assumed to be attributable to quality. This *overlap* method, in replacing the old item by a new one, takes the ratio of prices in a period to be a measure of their quality difference. It is implicitly used when new samples of items are taken. The assumption of relative prices equating to quality differences at the time of the splice is unlikely to hold if the old and replacement items are at different stages in their life cycles and different pricing strategies are used at these stages. For example, there may be deep discounting of the old item to clear inventories, and price skimming of market segments that will purchase new models at relatively high prices. As with comparable replacements, early substitutions are advised so that the overlap is at a time when items are at similar stages in their life cycles.

7.124 For the reasons discussed, the use of the *linked to show no change* method and the *carry-forward* method is not generally advised for making quality adjustment imputations, unless the implicit assumptions are deemed to be valid.

Figure 7.3 Flowchart for making decisions on quality change



Source: Chart developed from a version by Fenella Maitland-Smith and Rachel Bevan, OECD; see also a version in Triplett (2002).

High-technology and other sectors with a rapid turnover of models

7.125 The measurement of price changes of items unaffected by quality changes is primarily achieved by matching models, the above techniques being applicable when the matching breaks down. But what of industries where the matching breaks down on a regular basis because of the high turnover in new models of different qualities to the old ones? The matching of prices of identical models over time, by its nature, is likely to lead to a seriously depleted sample. There is both a dynamic universe of all items consumed and a static universe of the items selected for repricing (Dalén, 1998a). If, for example, the sample is initiated in December, by the subsequent May the static universe will be matching prices of those items available in the static universe in both December and May, but will omit the unmatched new items introduced in January, February, March, April and May, and the unmatched old ones available in December but unavailable in May. Two empirical questions show whether there will be any significant bias. First, is sample depletion substantial? Substantial depletion of the sample is a necessary condition for such bias. Second, are the unmatched new and unmatched old items likely to have quality-adjusted prices that substantially differ from those of the matched items in the current and the base periods?

7.126 The matching of prices of identical models over time may lead to the monitoring of a sample of models that is increasingly unrepresentative of the population of transactions. Some of the old models that existed when the sample was drawn are not available in the current period; and new models that enter the sample are not available in the base period. It may be that the models that are going out have relatively low prices, while the entrants have relatively high ones. By ignoring these prices, a bias is being introduced. Using old low-priced items and ignoring new high-priced ones has the effect of biasing the index downwards. In some industries, the new item may be introduced at a relatively low price and the old one may become obsolete at a relatively high price, serving a minority segment of the market (Berndt et al., 2003). In this case, the bias would take the opposite direction. The nature of the bias will depend on the pricing strategies of firms for new and old items.

7.127 This sampling bias exists for most products. Our concern here, however, is with product markets where the statistical agencies are finding the frequency of new item introductions and old item obsolescence sufficiently high that they may have little confidence in their results. First, some examples of such product markets will be given and then two procedures will be considered: the use of hedonic price indices (as opposed to the partial, hedonic patching discussed above) and chaining.

Some examples

7.128 Koskimäki and Vartia (2001) attempted to match prices of models of personal computers (PCs) over three two-month periods (spring, summer and autumn) using a sample of prices collected as part of standard price collection for the Finnish consumer price index. Of the 83 spring prices, only 55 matched comparisons could be made with the summer prices, and then only 16 continued through to the autumn. The sample of matched pairs became increasingly rapidly biased: of the 79 models in the autumn, the 16 matched ones had a mean processor speed of 518 MHz compared with 628 MHz for the remaining 63 unmatched ones; the hard disk sizes were, respectively, 10.2 and 15.0 Gigabytes, and the percentages of high-end processors (Pentium III and AMD Atl.) were 25 per cent and 49.2 per cent, respectively. Hardly any change in *matched* prices was found over this six-month period, while a hedonic regression analysis using all of the data found quality-adjusted price falls of around 10 per cent. Instructions to price collectors to hold onto models until forced replacements are required may thus lead to a sample that is increasingly unrepresentative of the population and is biased towards technically inferior variants. In this instance, the hedonic price changes fell faster since the newer models became cheaper for the services supplied.

7.129 Kokoski et al. (1999) used hedonic regressions in an empirical study of inter-area price comparisons of food products across urban areas in the United States using the United States consumer price index data. They found a negative sign on the coefficients of dummy variables for whether or not the sample items were from newly rotated samples (dummy variable = 1) or samples prior to rotation (dummy variable = 0). This indicated that quality-adjusted prices were lower for the newly included items compared with the quality-adjusted prices of the old items.

7.130 Silver and Heravi (2002) found evidence of sample degradation when matching prices of United Kingdom washing machines over a year. By December, only 53 per cent of the January basket of model varieties was used for the December/January index, although this accounted for 81.6 per cent of January expenditure. Models of washing machines with lower sales values dropped out faster. However, the remaining models in December accounted for only 48.2 per cent of the value of transactions in December. The active sample relating to the universe of transactions in December had substantially deteriorated. The prices of unmatched and matched models were found to differ, as were their vintage and quality. Even when prices were adjusted for quality using hedonic regressions, prices of unmatched old models were found to be lower than matched ones, there also being evidence of higher prices for unmatched new models. Quality-adjusted prices fell faster for the matched sample than for the full sample: about 10 per cent compared with about 7 per cent. Residuals from a common hedonic surface and their leverage were also examined. The residuals from unmatched new models were higher than matched ones, while residuals from unmatched old models were much lower. Unmatched observations had nearly twice the (unweighted) leverage as matched ones – their influence in the estimation of the parameters of the regression equation was much greater, and their exclusion more serious.

7.131 The above studies demonstrate how serious sample degradation can occur and how unmatched excluded items may be quite different from included ones. Two procedures for dealing with such situations will be considered: the use of hedonic price indices (as opposed to the partial, hedonic patching discussed above) and chaining. Both rely on a data set of a representative sample of items and their characteristics in each period. Price collectors might use a checklist of characteristics in gathering the data (Merkel, 2000). They will be asked to collect prices and characteristics of more than one item in each store, the items being the major or typical ones sold. If a new item is introduced which has or is likely to have substantial sales, then it is included as a replacement or even addition, and its characteristics are marked off against a checklist of salient characteristics. The list will be developed at the time of initiating the sample, and updated as required. Alternatively, market research agencies, web pages and trade associations may also be able to provide lists of models and their prices. Nevertheless, there is a need to collect transaction prices, as opposed to list prices.

Hedonic price indices

7.132 It is important to distinguish between the use of hedonic regressions to make adjustments for quality differences when a non-comparable substitute is used, as in paragraphs 7.90 to 7.115, and their use in their own right as *hedonic price indices,* which are measures of quality-adjusted price changes. Hedonic price indices are suitable when the pace and scale of replacements of items are substantial because, first, an extensive use of quality adjustments may lead to errors and, second, the sampling will be from a matched/replacement universe likely to be biased. With new models being continually introduced and old ones disappearing, the coverage of a matched sample may deteriorate and bias may be introduced as the price changes of new/old models differ from those of the matched ones. What is required is a sample to be drawn in each month and price indices constructed; but instead of controlling for quality differences by matching, they will be controlled for, or "partialled out", in the hedonic regression. Note that all the indices described below use a fresh sample of the data available in each period. If there is a new item in a period, it is included in the data set and its quality differences controlled for by the regression. Similarly, if old items drop out, they are still included in the data for the indices in the periods in which they exist.

Paragraphs 7.110 to 7.115 stress the need for caution in the use of hedonic regressions for quality adjustments; some of the theoretical and econometric aspects are considered in Chapter 21. This need for caution extends to the use of the results from hedonic indices, and the discussion is not repeated here for the sake of brevity.

7.133 In Chapter 17, theoretical price indices are defined and practical index number formulae are considered as bounds or estimates of these indices. Theoretical index numbers are also defined in Chapter 21 to include goods made up of tied characteristics, so something can be said about how such theoretical indices relate to different forms of hedonic indices. A number of forms are considered in Chapter 21; they are summarized below.

7.134 *Hedonic functions with dummy variables for time*. The sample covers the two time periods being compared, say t and t+2, and does not have to be matched. The hedonic formulation regresses the price of item i, p_i, on the k=2,….K characteristics of the items z_{ki}. A single regression is estimated on the data in the two time periods compared, the equation also including a dummy variable D^{t+2} being 1 in period t+2, zero otherwise:

$$\ln p_i = \beta_0 + \beta_1 D^{t+2} + \sum_{k=2}^{K} \beta_k z_{ki} + \varepsilon_i \qquad (7.29)$$

The coefficient β_1 is an estimate of the quality-adjusted price change between period t and period t+2. It is an estimate of the change in the logarithm of price, having controlled for the effects of variation in quality via $\sum_{k=2}^{K} \beta_k z_{ki}$. Note that an adjustment is required for $\beta_{1:}$ the addition of ½(standard error)^2 of the estimate, as discussed in Goldberger (1968) and Teekens and Koerts (1972). Two variants of equation (7.28) are considered. The first is the direct *fixed base version*, that compares period t with t+2 as outlined; January–February, January–March, etc. The second is a rolling *chained version* evaluated for period t with t+1; then again for t+1 with t+2, the links in the chain being combined by successive multiplication. A January–March comparison, for example, would be the January–February index multiplied by the February–March one. There is, of course, a *fully constrained version:* a single constrained regression for, say, January to December with dummy variables for each month, but this is impractical in real time since it requires data on future observations.

7.135 The above approach uses the dummy variables on time to compare prices in period 1 with prices in each subsequent period. In doing so, the β parameters are constrained to be constant over the period being compared. A fixed base, bilateral comparison using equation (7.29) makes use of the constrained parameter estimates over the two periods and, given an equal number of observations in each period, is a form of a symmetric average. A *chained* formulation would estimate $I_{1,4}$, for example, as: $I_{1,4} = I_{1,2} * I_{2,3} * I_{3,4}$. In each binary comparison for matched data, equal weight is also given to the data in each period.

7.136 There is no explicit weighting in these formulations and this is a serious disadvantage. In practice, "cut-off" sampling might be employed to include only the most important items. If sales data are available, a weighted least squares (WLS) estimator should be used, as opposed to an ordinary least squares (OLS) estimator. It is axiomatic in normal index number construction that the same weight should not be given to each price comparison, since some items may account for much larger sales revenues than others. The same consideration applies to these hedonic indices. Diewert (2002e) has argued for a preference for sales *value* weights over quantity weights. Two items may have sales equal to the same quantity, but if one is priced higher than another, its price changes should be accordingly weighted higher for the result to be meaningful in an economic sense. Additionally, Diewert (2002e) has shown that value *shares* should form the weights, since values will increase, in say period t+2, with prices, the residuals and their variance thus being higher in period *t*+2 than in *t*. This heteroscedasticity is an undesirable feature of a regression model, resulting in increased standard errors. Silver (2002) has further shown that a WLS estimator does not purely weight the observations by their designated weights, the actual influence given being also the result of a combination of the residuals and the leverage effect. The latter is higher as the characteristics of the observations diverge from the average characteristics of the data. Silver suggests that observations with relatively high leverage and low weights be deleted and the regression re-run.

7.137 *Period-to-period hedonic indices.* An alternative approach for a comparison between periods *t* and *t+2* is to estimate a hedonic regression for period *t+2,* and insert the values of the characteristics of each model existing in period *t* into the period *t+2* regression to predict, for each item, its price. This would generate predictions of the prices of items existing in period *t* based on their z_i^t characteristics, at period *t+2* shadow prices, $\hat{p}_i^{t+2}(z_i^t)$. These prices (or an average) can be compared with the actual prices (or the average of prices) of models in period *t*, $p_i^t(z_i^t)$ as, for example, a Jevons hedonic base period index:

$$P_{JHB} = \frac{\left[\prod_{i=1}^{N^t} \hat{p}_i^{t+2}(z_i^t)\right]^{1/N^t}}{\left[\prod_{i=1}^{N^t} p_i^t(z_i^t)\right]^{1/N^t}} = \frac{\left[\prod_{i=1}^{N^t} \hat{p}_i^{t+2}(z_i^t)\right]^{1/N^t}}{\left[\prod_{i=1}^{N^t} \hat{p}_i^t\right]^{1/N^t}} = \frac{\left[\prod_{i=1}^{N^t} \hat{p}_i^{t+2}(z_i^t)\right]^{1/N^t}}{\left[\prod_{i=1}^{N^t} p_i^t\right]^{1/N^t}} \qquad (7.30a)$$

7.138 Alternatively, the characteristics of models existing in period *t*+2 can be inserted into a regression for period *t*. Predicted prices of period *t*+2 items generated at period *t* shadow prices, $p_i^t(z_i^{t+2})$, are the prices of items existing in period *t*+2 estimated at period *t* prices and these prices (or an average) can be compared with the actual prices (or the average of prices) in period *t*+2, $p_i^{t+2}(z_i^{t+2})$; a Jevons hedonic current period index is:

$$P_{JHC} = \frac{\left[\prod_{i=1}^{N^{t+2}} p_i^{t+2}(z_i^{t+2})\right]^{1/N^{t+2}}}{\left[\prod_{i=1}^{N^{t+2}} p_i^t(z_i^{t+2})\right]^{1/N^{t+2}}} = \frac{\left[\prod_{i=1}^{N^{t+2}} \hat{p}_i^{t+2}\right]^{1/N^{t+2}}}{\left[\prod_{i=1}^{N^{t+2}} p_i^t(z_i^{t+2})\right]^{1/N^{t+2}}} = \frac{\left[\prod_{i=1}^{N^{+2t}} p_i^{t+2}\right]^{1/N^{t+2}}}{\left[\prod_{i=1}^{N^{t+2}} p_i^t(z_i^{t+2})\right]^{1/N^{t+2}}} \qquad (7.30b)$$

7.139 For a fixed base bilateral comparison using either equation (7.30a) or equation (7.30b), the hedonic equation is only estimated for one period, the current period *t*+2 in equation (7.30a) and the base period *t* in equation (7.30b). For reasons analogous to those explained in Chapters 15, 16 and 17, a symmetric average of these indices would have some theoretical support.

7.140 Note that a geometric mean of (7.30) uses all the data available in each period, as does the hedonic index using a time dummy variable in (7.29). If in (7.29) there is a new item in, say, period t+2, it is included in the data set and its quality differences controlled for by the regression. Similarly, if old items drop out, they are still included in the indices in the periods in which they exist. This is part of the natural estimation procedure, unlike using matched data and hedonic adjustments on non-comparable replacements when items are no longer available.

7.141 With the dummy variable approach, there is no explicit weighting in its formulation in (7.29), and this is a serious disadvantage. In practice, cut-off sampling might be employed to include only the most important items; or if expenditure data are available, a WLS as opposed to an OLS estimator might be used, with expenditure value shares as weights, as discussed in Appendix 21.1 to Chapter 21.

7.142 *Superlative and exact hedonic indices (SEHI).* In Chapter 17, Laspeyres and Paasche bounds are defined on a theoretical basis, as are superlative indices, which treat both periods' data symmetrically. These superlative formulae, in particular the Fisher index, are also seen in Chapter 16 to have desirable axiomatic properties. Furthermore, the Fisher index is supported by economic theory as a symmetric average of the Laspeyres and Paasche bounds, being found to be the most suitable such average on axiomatic grounds. The Törnqvist index is seen to be best from the stochastic viewpoint, and also does not require strong assumptions for its derivation from the economic approach as a superlative index. The Laspeyres and Paasche indices are found to correspond to (be exact for) underlying Leontief aggregator functions with no substitution possibilities, while superlative indices are exact for flexible functional forms, including the quadratic and translogarithmetic forms for the Fisher and Törnqvist indices, respectively. If data on prices, characteristics and quantities are available, analogous approaches and findings arise for hedonic indices (Fixler and Zieschang, 1992 and Feenstra, 1995). Exact theoretical bounds on a hedonic index have been defined by Feenstra (1995). Consider the theoretical index in Chapter 17, equation (17.3), but now defined only over items in terms of their characteristics z_i. The prices (and quantities) are still of items, but they are wholly defined through their characteristics $p_i(z_i)$. An arithmetic aggregation for a linear hedonic equation finds a Laspeyres upper bound (as quantities demanded *decrease* with increasing relative prices) given by:

$$\frac{\sum_{i=1}^{N} q_i^t \hat{p}_i^{t+2}}{\sum_{i=1}^{N} q_i^t p_i^t} = \sum_{i=1}^{N} s_i^t \left(\frac{\hat{p}_i^{t+2}}{p_i^t} \right) \geq \frac{C\left(u^t, p(z)^{t+2}\right)}{C\left(u^t, p(z)^t\right)} \qquad (7.31a)$$

where the right-hand-side expression is the ratio of the cost of achieving a period *t* level of utility (u^t), where utility is a function of the vector of quantities, i.e., $u^t = f\left(q^t\right)$. The price comparison is evaluated at a fixed level of period *t* quantities, and s_i^t are the shares in total value of expenditure on product *i* in period *t*, $s_i^t = q_i^t p_i^t \Big/ \sum_{j=1}^{N} q_j^t p_j^t$ and

$$\hat{p}_i^{t+2} \equiv p_i^{t+2} - \sum_{j=1}^{N} \beta_k^{t+2} (z_{jk}^{t+2} - z_{jk}^t) \qquad (7.31b)$$

are prices in period *t*+2 adjusted for the sum of the changes in each quality characteristic weighted by their coefficients derived from a linear hedonic regression. Note that the summation is over the same *i* in both periods, since replacements are included when an item is missing and equation (7.31b) adjusts their prices for quality differences.

7.143 A Paasche lower bound is estimated as:

$$\frac{\sum_{i=1}^{N} q_i^{t+2} p_i^{t+2}}{\sum_{i=1}^{N} q_i^{t+2} \hat{p}_i^{t}} = \left[\sum_{i=1}^{N} s_i^{t+2} \left(\frac{p_i^{t+2}}{\hat{p}_i^{t}} \right) \right]^{-1} \leq \frac{C\left(u^{t+2}, p(z)^{t+2}\right)}{C\left(u^{t+2}, p(z)^{t}\right)} \qquad (7.32a)$$

where $s_i^{t+2} = q_i^{t+2} p_i^{t+2} \Big/ \sum_{j=1}^{N} q_j^{t+2} p_j^{t+2}$ and $\hat{p}_i^{t} \equiv p_i^{t} + \sum_{j=1}^{N} \beta_k^{t} (z_{jk}^{t+2} - z_{jk}^{t})$ (7.32b)

which are prices in periods t adjusted for the sum of the changes in each quality characteristic weighted by its respective coefficients derived from a linear hedonic regression.

7.144 In Chapter 17 it is shown that Laspeyres P_L and Paasche P_P price indices form bounds on their respective "true" economic theoretic indexes. Using similar reasoning to that in Chapter 17 applied to equations (7.31a) and (7.32a), it can be shown that under homothetic preferences these true economic indices collapse into a single theoretical index $c(p^{t+2})/c(p^{t})$, and:

$$P_L \geq c\left(p^{t+2}\right) \big/ c\left(p^{t}\right) \geq P_P \qquad (7.33)$$

7.145 The approach is akin to that used for adjustments to non-comparable replacement items in equations (7.27a) and (7.27b), above. However, the SEHI approach first uses all the data in each period, not just the matched sample and selected replacements. Second, it uses the coefficients from hedonic regressions on changes in the characteristics to adjust observed prices for quality changes. Third, it incorporates a weighting system using data on the expenditure shares of each model and their characteristics, rather than treating each model as equally important. Finally, it has a direct correspondence to formulations defined from economic theory.

7.146 Semi-logarithmic hedonic regressions would supply a set of β coefficients suitable for use with these base and current period geometric bounds:

$$\prod_{i=1}^{N} \left(\frac{p_i^{t+2}}{\hat{p}_i^{t}} \right)^{s_i^{t+2}} \leq \frac{C\left(u, p(z)^{t+2}\right)}{C\left(u, p(z)^{t}\right)} \leq \prod_{i=1}^{N} \left(\frac{\hat{p}_i^{t+2}}{p_i^{t}} \right)^{s_i^{t}} \qquad (7.34a)$$

$$\hat{p}_i^{t} \equiv p_i^{t} \exp[\sum_{j=1}^{N} \beta_k^{t} (z_{jk}^{t+2} - z_{jk}^{t})]$$

$$\hat{p}_i^{t+2} \equiv p_i^{t+2} \exp[-\sum_{j=1}^{N} \beta_k^{t+2} (z_{jk}^{t+2} - z_{jk}^{t})] \qquad (7.34b)$$

7.147 In equation (7.34a) the two bounds on the respective theoretical indices have been shown to be brought together under an assumption of homothetic preference (see Chapter 17). The calculation of such indices is no small task. For examples of their application, see Silver and Heravi (2001a and 2003) for comparisons over time and Kokoski et al. (1999) for price comparisons across areas of a country. Kokoski et al. (1999) used a sample from a replacement universe of otherwise matched data from the United States Bureau of Labor Statistics consumer price index, though the sample benefited from rotation. Silver and Heravi (2001a and 2003) used scanner data for the universe of transactions via a two-stage procedure in which cells were defined according to major price-determining features such as all combinations of brand, outlet type and (for television sets) screen size – much like strata. There may be a gain in the efficiency of the final estimate since the adjustment is for within-strata variation, much in the way that stratified random sampling improves on simple random sampling. The average price in each matched cell could then be used for the price comparisons using equations (7.32a) and (7.34a), except that – to ensure that the quality differences in each cell from characteristics other than these major ones did not influence the price comparison – adjustments were made for quality changes using equations (7.32b) and (7.34b). This allowed all matched, old unmatched and new unmatched data to be included since, if the average price in, say, a cell of equation (7.32a) was increased because of the inclusion of a new improved item, equation (7.32b) would be used to remove such improvements, on average. Consider, for example, a brand X, 14-inch television set with stereo sound sold to multiple outlets. There might be matched cells for brand X television sets sold in multiples, but not matched cells also including stereo. The new model may have to be grouped in a cell with the brand X, 14-inch television sets sold in multiples, and the average price of the cells compared in equation (7.32a) or (7.34a), and making a quality adjustment for the stereo in the form of equation (7.32b) or (7.34b). The estimated coefficient for stereo would be derived from a hedonic equation estimated from data of other television sets, some of which possess stereo.

7.148 The above description illustrates how weighted index number formulae such as Laspeyres, Paasche, Fisher and Törnqvist might be constructed using data on prices, quantities and characteristics of an item. Silver and Heravi (2003) show that as the number of characteristics over which the summation takes place in equations (7.32a) and (7.34a) increases, the more redundant the adjustment in equations (7.32b) and (7.34b) becomes, until, when all combinations of characteristics are used in equations (7.32a) and (7.34a) as strata, the calculation extends to a matched models problem in which each cell uniquely identifies an item. For matched data, equations (7.32b) and (7.34b) serve no purpose; the aggregation in equations (7.32a) and (7.34a) would be over all items, and would reduce to the usual index number problem. Diewert (2003a), commenting on the method, explains why, when matching is relatively large, the results given are similar to those from superlative hedonic index numbers.

7.149 Weighted index number formulae might thus be constructed using data on prices, quantities and characteristics of an item when the data are not matched. This is because continuing with matched data may lead to errors from two sources: multiple quality adjustments from items no longer available and their non-comparable replacements; and sample selectivity bias from sampling from a replacement universe as opposed to a double universe.

The difference between hedonic indices and matched indices

7.150 In previous sections, the advantages of hedonic indices over matched comparisons are referred to in terms of the inclusion by the former of un-matched data. This relationship is discussed more formally here. Triplett (2002) argued and Diewert (2002e) showed that an unweighted geometric mean (Jevons) index for matched data gives the same result as a logarithmic hedonic index run on the same data. Consider the matched sample *m* and Z^{t+2} and Z^t as overall quality adjustments to the dummy variables for time in equation (7.29), that is, $\sum_{k=2}^{K} \beta_k z_{ki}$. The very first line in equation (7.35) below is shown by Aizcorbe et al. (2001) to equal the difference between two geometric means of quality-adjusted prices. The sample space $m = M^t = M^{t+2}$ is the same model in each period. Consider the introduction of a new model n introduced in period t+2 with no counterpart in t and the demise of an old model o so it has no counterpart in t+2. So M^{t+2} is composed of m and n, and M^t is composed of m and o, while M consists only of the matched models m. Silver and Heravi (2002) have shown the dummy variable hedonic comparison to now be:

$$\ln p^{t+2}/p^t = \left[m/(m+n)\sum_m \ln\left(p_m^{t+2} - Z_m\right)/m + n/(m+n)\sum_n \ln\left(p_n^{t+2} - Z_n\right)/n \right]$$
$$\times \left[m/(m+o)\sum_m \ln\left(p_m^{t} - Z_m\right)/m + o/(m+o)\sum_o \ln\left(p_o^{t} - Z_o\right)/o \right]$$
$$= \left[m/(m+n)\sum_m \ln\left(p_m^{t+2} - Z_m\right)/m - m/(m+o)\sum_m \ln\left(p_m^{t} - Z_m\right)/m \right]$$
$$\times \left[n/(m+n)\sum_n \ln\left(p_n^{t+2} - Z_n\right)/n - o/(m+o)\sum_o \ln\left(p_o^{t} - Z_o\right)/o \right] \quad (7.35)$$

7.151 Consider the second expression in equation (7.35). First, there is the change for the m matched observations. This is the change in mean prices of matched models m in period t+2 and t, adjusted for quality. Note that the weight in period t+2 for this matched component is the proportion of matched to all observations in period t+2. Similarly, for period t, the matched weight depends on how many unmatched old observations are in the sample. In the last line of equation (7.35), the change is between the unmatched new and the unmatched old mean (quality-adjusted) prices in periods t+2 and t. Thus matched methods can be seen to ignore the last line in equation (7.35) and will thus differ from the hedonic dummy variable approach in at least this respect. The hedonic dummy variable approach, in its inclusion of unmatched old and new observations, can be seen from equation (7.35) possibly to differ from a geometric mean of matched prices changes, the extent of any difference depending, in this unweighted formulation, on the proportions of old and new items leaving and entering the sample and on the price changes of old and new items relative to those of matched ones. If the market for products is one in which old quality-adjusted prices are unusually low while new quality-adjusted prices are unusually high, then the matched index will understate price changes (see Silver and Heravi, 2002 and Berndt et al., 2003 for examples). Different market behaviour will lead to different forms of bias.

7.152 If sales weights replace the number of observations in equation (7.35), then different forms of weighted hedonic indices can be derived, as explained in Chapter 21. Silver (2002) has also shown that the hedonic approach will differ from a corresponding weighted or unweighted hedonic regression in respect of the leverage and influence that the hedonic regression gives to observations.

Chaining

7.153 An alternative approach to dealing with products with a high turnover of items is to use a chained, say monthly, index instead of the long-term fixed base comparison. A chained index compares prices of items in period t with period t+1 ($\text{Index}_{t,,t+1}$) and then, as a new exercise, studies the universe of items in period t+1 and matches them with items in period t+2. These links ($\text{Index}_{t,,t+1}$ and $\text{Index}_{t,,t+2}$) are combined by successive multiplication, continuing to, say, $\text{Index}_{t,+5,t+6}$ to form $\text{Index}_{t,t+6.}$ Only items available in both period t and period t+6 would be used in a fixed base consumer price index. Consider the five items 1, 2, 5, 6 and 8 over the four months January–April, as shown in Table 7.1. The price index for January compared with February (J:F) involves price comparisons for all five items. For February–March (F:M) it involves items 1, 4, 5 and 8 and for March–April (M:A) six items: 1, 3, 4, 5, 7 and 8. The sample composition changes for each comparison as old items disappear and new items come in. Price indices can be calculated for each of these successive price comparisons using any of the unweighted formulae described in Chapter 21. The sample will grow in size when new products appear and shrink when old products disappear, changing in composition through time (Turvey, 1999).

7.154 Sample depletion may be reduced in long-run comparisons by the judicious use of replacement items. As discussed in Chapter 8, however, the replacement sample would only include a new item as and when a replacement was needed, irrespective of the number of new items entering the market. Furthermore, the re-placement item is likely to be either of a similar quality, to facilitate quality adjustment, and thus have relatively low sales, or of a different quality with relatively high sales, but requiring an extensive quality adjustment. In either case this is unsatisfactory.

7.155 Chaining, unlike hedonic indices, does not use all the price information in the comparison for each link. Items 2 and 6, for example, may be missing in March. The index makes use of the price information on items 2 and 6 when they exist, for the January–February comparison, but does not allow their absence to disrupt the index for the February–March comparison. It may be that item 4 is a replacement for item 2. Note how easily it is included as soon as two price quotes become available. There is no need to wait for rebasing or sample rotation. It may be that item 7 is a replacement for item 6. A quality adjustment to prices may be required for the February–March comparison between items 6 and 7, but this is a short-run one-off adjustment, the compilation of the index continuing in March–April using item 7 instead of item 6. *SNA 1993* (Chapter 16, para. 54) on price and volume measurement picks up on the point:

> In a time series context, the overlap between the products available in the two periods is almost bound to be greatest for consecutive time periods (except for sub-annual data subject to seasonal fluctuations). The amount of price and quantity information that can be utilized directly for the construction of the price or volume indices is, therefore, likely to be maximized by compiling chain indices linking adjacent time periods. Conversely, the further apart the two time periods are, the smaller the overlap between the ranges of products available in the two periods is likely to be, and the more necessary it becomes to resort to implicit methods of price comparisons based on assumptions. Thus, the difficulties created by the large spread between the direct Laspeyres and Paasche indices for time periods that are far apart are compounded by the practical difficulties created by the poor overlap between the sets of products available in the two periods.

7.156 The chained approach has been justified as the natural discrete approximation to a theoretical Divisia index (Forsyth and Fowler, 1981 and Chapter 16). Reinsdorf (1998) has formally determined the theoretical underpinnings of the index, concluding that in general chained indices will be good approximations to the theoretical ideal – though they are prone to bias when price changes “swerve and loop”, as Szulc (1983) has demonstrated (see also Forsyth and Fowler, 1981 and de Haan and Opperdoes, 1997).

7.157 The dummy variable hedonic index uses all the data in January and March for a price comparison between the two months. Yet the chained index ignores unmatched successive pairs, as outlined above; but this is preferable to its fixed base equivalent. The hedonic approach, in predicting from a regression equation, naturally has a confidence interval attached to such predictions. The width of the interval is dictated by the fit of the equation, the distance of the characteristics from their mean and the number of observations. Matching, chained or otherwise, does not suffer from any prediction error. Aizcorbe et al. (2001) undertook an extensive and meticulous study of high-technology goods (personal computers and semiconductors) using quarterly data for the period 1993 to 1999. The results from comparable hedonic and chained indices were remarkably similar over the seven years of the study. For, example, for desktop central processing units (CPUs) the index fell between the seven years from 1993:Q1 to 1999:Q4 by 60.0 per cent (dummy variable hedonic), 59.9 per cent (chained Fisher) and 57.8 per cent (chained geometric mean). The results differed only for quarters when there was a high turnover of items, and in these cases such differences could be substantial. For example, for desktop CPUs in 1996:Q4 the 38.2 per cent annual fall measured by the dummy variable hedonic method differed from the chained geometric mean index by 17 percentage points. Thus with little model turnover there is little discrepancy between hedonic and chained matched models methods and, for that matter, fixed base matched indices. It is only when binary comparisons or links have a high model turnover that differences arise (see also Silver and Heravi, 2001a and 2003).

7.158 Of course it is possible to make up for missing prices by using partial, patched hedonic estimates, as discussed above. Dulberger (1989) computed hedonic indices for computer processors and compared the results to those from a matched models approach. The hedonic dummy variable index fell by about 90 per cent over the period 1972 to 1984, about the same as for the matched models approach where missing prices for new or discontinued items were derived from a hedonic regression. However, when using a chained matched models approach with no estimates or imputations for missing prices, the index fell by 67 per cent. It is also possible to combine methods; de Haan (2003) used matched data when available and the time dummy only for unmatched data – his double imputation method.

Long-run and short-run comparisons

7.159 This section describes a useful formulation to aid quality adjustment. Its innovation arises from a possible concern with the long-run nature of the quality-adjusted price comparisons being undertaken. In the example in Table 7.1, prices in March were compared with those in January. Assumptions of similar price changes are required by the imputation method to hold over this period for long-run imputations – something that gives rise to increasing concern when price comparisons continue over longer periods, between January and October, January and November, January and December, and even subsequently. To help alleviate such concerns, this section considers a short-run formulation, mentioned in paragraph 7.42. Consider Table 7.5, which, for simplicity, has a single item A that exists throughout the period, an item B which is permanently missing in April, and a possible replacement item C in April.

Table 7.5 Example of long-run and short-run comparisons

Item	January	February	March	April	May	June
Comparable replacement						
A	2	2	2	2	2	2
B	3	3	4			
C				6	7	8
Total	5	5	6	8	9	10
Explicit adjustment						
A	2	2	2	2	2	2
B	3	3	4	**5/6** x **6=5**	**5/6** x **7=5.8**	**5/6**x**8=6.67**
C	**6/5**x3 **=3.60**			6	7	8

Total	5	5	6	8	9	10
Overlap						
A	2	2	2	2	2	2
B	3	3	4	**6** x **4/5=4.8**		
C			5	6	7	8
Total	5	5	6	6.8		
Imputation						
A	2	2	2.5	3.5	4	5
B	3	3	4	**3.5/2.5**x**4= 5.6**	**4/3.5**x **5.6=6.4**	**5/4** x **6.4=8**
Total	5	5	6.5	9.1	8.4	13

Figures in bold are estimated quality-adjusted prices described in the text.

Quality adjustment methods in short-run comparisons

7.160 A *comparable replacement* C may be found. In the previous example the focus was on the use of the Jevons index at the elementary level, since it is shown in Chapter 20 that this has much to commend it. The example here uses the Dutot index, the ratio of arithmetic means. This is not to advocate it, but only to provide an example using a different formulation. The Dutot index also has much to commend it on axiomatic grounds, but fails the commensurability (units of measurement) test and should only be used for relatively homogeneous items. The long-run Dutot index for April compared with January is:

$$\mathrm{P_D} \equiv \frac{\sum_{i=1}^{N} p_i^{Apr}/N}{\sum_{i=1}^{N} p_i^{Jan}/N}$$

which is 8/5= 1.60, a 60 per cent increase.

The short-run equivalent is the product of a long-run index up to the immediately preceding period, and an index for the preceding to the current period, i.e., for period *t*+4 compared with period *t*:

$$P_D \equiv \left[\frac{\sum_{i=1}^{N} p_i^{t+3}/N}{\sum_{i=1}^{N} p_i^{t}/N}\right] \times \left[\frac{\sum_{i=1}^{N} p_i^{t+4}/N}{\sum_{i=1}^{N} p_i^{t+3}/N}\right]$$

or for January with April:

$$\mathrm{P_D} \equiv \left[\frac{\sum_{i=1}^{N} p_i^{Mar}/N}{\sum_{i=1}^{N} p_i^{Jan}/N}\right] \times \left[\frac{\sum_{i=1}^{N} p_i^{Apr}/N}{\sum_{i=1}^{N} p_i^{Mar}/N}\right] \qquad (7.36)$$

which is, of course, $\frac{6}{5} \times \frac{8}{6} = 1.60$ as before.

7.161 Consider a *non-comparable replacement with an explicit quality adjustment*. Say, for example, that C's value of 6 in April is quality-adjusted to be considered to be worth only 5 when compared to the quality of B. The quality adjustment to prices may have arisen from an option cost estimate, a quantity adjustment, a subjective estimate or a hedonic coefficient, as outlined above. Say that the long-run comparison uses an adjusted January price for C, which is B's price of 3 multiplied by 6/5 to upgrade it to the quality of C, i.e., (6/5) x 3 = 3.6. From April onwards, the prices of the replacement item C can be readily compared to its January reference period price. Alternatively, the prices of C in April onwards might have been adjusted by multiplying them by 5/6 to downgrade them to the quality of B and enable comparisons to take place with item B's price in January: for April the adjusted price is 5/6 x 6=5; for May the adjusted price is 5.8 and for June it is 6.67 (see Table 7.5). Both procedures yield the same results for long-run price comparisons. The results from both methods (rounding errors aside) are the same for item B.

7.162 For the overall Dutot index, however, the results will differ, since the Dutot index weights price changes by their price in the initial period as a proportion of total price (see Chapter 20, footnote 27). The two quality adjustment methods will have the same price changes, but different implicit weights. The Dutot index in May is 9/5.6 = 1.607 if an adjustment is made to the initial (January) price or 7.8/5 = 1.56 if an adjustment is made to the current period (May) price. The short-run indices give the same results for each adjustment:

$\frac{8}{5.6}\times\frac{9}{8}=1.607$ using an adjustment to the initial (January) price, and

$\frac{7}{5}\times\frac{7.8}{7}=1.56$ using an adjustment to the current period (May) price.

7.163 The *overlap method* may also take the short-run form. In Table 7.5 there is a price for C in March of 5 that overlaps with B in March. The ratio of these prices is an estimate of their quality difference. A long-run comparison between January and April would be $\left(6\times\frac{4}{5}+2\right)/5$ = 1.36. The short-run comparison would be based on the product of the January to March and March to April link: $\frac{6.8}{6}\times\frac{6}{5}=1.36$.

7.164 At this unweighted level of aggregation it can be seen that there is no difference between the long-run and short-run results when items do not go missing, when comparable replacements are available, when explicit adjustments are made for quality or when the overlap method is used. The separation of short-run (most recent month-on-month) and long-run changes may have advantages for quality assurance to help spot unusual short-run price changes. But this is not the concern of this chapter. The short-run approach does, however, have advantages when imputations are made.

Implicit short-run comparisons using imputations

7.165 The use of the short-run framework has been mainly considered for temporarily missing values, as outlined by Armknecht and Maitland-Smith (1999) and Feenstra and Diewert (2001). Similar issues nevertheless arise in the context of quality adjustment. Consider again Table 7.5, but this time there is no replacement item C and item A's prices have been changed to show an upward trend. Item B is again missing in April. A long-run imputation for item B in April is given by (3.5/2)× 3 = 5.25. The price change is thus $(5.25+3.5)/5=1.75$ or 75 per cent. This is, of course, the same result as that obtained by simply using item A (3.5/2 = 1.75), since the implicit assumption is that price movements of item B, had it continued to exist, would have followed those of A. The assumption of similar long-run price movements may, in some instances, be difficult to support over very long periods. An alternative approach would be to use a short-run framework in which the imputed price for April is based on the (say, overall) mean price change between the preceding and current period, i.e. (3.5/2.5)×4 = 5.6 in the above example. In this case, the price change between March and April is (5.6+3.5)/(2.5+4) = 1.40. This is combined with the price change between January and March 6.5/5 = 1.30, to give the January to April change of 1.30×1.40 = 1.82, an 82 per cent increase.

7.166 Consider why the short-run result of 82 per cent is larger than the long-run result of 75 per cent. The price change for A between March and April of 40 per cent, upon which the short-run imputation is based, is larger than the average *annual* change of A, which is just over 20 per cent. The extent of any bias from this approach was found, above, to depend on the ratio of missing values, and the difference between the average price change of the matched sample and the quality-adjusted price change of the item that went missing, had it continued to exist. The short-run comparison is to be favoured if the assumption of similar price changes is considered more likely to hold than the long-run assumption.

7.167 There are data on price changes of the item that are no longer available, item B in Table 7.5, up to the period preceding the period in which it is missing. In Table 7.5, item B has price data for January, February and March. The long-run imputation makes no use of such data, simply assuming that price changes over the period of January to April, for example, are the same for B as for A. Let the data for B's prices in Table 7.5 (penultimate row) now be 3, 4 and 6 in January, February and March, respectively, instead of 3, 3 and 4. The long-run estimate for B in April is 5.25, as before. The estimated price change between March and April for B is now a fall from 6 to 5.25. A short-run imputation based on the price movements of A between March and April would more correctly show an increase from 6 to (3.5/2.5)× 6 = 8.4.

7.168 There may, however, be a problem with the continued use of short-run imputations. Returning to the data for A and B in Table 7.5, consider what happens in May. Adopting the same short-run procedure, the imputed price change is given in Table 7.5 as (4/3.5)×5.6 = 6.4 and for June as (5/4)×6.4=8. In the former case, the January to May price change is**:**

$$\left[\frac{(6.4+4)}{(5.6+3.5)}\right]\times\left[\frac{(5.6+3.5)}{(3+2)}\right]=2.08$$

and in the latter, for June:

$$\left[\frac{(8+5)}{(6.4+4)}\right]\times\left[\frac{(6.4+4)}{(3+2)}\right]=2.60$$

compared with long-run comparisons for May and June, respectively, of:

$$\left[\frac{((4/2)\times 3+4)}{(3+2)}\right]=2.00$$

$$\left[\frac{((5/2)\times 3+5)}{(3+2)}\right]=2.50$$

7.169 A note of caution is required. The comparisons here use an imputed value for item B in April and also an imputed value for May. The price comparison for the second term in equation (7.36) above, for the current versus immediately preceding period, uses imputed values for item B. Similarly, for the January to June results, the May to June comparison uses imputed values for item B for both May and June. The pragmatic needs of quality adjustment may of course demand this. If comparable replacements, overlap links and resources for explicit quality adjustment are unavailable, an imputation must be considered. However, using imputed values as lagged values in short-run comparisons introduces a level of error into the index which will be compounded with their continued use. Long-run imputations are likely to be preferable to short-run changes based on lagged imputed values, unless there is something in the nature of the industry that cautions against such long-run imputations. There are circumstances in which the price collector may believe the missing item to be temporarily missing, and the imputation is conducted in the expectation that production will subsequently continue; a wait-and-see policy is adopted under some rule, say that the item is missing for a maximum of three months, after which the item is deemed to be permanently missing. Such pragmatic situations require imputations to extend values over consecutive periods and call for the use of lagged imputed values to be compared with current imputed values, despite the fact that this is cautioned against, especially over several months. There is an intuitive feeling that the period over which this is undertaken should not be extensive. First, the effective sample size decreases as the use of imputation increases. Second, the implicit assumptions of similar price movements inherent in imputations are less likely to hold over the longer run. Finally, there is some empirical evidence, albeit from a different context, against the use of imputed values as if they were lagged actual values (see Feenstra and Diewert’s study (2001) using data from the United States Bureau of Labor Statistics for their International Price Program).

7.170 The above short-run approach will be developed in the next section, where weighted indices are considered. The practice of estimating quality-adjusted prices is usually carried out at the elementary item level. At this lower level, the prices of items may subsequently be missing, and replacements with or without adjustments and imputations are used to allow the series to continue. New items are also being introduced, as are newer varieties and switching of sales between sections of the index. The turmoil of changing quality is not just about maintaining similar price comparisons, but also about the accurate reweighting of the mix of what is consumed. Under a Laspeyres framework, the bundle is held constant in the base period, so any change in the relative importance of items consumed is held to be of no concern until the next re-basing of the index. Yet procedures for updating the weights are required to capture something of the very real changes in the mix of what is consumed. This is considered in Chapter 9. The concern here is with an equivalent higher-level procedure to the short-run adjustments discussed above. It is one particularly suited to countries where resource constraints prohibit the regular updating of weights through regular household surveys.

Single-stage and two-stage indices

7.171 Consider aggregation at the elementary level. This is the level at which prices are collected from a representative selection of outlets across regions in a period and compared with the matched prices of the same items in a subsequent period to form an index for, say, lamb. Each price comparison is equally weighted unless the sample design gives proportionately more chance of selection to items with more sales. The elementary price index for lamb is then weighted, and combined with the weighted elementary indices for other products to form the consumer price index. A Jevons elementary aggregate index for period *t*+6 compared with period *t*, for example, is given as:

$$P_J \equiv \prod_{i \in N(t+6) \cap N(t)} \left(p_i^{t+6} / p_i^t \right) \tag{7.37}$$

Compare this with a two-stage procedure:

$$P_J \equiv \prod_{i \in N(t+5) \cap N(t)} \left(p_i^{t+5} / p_i^t \right) \prod_{i \in N(t+6) \cap N(t+5)} \left(p_i^{t+6} / p_i^{t+5} \right) \tag{7.38}$$

7.172 If an item is missing in period *t*+6, an imputation may be undertaken. If equation (7.37) is used, the requisite assumption is that the price change of the missing item, had it continued, is equal to that of the average of the remaining items over the period *t* to *t*+6. In equation (7.38), the missing item in period *t*+6 may be included in the first stage of the calculation, between periods *t* and *t*+5, but excluded in the second stage, between periods *t*+5 and *t*+6. The requisite assumption is that price changes between *t*+5 and *t*+6 are equal. Assumptions of short-run price changes are generally considered to be more valid than their long-run counterparts. The two-stage framework also has the advantage of including in the worksheet prices for the current period and the immediately preceding one which, as is shown in Chapter 9, promotes good data validity checks.

7.173 Feenstra and Diewert (2001) applied a number of mainly short-run imputation procedures to price comparisons for the United States Bureau of Labor Statistics International Price Program. Although such price indices are not the direct interest of this manual, the fact that about one-quarter of the individual items tracked did not have price quotations in any given month makes it an interesting area in which to explore the results from different imputation procedures. When using the two-stage procedure, Feenstra and Diewert (2001) advise against carrying forward imputed period prices as if they were actual values, for the subsequent price comparison. The resulting price relatives for the subsequent period based on prior imputations had a standard deviation about twice that of price relatives where no imputation was required, leading the authors to conclude that such a practice introduced a significant amount of error into the calculation. Feenstra and Diewert (2001) found that higher variances of price changes arose from long-run imputation compared with the short-run imputation method. They also found, from both theory and empirical work, that when actual prices become available in a future data set and were used to interpolate back on a linear basis the missing prices, then such estimates lead to much lower variances than the short-run imputation approach. Such linear interpolations, however, require the statistical agency to store past information until a price quote becomes available, interpolate back the missing price, and then publish a revised consumer price index.

Appendix 7.1 Data on personal computers, obtained from United Kindgom Compaq and Dell Web sites, July 2000, to illustrate hedonic regression

PRICE (£)	SPEED(MHz)	RAM ,MB.	HD,M B.	DELL	PRESARIO	PROSIGNIA	CELERON	PENTIUM III	CD-RW	DVD	DELL*SPEE D (MHz)
2123	1000	128	40	0	1	0	0	0	0	0	0
1642	700	128	40	0	1	0	0	0	0	0	0
2473	1000	384	40	0	1	0	0	0	0	0	0
2170	1000	128	60	0	1	0	0	0	0	0	0
2182	1000	128	40	0	1	0	0	0	0	1	0

2232	1000	128	40	0	1	0	0	0	1	0	0
2232	1000	128	40	0	1	0	0	0	0	0	0
1192	700	384	40	0	1	0	0	0	0	0	0
1689	700	384	60	0	1	0	0	0	0	0	0
1701	700	384	40	0	1	0	0	0	0	1	0
1751	700	384	40	0	1	0	0	0	1	0	0
1851	700	384	40	0	1	0	0	0	0	0	0
2319	933	128	15	0	0	0	0	1	0	0	0
2512	933	256	15	0	0	0	0	1	0	0	0
2451	933	128	30	0	0	0	0	1	0	0	0
2270	933	128	10	0	0	0	0	1	0	0	0
2463	933	256	10	0	0	0	0	1	0	0	0
2183	933	64	10	0	0	0	0	1	0	0	0
1039	533	64	8	0	0	1	1	0	0	0	0
1139	533	128	8	0	0	1	1	0	0	0	0
1109	533	64	17	0	0	1	1	0	0	0	0
1180	533	64	8	0	0	1	1	0	1	0	0
1350	533	128	17	0	0	1	1	0	1	0	0
1089	600	64	8	0	0	1	0	1	0	0	0
1189	600	128	8	0	0	1	0	1	0	0	0
1159	600	64	17	0	0	1	0	1	0	0	0
1230	600	64	8	0	0	1	0	1	1	0	0
1259	600	128	17	0	0	1	0	1	0	0	0
1400	600	128	17	0	0	1	0	1	1	0	0
2389	933	256	40	0	1	0	0	1	0	0	0
1833	733	256	40	0	1	0	0	1	0	0	0
2189	933	128	40	0	1	0	0	1	0	0	0
2436	933	256	60	0	1	0	0	1	0	0	0
2397	933	256	40	0	1	0	0	1	0	1	0
2447	933	256	40	0	1	0	0	1	1	0	0
2547	933	256	40	0	1	0	0	1	0	0	0
2845	933	384	60	0	1	0	0	1	0	0	0
2636	933	384	60	0	1	0	0	1	0	0	0
1507	733	64	30	0	1	0	0	1	0	0	0
1279	667	64	10	1	0	0	0	1	0	0	667
1379	667	128	10	1	0	0	0	1	0	0	667
1399	667	64	30	1	0	0	0	1	0	0	667
1499	667	128	30	1	0	0	0	1	0	0	667
1598	667	128	30	1	0	0	0	1	1	0	667
1609	667	128	30	1	0	0	0	1	0	1	667
1389	667	64	10	1	0	0	0	1	0	1	667
999	667	64	10	1	0	0	1	0	0	0	667
1119	566	64	30	1	0	0	1	0	0	0	566
1099	566	128	10	1	0	0	1	0	0	0	566
1097	566	64	10	1	0	0	1	0	1	0	566
1108	566	64	10	1	0	0	1	0	0	1	566
1219	566	128	30	1	0	0	1	0	0	0	566
1318	566	128	30	1	0	0	1	0	1	0	566
1328	566	128	30	1	0	0	1	0	0	1	566
1409	566	128	10	1	0	0	0	1	0	0	733
1809	733	384	10	1	0	0	0	1	0	0	733
1529	733	128	30	1	0	0	0	1	0	0	733
1519	733	128	10	1	0	0	0	1	0	1	733
1929	733	384	30	1	0	0	0	1	0	0	733
2039	733	384	30	1	0	0	0	1	0	1	933
2679	933	128	30	1	0	0	0	1	0	0	933
3079	933	384	10	1	0	0	0	1	0	0	933
2789	933	128	10	1	0	0	0	1	0	1	933
3189	933	384	10	1	0	0	0	1	0	1	933