TEMPORARILY AND PERMANENTLY MISSING PRICES AND QUALITY CHANGE

6

Introduction

6.1 Chapter 6 focuses on the treatment of temporarily and permanently missing varieties and their prices. While Chapter 5 focuses on the collection of data, Chapter 6 highlights the important role of the price collector in the context of the treatment of missing prices and starts by providing an overview of the matched-model method (MMM). While the MMM serves as the underlying method regarding the treatment of missing prices, the chapter describes how the MMM can potentially fail, the consequences of this failure, and how to deal with the effects of such failure on price measurement.

6.2 Temporarily missing prices and the methods used for the treatment of missing varieties are reviewed in this chapter. The concept of quality is defined and discussed. Explicit (direct) and implicit (indirect) methods for quality adjustment are identified and described.

6.3 Some introductory notes are provided on general measurement issues including the use of additive versus multiplicative quality adjustments, price reference versus current period quality adjustment, short-term versus long-term comparisons, and geometric aggregation formula. Finally, this chapter considers the need of price measurement in product markets with a rapid turnover of models,[1] usually in the electronic and high-technology product markets.

Background

6.4 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, a representative sample could be taken of the varieties of goods and services households consumed in a reference period 0, their prices recorded and compared with the prices of the same matched varieties in subsequent periods. In this way, the prices of like would be compared with like. In practice, some complications do exist. Varieties change in quality over time and replacements are of a different quality compared to the original. New and old models of varieties appear and disappear.

6.5 Changes in the quality of varieties should be treated as changes in the volume of the goods or service provided, as opposed to changes in the price. For example, increases over time in the concentration of a detergent (number of washes per one kilogram packet), faster internet service

(megabits per second, Mbps), and inclusion of a warranty in the price of a dishwasher, all contribute to effective decreases in price; consumers get more for their money. Similarly, quality decreases, for example, less legroom in economy flights, when prices remain constant, are effective increases in price. A volume change for an individual variety may be comprised of a quantity *and* quality change. The change in the variety's nominal value of consumption expenditure is the product of its price and volume change. It follows that the price change is the change in value divided by the change in volume.

6.6 National statistical offices (NSOs) go to great lengths to ensure measured price changes are not influenced by changes in the quality of items. By measuring the price change of a fixed, constant-quality basket of goods and services, NSOs use the MMM. When updating the basket, price collectors visit selected outlets with broad details of an item and identify the most popular, regularly stocked varieties sold in each of the outlets. Next, they develop a detailed description of the variety including all the price-determining characteristics (for example, brand or size) and record the price. This specification must be sufficiently detailed to include all price-determining characteristics to define a unique, specific variety. The detailed specification allows the price collector to easily identify the variety or model in subsequent periods and record its matched price.

6.7 The measurement of changes in the level of consumer prices using the MMM is appropriate when variety prices are not missing. However, the use of the MMM is complicated by *temporarily unavailable* prices, for example for one, two, or three months, because of a variety being out of stock and not yet replenished. A matched price is unavailable in these intervening months. The treatment of prices of temporarily missing varieties is considered in more detail in paragraphs 6.52–6.72, but typically requires the missing variety's price to be imputed for the month(s) in which it is missing using the price changes of similar goods or services, or price changes drawn from a higher level of aggregation. Actual prices are then compared with imputed prices for the measurement of the consumer price index (CPI).

6.8 If varieties become *permanently unavailable* a replacement variety is required that is preferably comparable with regard to the price-determining characteristics of the missing variety. If the replacement is of a comparable quality (that is, possesses the same price-determining characteristics), its price can be directly compared with the last actual or imputed price for the missing variety. If the replacement is noncomparable, for example, it is of a better quality, the improvement in quality has to be explicitly quantified with regard to its "worth" or its contribution to the price. Using this value, compilers make a price adjustment to reflect the difference in quality allowing the price

---

[1] The term "model" of a good or service is used in this chapter mainly in the context of high-technology goods such as laptops, household appliances, or cars. This use of the term "model" follows general usage and refers to a specific variety whose characteristics are updated regularly.

of like be compared with like. If a reliable explicit quality adjustment to the price is not possible, for data or resource reasons, implicit methods of quality adjustment are available. Details of explicit and implicit methods of quality adjustment are provided in paragraphs 6.90–6.188, and methods for dealing with new and disappearing goods and services are illustrated in Chapter 8.

6.9 Products that are "strongly" seasonal, that is, missing in particular months when out of season but expected to return in the next season, could also be treated as temporarily missing and imputed. Chapter 11 describes in more detail the different options that can be used for the treatment of seasonal products. Strongly seasonal products include some fresh fruits, vegetables, and clothing. Also considered in Chapter 11 is the treatment of "weakly" seasonal products: available throughout the year but whose prices, sales, and quality fluctuate. The prices of weakly seasonal products are not missing and are treated differently from strongly seasonal ones, as discussed in Chapter 11.

6.10 The matching of models facilitates the measurement of constant-quality price change. When the matching breaks down because of missing prices, temporary imputations are necessary until the variety's temporarily missing price becomes available or a replacement can be introduced, thus helping to update the sample. However, there are product markets where the matching breaks down on a regular basis because of high turnover of models, with new models of different quality compared to the old ones, such as laptop computers. In this case, a failure to match and replace models would lead to a seriously depleted and unrepresentative sample. Yet, a continual process of linking-in new replacement varieties has been found to lead to a bias in CPI measurement. Paragraphs 6.141–6.177 of this chapter outline an alternative approach making use of hedonic regressions.

Potential Errors in the Matched-Model Method

6.11 Three potential sources of error arise from the MMM: missing varieties, representativity of sample space, and new products.

Missing Varieties

6.12 The first source of error in the MMM, and the focus of this chapter, occurs when a variety is no longer available in the outlet. It may be temporarily out of stock, discontinued, or one or more of the price-determining characteristics may have changed. Whatever the reason, the variety is effectively missing in the current period and a price cannot be collected. The variety's price may be missing for other reasons: it may be a seasonal variety or one whose price does not need to be recorded so frequently, or it may be a custom-made good or service, supplied each time to the customer's specification.

6.13 It is necessary to distinguish between varieties that are permanently and temporarily missing. Varieties that are *temporarily* missing are varieties that are not available and not priced in the current period, but that are available and priced in subsequent periods. The treatment of varieties missing because demand and supply are seasonal, as is the case with some fruits and vegetables, is described in Chapter 11.

6.14 The different methods for the treatment of missing prices, and the implied assumptions, are listed in Figure 6.3 and discussed in some detail in paragraphs 6.90–6.234. By definition, the prices of the unavailable varieties cannot be determined and the accuracy of some of the assumptions about their price changes is difficult to establish. The matching of prices of varieties allows for the measurement of price changes unaffected by quality changes. When varieties 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 Issues

6.15 There are three sampling concerns when using the MMM. First, the MMM and the use of replacements are designed to meet the needs of constant-quality price measurement and while the sample of varieties priced might initially be designed to be representative of price changes of the population of varieties, it is effectively following a static sample of varieties that, over time, can become increasingly unrepresentative. The matching of prices of identical varieties over time, by its nature, is likely to lead to the monitoring of a sample of varieties increasingly unrepresentative of the population. The sample deteriorates over time because the MMM fails to incorporate new models/varieties into the sample, except as replacements to obsolete ones. For example, substantial developments in telecommunication hardware and services, reflected in the growing number of models available, are excluded from the sample covered by the CPI. This omission would not be problematic if the (implicit) quality-adjusted price changes of the excluded varieties were similar to those of the included matched-model sample. However, this is unlikely to be the case. The (quality-adjusted) prices of old models being dropped may be relatively low and the (quality-adjusted) prices of new ones relatively high as part of a sales strategy of dumping old models at relatively low price to make way for the introduction of new models priced relatively high.

6.16 A second sampling concern with the use of the MMM relates to the timing of the substitution and to when a replacement variety is chosen to replace an old one. In general, the prices of varieties continue to be monitored until they are no longer sold. This means that old varieties with limited sales continue to be monitored and included in the sample. Such varieties may show unusual price changes as they approach 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 their life cycle. The (implicit or explicit) weight of end-of-cycle varieties in the index would thus remain relatively high, being based on their sales share when they were sampled. Furthermore, new unmatched varieties with possibly relatively large sales would be ignored. Consequently, greater weight would be given to the unusual price changes of matched varieties at the end of their life cycle.

122

6.17 The final sampling concern with the use of the MMM results from the price collector collecting prices until the variety is no longer available, thus forcing a replacement. Data collectors replace the missing discontinued variety with the most popular or typically consumed variety. This improves the coverage and representativity of the sample. But it also makes reliable quality adjustments of prices between the old and new popular varieties 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 variety is in the final stages of its life cycle and the other in its first. Furthermore, the technical differences between the varieties 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 varieties are unlikely to meet assumptions of "similar price changes to existing varieties or classes of varieties," as required by the imputation methods. Many of the methods of dealing with quality adjustment for unavailable varieties may be better served if the switch to a replacement variety is made earlier rather than later. Sampling issues are closely linked to quality-adjustment methods. This topic is discussed in Chapter 8.

6.18 This chapter references the need to permanently replace missing varieties to ensure that the sample of varieties does not become unrepresentative. Samples of representative varieties and outlets are generally updated when an index is updated. Where there is a lengthy period between rebasing, the sample can become seriously deteriorated. It is feasible to update/rotate the sample between periods of revising the index and Chapter 7 outlines how this can be done in the context of maintaining the representativity of the sample. Chapter 6 refers to the need for regular updating of the sample which could be achieved through sample rotation.

New Products

6.19 Another potential source of error when using the MMM arises when a new product is introduced into the marketplace. When a really new product is introduced, there is an immediate gain in welfare or utility as demand switches from the old variety to the new variety. For example, the introduction of mobile telephones represented a completely new good that led to an initial gain in utility or welfare to consumers as they switched from the old (landlines) to the new technology. This gain from the introduction of mobile telephones, and subsequently of increasingly smarter telephones, 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 mobile telephones 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 neither well developed nor practical for CPI compilation, as referenced in more detail in Chapter 7. For a CPI built on the concept of a base period and a fixed basket, this situation, strictly speaking, does not represent a

problem. The new product 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 households purchase. 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, there is no doubt that it would be conceptually appropriate to include the new product and any welfare gain from its introduction, though as outlined in Chapter 8, this is highly problematic in practice.

Useful Concepts for the Treatment of Missing Prices

Multiplicative versus Additive Adjustment

6.20 Explicit quality adjustments to prices can be made by adding a fixed amount (additive adjustment) or multiplication by a ratio (multiplicative adjustment). For example, consider m, an old variety, and n its replacement; for a price comparison over periods t, $t + 1$, and $t + 2$, the price of m is only available in periods t and $t + 1$ and n only available in periods $t + 1$ and $t + 2$. The measurement of constant-quality price change between periods t and $t + 1$ is based on variety m, p_m^{t+1} / p_m^t, and the price change between periods $t + 1$ and $t + 2$ on variety n, p_n^{t+2} / p_n^{t+1}. Although not necessary for the compilation of the price index, the previous calculation can be elaborated in an equivalent, though more complex, form that enables the nature of the quality adjustment from m to n to be identified and the multiplicative formulation demonstrated.

6.21 A price relative over periods t, $t + 1$, and $t + 2$ requires an overlap ratio p_n^{t+1} / p_m^{t+1} to be used as a measure of the relative quality difference between the old variety and its replacement. This ratio could then be *multiplied* by the price of the old variety in period t, p_m^t to obtain the quality-adjusted prices p_n^{*t} as outlined later in equation (6.6), and illustrated in Table 6.1:

6.22 Such multiplicative formulations are generally recommended, as the adjustment is invariant to the absolute value of the price. If the overlap ratio equals, for example, 1.2, the new variety costs 20 percent more than the old. There may be some varieties 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. For example, retailers selling on websites may include free shipping in the price. In some instances, the cost of shipping may remain the same in the short to medium-term irrespective of what happens to the price of the variety (exclusive of shipping). If the price no longer includes free shipping, this fall in quality should be valued as a fixed additive sum.

Table 6.1   Example of a Replacement Variety with Overlap

	t	$t + 1$	$t + 2$
Old Variety m	p_m^t	p_m^{t+1}	
Replacement Variety n	p_n^{*t}	p_n^{t+1}	p_n^{t+2}

123

Price Reference versus Current Period Adjustment

6.23   Two variants of the approaches to quality adjustment are to make the adjustment either to the price in the price reference period or to the price in the current period. For example, in the overlap method, described previously, the implicit quality-adjustment coefficient was used to adjust p_m^t to p_n^{*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 variety p_m^{t+2} to obtain the quality-adjusted price p_n^{*t+2}. The first approach is more straightforward since, once the reference period price has been adjusted, no subsequent adjustments are required. Each new price of a replacement can be compared with the adjusted reference price.

Long-Term versus Short-Term Comparisons

6.24   Much of the analysis of quality adjustments in this Manual is undertaken by comparing prices between two adjacent periods, say, month-on-month period t prices with those in a subsequent period $t + 1$. For long-term comparisons, the price reference period is taken as, for example, period t and the index compiled by comparing prices in period t first with $t + 1$; then t with $t + 2$; t with $t + 3$, and so on. The short-term framework allows long-term comparisons built up as the product of links: t first with $t + 1$; then $t + 1$ with $t + 2$; $t + 2$ with $t +3$, and so on; built up as a sequence of links joined together by successive multiplication. This chapter focuses on short-term comparisons, for reasons of their inherently better properties and for focus of exposition; In particular, as outlined in paragraphs 6.51–6.71, the short-term approach enables superior imputations to be made for temporarily missing prices and facilitates the incorporation of replacement varieties as and when an old variety's price is permanently missing. The short-term approach is generally recommended.

Treatment of Missing Variety Prices and Quality Adjustment within an Elementary Aggregate: Short-Term Comparisons

6.25   This chapter uses a short-term framework of comparing period-on-period prices rather than a long-term framework of comparing the current period's price with a fixed price reference period. The use of matching is particularly problematic for long-term price comparisons. For long-term price comparisons a selection of representative models in period 0, for example, the year (or a month in) 2020, have their prices compared with those in January 2021; in February, for the 2020-February price relative; in March, for the 2020-March price relative; continuing for what may be, in some countries, several years. The sample is increasingly depleted over time as 2020 varieties become obsolete.

6.26   The short-term approach has several advantages. An illustration of the short-term against the long-term approach is included in Chapter 7. This chapter is concerned with temporarily and permanently missing variety prices. Temporarily missing variety prices return to the sample, as do seasonal ones (as described in Chapter 11), and there is no issue with maintaining the sample. However, permanently missing variety prices need a replacement variety so that, over time, the sample does not become increasingly

depleted and degraded. Yet such one-on-one replacement is unlikely to be sufficient to maintain the representativity of the sample, which may be based on variety selection when updating the index (Chapter 9) or sample rotation that may for some countries be many years ago. Since this initiation, many newer varieties/products may have been introduced and old ones become obsolete. Maintaining the representativity of the sample is addressed in Chapter 7 and a related use of web-based and scanner data sources in Chapter 10. The treatment of permanently missing variety prices was set within the context of an elementary aggregate, where weights were neither available nor used. A modified Lowe price index number is used. The formulas are presented in more detail in Chapters 7 and 8 (see paragraphs 8.109–8.116).

6.27   The example in Table 6.3 might be used to illustrate an individual elementary aggregate with geometric mean prices compiled from several outlets. If a variety is temporarily not available, an imputation can be based on a short-term month-on-month price relative, rather than a long-term price that might assume similar price movements over several years. Similarly, for permanently missing prices where imputations are used to form an overlap comparison for the missing variety price and its replacement, assumptions based on similar short-term month-on-month price movements are more reasonable than the less plausible ones based on long-term price movements.

6.28   This mechanism facilitates the inclusion of new specifications when old specifications become obsolete and enables the index to better represent the dynamic changes taking place in consumer choice. A direct comparison between the price of a new replacement variety specification in April 2020 with its old specification in 2019 is likely to be challenging given the quality differences between the two variety specifications over a long period.

6.29   The short-term formula will differ (be improved) from its fixed-base long-term counterpart since the monthly imputations will differ.

6.30   This chapter's work has for the large part been concerned with short-term price relatives compiled within an elementary aggregate, i. The larger picture of weighted aggregation across elementary aggregates is for this context of missing prices and sample representativity, considered in Chapter 7, with illustrative calculations of the aggregation formulas in Chapter 8, and the introduction of new weights in Chapter 9.

Aggregation Formula for Elementary Price Indices

6.31   A ratio of geometric means (that is, the Jevons price index number formula) is used to measure price changes at the unweighted level of the elementary aggregate. Alternative formulas include a ratio of arithmetic means (that is, the Dutot price index number formula) and an arithmetic average of price ratios—the Carli price index number formula. The Jevons price index formula is used here for reasons of its better properties and focus of exposition. Chapter 9 provides detail, an illustration, and the relative merits of the use of the Dutot and Carli price index number formula.

6.32   With scanner and other such data, information on prices, expenditure values, and quality characteristics will often be available for most individual models sold by major outlets.

124

This availability of data on transaction values allows weights to be used at this detailed level of aggregation and thus the use of weighted price index formulas as outlined in Chapter 9.

The Role of Price Collectors

6.33 Price collectors have a critical role to play in the treatment of missing price observations. They observe and record that a price is missing; whether it is temporarily or permanently missing; if permanently missing, whether a comparable or noncomparable replacement is available, and in the latter case, the price and details of the replacement variety. When selecting a sample of prices, the outlets are visited in a process referred to as initiation. During the initiation phase of price collection, collectors identify the detailed specifications of representative varieties sold. For example, for the general class of "large white bread, unsliced," the more detailed, "large loaf, white, unsliced, Brand A, 800 gm" may be selected and its details entered along with its price for subsequent periodic repricing. Ideally, price collectors should have in their possession a checklist of these specifications when visiting outlets in subsequent periods. The detailed specifications serve many purposes including (1) to help identify the variety to be priced; (2) to review and verify the variety's specification to ensure there have been no changes in the price-determining characteristics; and (3) if the variety is noncomparable, to use the specifications to identify a replacement variety to be priced and record any changes in the price-determining specifications. Initiation activities occur only when it is necessary to select a new variety for pricing.

6.34 The price collector also plays an important role in determining whether the missing price should be treated as temporarily or permanently missing. A variety's price is considered temporarily missing if the same variety is likely to return to the market within a reasonable time period. On finding that the specified variety is not available for immediate sale, the price collector should check with the manager or informed member of outlet staff whether it is temporarily or permanently missing. If temporarily missing, record the expected duration: one, two, or more periods should be recorded along with the reason for it being unavailable and an indication of the likelihood of its return.

6.35 Temporarily missing varieties have their prices imputed; permanently missing ones require a replacement. As these are different issues requiring different treatments, it is important for the price collector to establish whether the unavailability of the variety is temporary or permanent. Consider the case of a monthly CPI. When a price is temporarily missing, it should be imputed using an overall mean imputation, a targeted mean imputation, or a class mean imputation. Some NSOs use a method referred to as carryforward (repeating the last observed price). As stressed in paragraph 6.65, this is not recommended.

6.36 Permanent unavailability occurs when the variety is withdrawn from the market with no prospect of returning. In some instances, it might be absent the next month and confirmed by the outlet manager or informed staff that it is not going to be replaced. With such information the price collector should immediately look to collecting the price and specifications of a replacement variety. In other cases, if a variety is out of stock, for example during three consecutive months, the price collector should be instructed to choose

Table 6.2 Illustrative Variety Codes for Price Collector for Missing Values

Collection Code	Description
T	**Temporarily Missing:** The Variety Is Unavailable but Is Expected to Be Available Again in the Near Future
P	**Permanently Missing:** The Variety Is No Longer Available and Is Unlikely to Return
S	**Seasonally Missing:** The Variety (Product Group) Is Strongly Seasonal and Is out of Season
C	**Comparable Replacement:** A Replacement Variety That Is Comparable to the Old Variety in All Major Aspects
NC	**Noncomparable Replacement:** A Replacement Variety That Is Not Comparable to the Old Variety

a replacement that matches as closely as possible the missing variety's specification. After being unavailable for three months and with no indication that the variety will return, it is treated as being permanently missing and a replacement variety is sought. There may be particular products or circumstances in which the three-month rule can be relaxed, such as a temporary withdrawal of products for health reasons, national emergencies, or the logistics of restocking where there is a sound basis for a belief that the product will return in the near future, but not within three months.

6.37 Decisions on the treatment of missing prices are made by CPI staff based on information provided by the price collector and, in some instances, by a follow-up telephone contact or visit to the outlet.

6.38 Variety prices may be missing for products because they are seasonal and out of season. Products that are out of season, but expected to return in the next season, are treated differently from those considered in this chapter, and their treatment is the subject of Chapter 11.

6.39 Data collection codes, such as those illustrated in Table 6.2, should be used to justify or explain each missing price to ensure proper treatment. Metadata should be collected on those products in which, for example, there is a high level of missing prices of different forms and the extent to which the prices are missing. Illustrative variety codes are given in Table 6.2 and NSOs should build on the detail required to meet their specific needs.

The Treatment of Temporarily and Permanently Missing Variety Prices

6.40 To measure aggregate price changes, a representative sample of varieties is selected from a sample of outlets, along with detailed item descriptions that define each variety and its specifications. For each selected variety, detailed item specifications, or descriptions, define a unique, specific variety that will be priced each period. The detailed specifications are included on the repricing form each period and serve as a prompt to help ensure that the same varieties are being priced. Detailed checklists of variety descriptions should be used, as any lack of clarity in the specifications may lead to errors. Attention should also be devoted to ensuring that the specifications used are not just to identify the variety on a subsequent visit, for example its location in the outlet, but contain all pertinent, price-determining characteristics, otherwise it will not be possible to identify if changes in quality have occurred.

125

6.41 The MMM succeeds in ensuring the prices of like being compared with like. The detailed variety specifications facilitate this process, and the method ensures that the measurement of price change is not influenced by quality changes. However, when a price is missing there is potential for mismeasurement. The treatment of missing prices depends on whether the variety is temporarily missing (that is, the product will be available in the near future) or permanently missing (that is, the variety will not be available in the future). Each month, all imputed prices are used in the calculation of the index. For example, an imputed March price would be compared with the actual price collected in February in order to calculate the price change from February to March. If the variety returns in April, the imputed March price is compared with its actual price collected in April to calculate the March to April price change. All permanently missing prices require a replacement.

6.42 When a variety is missing in a month a number of approaches may be used. Details of each method are provided in separate sections of this chapter. While the terminology may differ between NSOs, the methods include:

- *Imputation*. The price change for the missing variety is assumed to be the same as the price changes of all varieties in the item group, or targeted similar ones. Such imputations can be used for temporarily missing varieties. Permanently missing varieties, however, require a comparable or noncomparable replacement.

- *Direct comparison*. If a variety is permanently missing and a replacement variety 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 price of the unavailable variety. Any difference in price level between the new and the old is judged to be price change and not quality change.

- *Explicit quality adjustment*. If a replacement variety is noncomparable (that is, there are identifiable quality differences) estimates of the impact of the quality differences on the price enable quality-adjusted price comparisons to be made between the old and the new varieties.

- *Implicit quality adjustment—overlap*. If a replacement variety is noncomparable and no information is available, or resources are too limited to allow reasonable explicit estimates to be made of the impact of a quality change on the price, the price difference between the old variety and its replacement in the overlap period is then taken to be a measure of the quality differential.

6.43 Specific attention needs to be devoted to products with relatively big weights, where large proportions of varieties are turned over. Some of the methods are not straightforward and require a higher level of expertise and experience. Quality adjustment should be implemented by developing a methodical approach on a product-by-product basis. Such concerns should not be used as excuses for failing to attempt to estimate quality-adjusted prices. Ignoring quality change results in an implicit quality adjustment. This approach assumes that any difference in price is pure price change and not changes because of a difference in quality. Such an implicit approach may not be appropriate and may even be misleading.

6.44 The treatment of missing price quotes is divided into two types, depending on if these are for temporarily or permanently missing variety prices. These two types cannot always be readily identified and treated accordingly. Some form of mechanism or rule is required to enable a transition from temporary to permanently missing variety prices. A price collector, supported by the head office, would regard a variety as permanently unavailable if verified by an informed member of the outlet's staff or, following a three-month period, the variety is no longer available and there is no evidence that it will reappear. A price index suffers from sample depletion if an increasing number of temporary missing prices are imputed over a lengthy period. Once judged to be a "permanently missing" variety type a replacement is found. If noncomparable, a quality adjustment is made, and imputations are no longer necessary.

6.45 It may also be the case that a replacement, comparable or not, is unavailable (for example, for a videocassette player when it became obsolete). The topic of introducing new goods and services and removing obsolete ones is the subject of Chapter 7. Alternatively, a price collector may find the variety to be permanently missing in the outlet visited, and with no comparable or noncomparable replacement, although the variety is sold in other outlets. An informed outlet staff may inform the price collector that the product is no longer being stocked. For example, considering bicycles in a sports shop, a price index for bicycles could be continued by imputing the price change for this outlet using the price change in other outlets as described in paragraphs 6.51–6.71. Such a procedure depletes the sample and in the longer term this should be remedied by a forced outlet replacement, sample rotation, or rebasing (see Chapter 7).

6.46 Adjustments to prices are not a simple matter of applying routine methods to prices in specified product areas, and alternative approaches are suggested in this chapter. Some approaches are 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 be required for the successful implementation of quality adjustments.

The Treatment of Temporarily Missing Price Observations

Overall Mean Imputation

6.47 The overall mean imputation method uses the price changes of other similar varieties as estimates of the price change of the missing variety. Considering a Jevons elementary price index (that is, a geometric mean of price relatives, equivalent to the ratio for geometric means of prices),[2] the price of the missing variety in the current period, $t + 1$, is imputed by multiplying its price in the immediately preceding period t by the geometric mean of the price relatives of the remaining matched varieties in the product group between these two periods. This method provides the same

---

[2]Chapter 8 provides further examples of imputations and replacements using the Jevons index, as compared with the Dutot (ratio of arithmetic means) and Carli indices (mean of the ratio of price relatives), and outlines the relative merits of the three indices.

result as simply dropping the variety that is missing from both periods from the calculation. In practice, the series is continued in the database by including the imputed prices; and, as described in Table 6.3, this forms a complete table of the variety prices in outlets A to F. The imputations are based on assumptions of similar price movements.

6.48   In the example in Table 6.3, a product with broad specifications is sold in six outlets, A to F, with different detailed outlet-specific specifications adopted for each outlet. The price reference period is December 2019 with successive prices collected for each outlet's specification in January, February, March, April, May, June, and July 2020. The price collector finds the variety temporarily missing in outlet F's price collection for March 2020, and likely to remain missing for the next month or two, but to return thereafter. In Table 6.3, the figures in bold for prices of March to May 2020 in outlet F represent two alternative imputation methods, as explained in the following paragraphs.

6.49   The Jevons index number formula is shown in equation 6.2 as a direct or long-term index comparing, in its second to last term, the geometric mean of the prices of each matched variety in the current month t with the geometric mean of the prices in the price reference period, hereafter referred as period 0, and in the last term, for the example in Table 6.3, July 2020 with the price reference period (= 100) of December 2019.

$$I_j\left(I^0, I^t\right) = \prod_{i=1}^{nN}\left(\frac{p_i^t}{p_i^0}\right)^{\frac{1}{n}} \equiv \frac{\prod_{i=1}^{N}\left(p_i^{Jul2020}\right)^{\frac{1}{n}}}{\prod_{i=1}^{N}\left(p_i^{Dec2019}\right)^{\frac{1}{n}}} \quad (6.2)$$

where p = price.

6.50   In practice, the use of the Jevons formula in this long-term form is not advised. Instead, a short-term formulation is recommended as the product of month-on-month Jevons indices. The short-term cumulative Jevons index for December 2019 = 100 to July 2020 is

$$I_J\left(I^{Dec2019}, I^{Jul2020}\right) = \frac{\prod_{i=1}^{N}\left(p_i^{Jan2020}\right)^{\frac{1}{n}}}{\prod_{i=1}^{N}\left(p_i^{Dec2019}\right)^{\frac{1}{n}}} \times \frac{\prod_{i=1}^{N}\left(p_i^{Feb2020}\right)^{\frac{1}{n}}}{\prod_{i=1}^{N}\left(p_i^{Jan2019}\right)^{\frac{1}{n}}}$$

$$\times \frac{\prod_{i=1}^{N}\left(p_i^{Mar2020}\right)^{\frac{1}{n}}}{\prod_{i=1}^{N}\left(p_i^{Feb2019}\right)^{\frac{1}{n}}} \times .... \times \frac{\prod_{i=1}^{N}\left(p_i^{Jul2020}\right)^{\frac{1}{n}}}{\prod_{i=1}^{N}\left(p_i^{Jun2019}\right)^{\frac{1}{n}}}$$

$$(6.3)$$

6.51   The long-term and short-term approaches in equations 6.2 and 6.3 provide the same result as the numerator of each term on the right-hand side of equation 6.3 cancels with the denominator of the next. However, a major advantage of the short-term formulation is that when an individual variety is missing, its price can be imputed using the month-on-month price changes of similar or higher-level aggregates rather than the long-term comparison of the current month to the price reference period, which may be several years ago.

6.52   The right-hand side of equation 6.3 requires monthly geometric mean prices to be calculated for the numerators and denominators for each month-on-month comparison. Table 6.3 shows the geometric means for the price reference period (December 2019, January 2020, and February 2020); missing prices for March, April, and May; and prices returning in June and July. The first task is to impute the missing price for March 2020. This is obtained using the ratio of geometric mean prices for each of the

Table 6.3   Temporarily Missing Price Observations and Imputed Prices

Outlets		Price Reference Period							
		Dec.-19	Jan.-20	Feb.-20	Mar.-20	Apr.-20	May-20	Jun.-20	Jul.-20
A	Supermarket	5.25	5.25	5.49	5.49	5.49	5.49	5.49	5.49
B	Supermarket	5.10	5.10	5.10	5.25	5.25	5.25	5.25	5.25
C	Supermarket	5.20	5.20	5.20	5.20	5.20	5.25	5.25	5.25
D	Independent Trader	5.49	5.49	5.49	5.65	5.75	5.80	5.80	6.00
E	Independent Trader	5.99	6.50	6.50	6.90	6.90	6.90	6.90	7.00
F	Independent Trader	5.99	5.99	5.99	**6.13**	**6.15**	**6.17**	6.25	6.25
Overall Mean Imputation: F Mar:May									
Geometric Mean: A:E				5.54	5.67	5.69	5.71		
Geometric Mean: A:F		5.49	5.57	5.61	5.74	5.76	5.78	5.79	5.84
Short-Term Price Relatives: A:F		1.00000	1.01371	1.00748	1.02377	1.00352	1.00365	1.00198	1.00864
Long-Term Indices as Product of Short Term		100.00	101.37	102.13	104.56	104.92	105.31	105.52	106.38
Targeted Imputation: Independent Traders									
D	Independent Trader	5.49	5.49	5.49	5.65	5.75	5.80	5.80	6.00
E	Independent Trader	5.99	6.50	6.50	6.90	6.90	6.90	6.90	7.00
F	Independent Trader	5.99	5.99	5.99	**6.26**	**6.32**	**6.34**	6.25	6.25
Geometric Mean: A:C & D:F		5.49	5.57	5.61	5.76	5.79	5.81	5.79	5.84
Short-Term Price Relatives: A:F		1.00000	1.01371	1.00718	1.02731	1.00440	1.00377	0.99753	1.00808
Long-Term Indices as Product of Short Term		100.00	101.37	102.13	104.92	105.38	105.78	105.52	106.37

Bold: Imputed values

matched sample of outlets A to E in February and March to provide the short-term price relative:[3]

$$P_J\left(P^{Feb\,2020}, P^{Mar\,2020}\right) = \frac{\prod_{i=A}^{E}\left(p_i^{Mar\,2020}\right)^{\frac{1}{5}}}{\prod_{i=A}^{E}\left(p_i^{Feb\,2020}\right)^{\frac{1}{5}}}$$

$$= \frac{(5.49 \times 5.25 \times 5.20 \times 5.65 \times 6.90)^{\frac{1}{5}}}{(5.49 \times 5.10 \times 5.20 \times 5.49 \times 6.50)^{\frac{1}{5}}}$$

$$= \frac{5.67}{5.54} = 1.02347 \qquad (6.4)$$

This change in the (geometric) mean price for the matched prices A to E is from February to March. The increase of 1.02347 when multiplied by the February price of 5.99 yields an imputed March price of $1.02347 \times 5.99 = 6.13$.

6.53 The price collector subsequently finds April and May prices for variety F to be temporarily missing. The respective imputed prices are: $1.0353 \times 6.13 = 6.15$ and $1.00351 \times 6.15 = 6.17$. The imputed prices are entered in Table 6.3, highlighted in bold, providing a complete table of prices for outlets A to F over the price reference period and subsequent months.

6.54 The short-term month-on-month price relative for all outlets A to F for February to March 2020 is given as follows:

$$P_J\left(P^{Feb\,2020}, P^{Mar\,2020}\right) = \frac{\prod_{i=A}^{F}\left(p_i^{Mar\,2020}\right)^{\frac{1}{6}}}{\prod_{i=A}^{F}\left(p_i^{Feb\,2020}\right)^{\frac{1}{6}}} = \frac{5.74}{5.61}$$

$$= 1.02317 \qquad (6.5)$$

6.55 The short-term price relatives for outlets A to F calculated from imputed prices, as in equation 6.5, provides the same measure as the price relative calculated from outlets A to E since the price for outlet F is computed from the price change for outlets A to E. Other short-term price relatives are shown in Table 6.3. The long-term price index (December 2019 = 100) to July 2020 is shown in Table 6.4 and in Table 6.3 as the cumulative product of short-term relatives.

Targeted Mean Imputation

6.56 The overall mean imputation is based on assuming that the price change of the temporarily missing variety is similar to the overall price change at a higher level of aggregation. A targeted form of the method would use price movements of an elementary aggregate or an aggregate of similar varieties, that is, varieties expected to have similar short-term price changes. The sample of observations used for the targeting may be specific to a type of outlet and region and cluster of features, for example, "up-market" television sets. It would generally be a subset of varieties within a higher level of aggregation. The decision to target the imputation using similar varieties in the subset or to use

Table 6.4   Overall Mean and Targeted Mean Imputations

	Overall Mean Imputation	Targeted Mean Imputation
	(a)	(b)
Dec.-19	100	100
Jan.-20	100.00 × 1.013712 = 101.37	100.00 × 1.013712 = 101.37
Feb.-20	101.37 × 1.007478 = 102.13	101.37 × 1.007181 = 102.13
Mar.-20	102.13 × 1.023765 = 104.56	102.13 × 1.026738 = 104.92
Apr.-20	104.56 × 1.003515 = 104.92	104.92 × 1.004396 = 105.38
May-20	104.92 × 1.003652 = 105.31	105.38 × 1.003766 = 105.78
Jun.-20	105.31 × 1.001977 = 105.52	105.78 × 0.997531 = 105.52
Jul.-20	105.52 × 1.008081 = 106.37	105.52 × 1.008081 = 106.37

a wider subset of the higher level of aggregation will depend in part on the adequacy of the sample size for the subset of similar varieties and the homogeneity of the elementary aggregate at the higher level.

6.57 Column B in Table 6.4 shows imputed prices for the missing variety prices in outlet F for March, April, and May 2020, based on adjusting the preceding period's price by the price movements of the remaining matched pairs of prices at other independent traders, D and E, rather than all outlets, presented in column A for comparison. Changes in the geometric mean price relative as applied to adjust the preceding period's price are:

$$5.99 \times (5.65 \times 6.90)^{1/2}/(5.49 \times 6.50)^{1/2} =$$
$$5.99 \times 1.04522 = 6.26 \text{ for March 2020; and}$$

$$6.26 \times (5.75 \times 6.90)^{1/2}/(5.65 \times 6.90)^{1/2} =$$
$$6.26 \times 1.00881 = 6.32 \text{ for April 2020; and}$$

$$6.32 \times (5.80 \times 6.90)^{1/2}/(5.75 \times 6.90)^{1/2} =$$
$$6.32 \times 1.00434 = 6.34 \text{ for May 2020} \qquad (6.6)$$

6.58 The price index is compiled as the cumulative product of the short-term price relatives, as shown in Tables 6.3 and 6.4, again highlighted in bold.

6.59 The higher levels used at this elementary stage of aggregation are country-specific and follow the country's CPI aggregation structure, as described in Chapter 8 (paragraphs 8.9–8.10) and Figure 8.1. The higher level might be a region and type of outlet. For example, in Figure 8.1, this would be Brand A of whole grain bread sold in supermarkets in the Northern region; if there is an insufficient sample size for Brand A, similar Brands A and B, or all brands, might be used for all types of outlets in the region. Imputation of the missing price by the average change of the available prices may be applied for elementary aggregates where the prices can be expected to move in the same direction. The imputation can be made using all of the remaining prices in the elementary aggregate. This is numerically equivalent to omitting the variety for the immediate period, but it is necessary to make the imputation (see 6.69–6.71).

6.60 An imputed price should always be directly compared with the actual price on the variety's return as this provides a self-correcting measure. For example, if the imputation was not accurate and showed decreases in prices over the period, when in fact the price of the variety sold elsewhere or, if not sold, being held in storage, was increasing, then a direct comparison between the last imputed and the returning actual price would bring the index back to

its longer-term trend. The long-term price indices in Table 6.3 using an overall imputation for June and July 2020 are 105.42 and 106.37, respectively. These are the same results as those given in Table 6.3 for targeted imputations. Both series have self-corrected the imputations to return to the long-term price changes as properly measured using actual prices in outlet F. The overlap method described in paragraphs 6.90–6.118 links in a replacement variety's price change and can be used for permanently missing varieties. The overlap method does not have this self-correcting feature and should not be used for temporarily missing varieties.

Carryforward Imputation

6.61   Carrying forward the last observed price should be avoided and is acceptable only in the case of fixed or regulated prices. Special care needs to be taken in periods of high inflation or when markets are changing rapidly as a result of a high rate of innovation and product turnover. While simple to apply, carrying forward the last observed price biases the resulting index toward zero change. In addition, when the price of the missing variety is recorded again, there is likely to be a large compensating step change in the index to return to its proper value. The adverse effect on the index will be increasingly severe if the variety remains unpriced for some length of time. In the example in Table 6.3, if carryforward imputation was used the missing prices in March 2020 would be imputed with the February 2020 price of 5.99, as would be the imputed prices in April and May 2020. In June, on the variety's return, there would be a step increase in price from May to June of 5.99 to 6.25. In general, carrying forward is not an acceptable procedure or solution to this problem. Exceptions may be well-established and well-advertised periodic increases of fixed or controlled prices and tariffs.

General Considerations

6.62   As a general principle, temporarily missing prices require an explicit imputation entered into the data compilation. The overall mean imputation should, by default, be based on a higher level of aggregation; however, it also may refer to a variety within a region or type of outlet. The higher level of aggregation may comprise more than one variety, some with different price changes. For example, if the missing price observation is for "canned tuna," where the higher weighted level aggregate is "canned fish" which includes "canned tuna" and "canned salmon," then subject to a sufficient sample size, the imputation should be based on price movements of "canned tuna."

6.63   An overall mean imputation benefits from an automation of its implementation, and the transparency of this methodology contributes to the integrity of the index. By using an overall imputation, NSOs guard themselves against criticisms of influencing the CPI by their choice of "similar varieties," particularly where there are missing prices for heavily weighted elementary aggregates. However, this method should not be used when there are strong a priori or empirical grounds to believe a targeted imputation would improve the results. NSOs should have retrospective monthly price data available at higher levels of aggregation and be able to examine differences in

short-term month-on-month price changes between the missing variety's price changes, and price changes of similar varieties and higher levels of aggregation, to choose between an overall or targeted imputation accordingly. In the much-simplified illustration of Table 6.3, price changes of supermarkets are very different from independent traders and a target imputation is illustrated for outlet F using independent traders. Outlet F might also have been imputed using the price index at a higher level of aggregation or even a single outlet's price. Similar principles of aggregation apply.

6.64   Typically, a price collector reports a variety price as temporarily missing; this is then passed to the head office for confirmation and then, perhaps, further at a higher level. If confirmed, a decision is made as to use a targeted or overall imputation. If the overall mean imputation is chosen, an appropriate computer routine is applied which enters the imputed price with a designation as being imputed into the data system. If a targeted imputation is used, the CPI compiler selects the variety rows regarded as likely to have similar price changes, and the imputation is applied accordingly. For quality assurance, the computer routine should record the decisions made and tabulate the number of temporarily missing varieties by elementary aggregate and their treatment.

6.65   Imputations are preferable to simply omitting the missing price observation for the calculation of the index. For example, consider a variety priced at 4 in January, temporarily missing in February, and returning at 6 in March. If the price change between January and February for the remaining varieties in the elementary aggregate was 25 percent (a price relative of 1.25). The imputed variety price for February would be $1.25 \times 4 = 5$. The actual price is not known, but the imputation serves as a benchmark without impairing the 25 percent price index measure based on observed matched prices. The February to March price relative for the missing variety is $6/5 = 1.20$. This is referred to as a self-correcting imputation; it self-corrects in the sense that the long-term February to March calculation would allow the index to return to its appropriate level: $5/4 \times 6/5 = 6/4 = 1.5$. Similar principles apply for the treatment of seasonal products as outlined in Chapter 11. Simply omitting the February price, and basing the February to March price relative on matched prices in these two months, would not have this self-correcting property.

6.66   Omitting a missing variety's price from the calculation is equivalent to an overall imputation using the other matched prices in the elementary aggregate. The entering of an imputed value allows for the flexibility of using a targeted imputation, if desired.

6.67   The designation of an imputed value by a different tab or color clearly shows in the prices database the extent to which imputations are used, the length of their runs, and product codes where they are overused. Summary counts of imputed prices should be monitored by type, product code, frequency, and duration as part of quality assurance. This is particularly important to identify if or where temporarily missing variety prices have continually imputed values well over a three-month run, with the possibility of continuing sample depletion as other variety prices go missing.

The Nature of Quality Change

6.68   Understanding the meaning of quality change requires well-defined conceptual and theoretical background, so that adjustments to prices for quality differences are made against this framework, as described in this section.

6.69   Over time, the quality of what is produced changes. For example, new automobiles typically have an increasing number of options and become more reliable, durable, safe, powerful, and economical. Another example is smartphones, as each new model includes faster processing speeds and power, more memory, improved screen resolution, and other technological advances such as facial recognition. In matching the prices of a sample of models selected in a price reference period with the same models in subsequent periods, the quality mix remains constant to avoid affecting the price measurement with quality differences. However, the resulting sample of models gives less emphasis to newer models that have benefited from more recent technological change and have different price changes given the quality of services they provide.

6.70   Observed changes in prices arise in theory from several sources, including quality changes, changes in tastes and preferences, and changes in the technology of producers. Price differences of similar products are often taken to be measures of differences in quality. However, differences in prices are often observed for varieties of the same quality. This may arise for a number of reasons:

- Some consumers may be unaware of the availability of the same varieties at lower prices, as there may be "search costs" to exploring the market to discover lower priced varieties.

- There may be price discrimination because the seller is able to charge different prices to different categories of consumers, such as movie tickets for children and senior citizens.

- Prices may be sticky with some retailers changing their prices infrequently to avoid the costs of doing so, including adverse customer reaction, or as strategic competitive behavior, such as loss leaders, leading to different retailers changing prices at different times.

- Cases where there are parallel markets, an official market subject to government or official control at which products are rationed and an unofficial unregulated market. The unofficial market may be at a lower price because it avoids taxes and regulations, or at higher price since the official price is a subsidized one, but has limited, possibly varying, quantities available for sale (*System of National Accounts 2008*, paragraphs 15.64–15.75).

6.71   In addition to the changing mix of the characteristics of varieties, 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. The same product provided at a different and more convenient location may result in a higher price and be considered to 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 products at these times and reflect supply-side pressure. Other differences, including the conditions of sale and circumstances or environment in which the products are supplied or delivered, can make an important contribution to differences in quality. For example, a retailer may attract customers by providing free delivery, financing, or better variety, by being more accessible, by offering shorter order times, smaller tailor-made orders, clearer labeling, 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. Although these sorts of benefits are not always specified in the variety description, such quality improvements should conceptually not be outside the scope of the index.

6.72   To consider how to adjust prices for quality changes, it is first necessary to define quality. While there may be an intuition as to whether a variety 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, a variety of clothing is sampled and, after a few periods, is missing. One option is to replace it with a similar variety, but the most comparable option may have more cloth in it, or have a lining, a different color or buttons, better stitching, or be considered more fashionable. There is a need to put a price estimate on the difference in quality between the old and new varieties so that the price of like can be compared with like. To propose or criticize a quality-adjustment procedure requires some concept of what is ideally required and how it is done in practice.

6.73   In Chapter 4 of *Consumer Price Index Theory*, a cost of living index is defined as the ratio of the minimum expenditure in the base and current period required to achieve a given standard of living or utility. Quality adjustments to prices involve attempting to measure the price change for a product that has exhibited some change in its characteristics from an earlier period that provides a different level of utility to the consumer. Equating of the value of a quality change with the change in utility derived by the consumer, while falling naturally under a cost of living index framework, is not exclusive to it. A cost of goods index can also benefit from regarding quality in this way. While a cost of goods index requires the pricing of a fixed basket of products, some varieties will become unavailable and the replacement varieties selected to maintain the sample may not be of the same quality. A cost of goods index based on a fixed-basket concept has the pragmatic need to adjust for quality differences when a variety becomes unavailable, and in the definition of a fixed-basket index does not preclude differences in utility being used as a guideline. If variety A is better than its old version, variety B, it is because it delivers more utility to the consumer who in turn is willing to pay more.

6.74   The definition of a quality change is based on equating some change in characteristics to a different level of utility provided. For example, consider the case in which a new variety with improved quality is made available, and the consumer has to choose between the old and new varieties in period t. If both varieties were offered to the consumer at the same price, $p^t = 100$, the consumer would naturally prefer the new variety. If the price of the old variety was

then progressively reduced until $p^{t*} = 75$, at which the consumer was indifferent between purchasing the old variety (at $p^{t*} = 75$) and the new variety (at $p^t = 100$), the consumer might select the old variety or the new one and would obtain the same utility from both. Any further decrease below $p^{t*} = 75$ would cause the consumer to switch back to the old variety.

6.75   The difference between p^t and p^{t*} would be a measure of the additional utility that the consumer placed on the new variety as compared with the old. It would measure the maximum amount that the consumer was prepared to pay for the new variety over and above the price of the old variety. In economic theory, if consumers are indifferent between two purchases, the utility derived from them is the same. The difference between p^t and p^{t*} (75 and 100) must therefore arise from the consumers' valuation of the difference in utility they derive from the two varieties: their quality difference.

6.76   The utility-based framework provides insights into the question of how consumers might choose between varieties of different qualities. Consumers derive more utility from a variety of higher quality than from a variety of lower quality, and thus they prefer it. But this does not explain why consumers buy one variety rather than the other. For this, it is also necessary to know the relative price of one variety with respect to the other, since if the lower-quality variety is cheaper, it may still be purchased ($p^{t*} \leq 75$ in the previous example) and to determine the price below which the old quality would be purchased.

Permanently Missing Price Observations

6.77   Tables 6.5A–C illustrate the treatment of permanently missing price observations. In these tables, prices are observed for outlets A to E over the seven months of December 2019 to June 2020. For outlet F, the price observations for the variety is reported as permanently missing from May 2020, and the price collector has to find a replacement variety (the prices in bold for May and June are imputed as explained in the following text). The use of replacement varieties maintains the original sample sizes at the last rebasing and the representativity of the varieties selected.

Informed outlet staff should be asked to confirm that the missing variety is permanently missing and help identify a best-selling replacement variety, its specifications, and how these specifications differ from those of the old variety. For logistical reasons the selected replacement variety would be expected to have high sales for the foreseeable future. The CPI staff at the head office would confirm or reject the choice of replacement variety.

Comparable Replacement

6.78   The price collector should use the missing variety's specification and identify a comparable variety with the same specifications, for example, a washing machine with the same spin speed, capacity, brand, or equivalent brand. If a comparable replacement exists, its detailed specifications should be confirmed by the price collector against the specifications for the missing variety. Any changes in the specification deemed to be not sufficiently price-determining should be noted for the head office to confirm, for example, color and trim.

6.79   The comparable replacement method requires the price collector to make a judgment that the replacement is of a similar quality to the old variety and any price changes are unaffected by changes resulting from quality differences. In Table 6.5A, there is a comparable variety F1 to replace variety F, also from outlet F. The replacement variety is considered by the price collector and confirmed by the head office as being directly comparable, and its prices in May and June (respectively, 6.20 and 6.25) are entered into the data system as a continuation of the outlet F series. The price index is calculated using short-term price relatives built into a long-term price index. The price index in July 2020 (December 2019 = 100) is 105.52, a 5.52 percent increase over this period. The price index remains as a constant-quality index since in May and June 2020 the prices of like quality varieties continue to be compared with like.

6.80   A common practice of manufacturers of electronic goods, such as televisions, household appliances, computers and computer-related hardware and software, and of automobiles is to have major quality changes in some years but relatively minor ones in other years. A new "comparable" model would have a new model number with a new

Table 6.5A   Illustration of Treatment of Comparable Replacements

Outlets		Price Reference Period						
		Dec.-19	Jan.-20	Feb.-20	Mar.-20	Apr.-20	May-20	Jun.-20
A	**Supermarket**	5.25	5.25	5.49	5.49	5.49	5.49	5.49
B	**Supermarket**	5.10	5.10	5.10	5.25	5.25	5.25	5.25
C	**Supermarket**	5.20	5.20	5.20	5.20	5.20	5.25	5.25
D	**Independent Trader**	5.49	5.49	5.49	5.65	5.75	5.80	5.80
E	**Independent Trader**	5.99	6.50	6.50	6.90	6.90	6.90	6.90
F	**Independent Trader**	5.99	5.99	5.99	6.13	6.15	**6.20**	**6.25**
Permanently Missing: F, May, June; F1: Comparable Replacement								
Geometric Mean: A:F		5.49	5.57	5.61	5.74	5.76	5.79	5.79
Short-Term Price Relatives: A:F		1.00000	1.01371	1.00748	1.02370	1.00347	1.00440	1.00134
Long-Term Indices as Product of Short Term		100.00	101.37	102.13	104.55	104.91	105.37	105.52

Values in bold are imputed.

131

production run, though physically not much has changed. The method of comparable replacement relies on the efficacy of the price collectors and head officer and, in turn, on the completeness of the specifications used as a description of the varieties. NSOs may tend toward designating replacements as comparable since they are concerned with sample sizes being reduced by dropping varieties, and the intensive use of resources to introduce noncomparable replacements or make explicit estimates of quality differences. The use of varieties of a comparable specification has practical advantages. However, if the quality of varieties is improving, the preceding variety will have inferior quality compared 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 replacements are accepted as being so despite quality differences, and the weight attached to those varieties. Chapter 8 includes proposals to monitor types of quality-adjustment methods by product group providing a basis for a strategy for applying explicit adjustments where they are most needed.

Noncomparable Replacements

6.81 Noncomparable replacements result when the price-determining characteristics of the replacement variety are different from those of the old variety. This means that the collected price of the replacement variety cannot be compared directly to the price of the old variety because the difference in these prices reflects not only pure price change, but also differences because of changes in quality. Noncomparable replacements require some form of quality adjustment.

6.82 Methods of quality adjustment for prices are classified into implicit (or indirect) quality-adjustment methods and explicit (or direct) methods. Implicit and explicit methods are discussed in paragraphs 6.90–6.188. Both decompose the price change between the old variety and its replacement into quality and pure price changes.

6.83 Implicit (or indirect) quality-adjustment methods estimate the pure price change component of the price difference between the old and new products based on the price changes observed for similar products. The difference between the estimate of pure price change and the observed

price change is considered as change due to quality difference. The most commonly used implicit method is the overlap method. The replacement's price change is linked to the old variety's price change using an overlap period that includes both the old and the replacement variety's price. Where an overlap price for the replacement variety does not exist, it might be imputed.

6.84 Explicit (or direct) quality-adjustment methods directly estimate the value of the quality difference between the old and new product and adjust one of the prices accordingly. Pure price change is then estimated as the difference in the adjusted prices. Explicit methods include quantity adjustments, option/feature costs, and "patched" hedonic regression methods.

6.85 Some of these methods are complex, costly, and difficult to apply. The methods used should as far as possible be based on objective criteria.

Implicit Methods of Quality Adjustment

6.86 This section discusses the following implicit methods for obtaining adjusting for quality differences: overlap method, class mean imputation, and link-to-show-no-change.

Overlap Method

The Use of an Overlap Price

6.87 A numerical illustration of the overlap method is shown in Table 6.5B. In this example, in outlet F there is a preexisting old model F available up to April 2020, and a noncomparable new replacement model, F2, available from May 2020, with the same actual price of 5.25 in both May and June. The prices are lower than would be expected from the prices of F, but this is a noncomparable replacement: it is a replacement variety with a major share and is expected to remain on the market in the foreseeable future.

6.88 The overlap method requires a price for both the old and new models in an overlap period. In the example in Table 6.5B, F exists up to and including April, F2 exists in May, June, and thereafter. The question is how to determine an overlap price in this case. One source of information for this overlap price (of 5.25 in April) is the price collector.

Table 6.5B   Illustration of Treatment Using the Overlap Method, Noncomparable Replacements: Actual Preceding Period Price

Outlets		Price Reference Period						
		Dec.-19	Jan.-20	Feb.-20	Mar.-20	Apr.-20	May-20	Jun.-20
A	Supermarket	5.25	5.25	5.49	5.49	5.49	5.49	5.49
B	Supermarket	5.10	5.10	5.10	5.25	5.25	5.25	5.25
C	Supermarket	5.20	5.20	5.20	5.20	5.20	5.25	5.25
D	Independent Trader	5.49	5.49	5.49	5.65	5.75	5.80	5.80
E	Independent Trader	5.99	6.50	6.50	6.90	6.90	6.90	6.90
F	Independent Trader	5.99	5.99	5.99	6.13	6.15		
Permanently Missing: F, May, June; F2: Noncomparable Replacement								
F2	Noncomparable Replacement					**5.25**	5.25	5.25
Geometric Mean: April, May, June: A:E,F2						5.61	5.63	5.63
Short-Term Price Relatives		100.00	1.0137	1.0075	1.0237	1.0035	1.0030	1.0000
Long-Term Indices as Product of Short Term		100.00	101.37	102.13	104.55	104.91	105.23	105.23

Values in bold are imputed.

Use the Actual Price of the Replacement in the Preceding Period, If It Exists

6.89   The price collector may have anticipated falling sales of a variety and the switch of consumers to a new model, brand, or variety, and recorded the price for the replacement prior to its adoption, obtaining in this way an overlap price. In the example in Table 6.5B, the price collector would start to record the price of the new variety F2 in April rather than May. Price collectors should be trained to anticipate such changes, to corroborate them with informed outlet staff, and to relay the information to the head office for possible action. In the example, the price collector would have seen diminishing sales and a poorer positioning for display in the outlet for the variety F, and it would have been apparent that it was to be replaced by F2. Outlet staff would confirm that F2 was to effectively replace F as a model aimed at that the same segment of the market, and the price for F2 and its quality characteristics should have been recorded alongside that of F to provide an overlap April price to facilitate the introduction of F2 in May. As a general principle, the replacement of models is best not undertaken when the old model has limited sales and is at the end of its life cycle.

6.90   Alternatively, the price collector may have asked informed outlet staff in May whether the new model was sold in the previous month to obtain an overlap price for April, or if sold in other outlets, whether there is a pricing agreement with the supplier that this outlet would have followed had it been supplied to them, and what would have been the price. The head office should confirm such details by visit or contact with informed staff of outlet F. Table 6.5B shows 5.25 to be entered as an *estimated* price in April 2020 for the new model, to provide an overlap price in April for the old and replacement models, F and F2.

6.91   The price index is measured through and including April 2020 using the prices of the old model: from Table 6.5A, the geometric mean of the prices of the old model in April 2020 is 5.76, and in March it is 5.74, with a short-term price relative of 5.76/5.74 = 1.0035. The price index to April 2020 is the cumulated product of the old index's price relatives: for April 2020, it is 104.91 = 104.55 × 1.0035 (December 2019 = 100).

6.92   In Table 6.5B, the index from May onward no longer uses F, but switches to F2. For this, there is a need for the overlap prices: average prices up to and including April using the old variety F and for May, June, and onward, using the new variety F2. The prices for outlet F in May and June (5.25 in both) are based on F2. The overlap in April for the new model F2 is 5.25. This overlap price is used to calculate the index for May. Using the overlap price, the short-term price relative for April to May in outlet F is 5.25/5.25 = 1.00000. This completes the price table. The geometric means are calculated as before, and their ratios form the short-term price relatives, and, in turn, the cumulative product of the price relatives, starting in December 2019 is the price index, at 105.23 for June 2020 (December 2019 = 100). The index from December to April is calculated using A–F and the index beginning in May onward is based on the prices of A–E and F2.

Imputed Overlap Prices

6.93   In the previous example, the new replacement variety may not have existed in April or the price collector may not have been able to obtain a reliable estimate of its price. In this situation, the overlap price in April can be imputed. The validity of this imputation is critical to the quality-adjustment methodology. In Table 6.5B, there was a price of 6.15 for the old model and 5.25 for the new model for April 2020. As shown in equation 6.6, this method implicitly attributes the price difference in the common April overlap period as an indicator of the quality difference between the old and new models.

6.94   If the new model was not sold in April, an imputation for the May price of the old model can be made to provide an overlap price. The imputation may be an overall mean or a targeted imputation following the principles outlined for temporarily missing variety prices (as described in paragraphs 6.52–6.72 and illustrated in Table 6.3). Table 6.5C illustrates a *forward* imputation of the old model's price to provide an estimate of the price in May 2020 had it existed then. The imputed price is given in Table 6.5C as 6.17. It is calculated by taking the ratio (relative) of the geometric mean of the May prices of outlets A to E to the geometric

Table 6.5C   Illustration of Treatment Using the Overlap Method, Noncomparable Replacements: Imputed Succeeding Period Price

Outlets		Price Reference Period						
		Dec.-19	Jan.-20	Feb.-20	Mar.-20	Apr.-20	May-20	Jun.-20
A	Supermarket	5.25	5.25	5.49	5.49	5.49	5.49	5.49
B	Supermarket	5.10	5.10	5.10	5.25	5.25	5.25	5.25
C	Supermarket	5.20	5.20	5.20	5.20	5.20	5.25	5.25
D	Independent Trader	5.49	5.49	5.49	5.65	5.75	5.80	5.80
E	Independent Trader	5.99	6.50	6.50	6.90	6.90	6.90	6.90
F	Independent Trader	5.99	5.99	5.99	6.13	6.15	**6.17**	
Permanently Missing: F, May, June; F3: Noncomparable Replacement								
F3	Noncomparable Replacement						5.25	5.25
Geometric Mean: May, June: A:F						5.76	5.78	
Geometric Mean: May, June: A:E,F3							5.63	5.63
Short-Term Price Relatives		100.00	1.0137	1.0075	1.0237	1.0035	1.0037	1.0000
Long-Term Indices as Product of Short Term		100.00	101.37	102.13	104.55	104.91	105.30	105.30

Values in bold are imputed.

133

mean of the April prices for the same outlets, 5.78/5.76 = 1.0037, and multiplying this by the old variety's price in April. The index in Table 6.5C is calculated by multiplying the short-term May to June price relative for the new replacement F3, 5.63/5.63 = 1.0000, by the value of the long-term index for May, 105.30. The 5.63 are the geometric means of the prices in outlets A to E and that of the replacement F3.

Forward versus Backward Imputation

6.95 In Table 6.5B, the overlap took place in April, while in Table 6.5C, it was in May. In Table 6.5B, an actual price was sought and found for the new (replacement) variety in April. There would be no equivalent May price for the permanently missing old variety. Given that an actual price may be preferred to an imputation based on the price movements of varieties priced in other outlets, the use of a *backward* price to provide an overlap in April can be used. In Table 6.5C, a *forward imputation* is used for the old variety's price in May, which is also an acceptable means of making an imputation. But one could apply a *backward imputation* for the new (replacement) variety's price in April to provide an overlap in April as the one in Table 6.5C. It is relatively straightforward to demonstrate algebraically that the result would be the same if the index is calculated either way. Both methods impute the missing price using the price movements of varieties in the outlets A to E for which prices exist in both April and May. The backward imputation is simply the inverse of the forward one. The ratio of prices in the overlap periods, as shown by equation 6.6, is the implicit measure of the quality differential between the old variety and the replacement variety, will be numerically the same for backward and forward imputations. As a result, either forward or backward imputations can be made.

Class Mean Imputation

6.96 The overall imputation method has many advantages considering the use of resources. It can be automated as a default measure to readily link-in replacement varieties keeping the sample up to date. However, the (forward) imputation assumes that the price movements of existing continuing varieties would be the same as that for the old variety, if it had continued to exist. A backward imputation assumes the price movement of the new (replacement) variety would be the same as existing continuing varieties had the new variety existed in the period prior to its introduction. Such assumptions are unlikely to be valid for high-technology goods being replaced at the end of their life cycle. An alternative imputation procedure designed to mitigate this problem is the class mean imputation which in principle, is more suited to this context of imputing replacement variety prices for permanently missing varieties, as opposed to temporarily missing ones.

6.97. *The class mean imputation method* is a specifically designed targeted imputation used to introduce a replacement when a variety's price is permanently missing. The class mean method of implicit quality adjustment to prices arose from concerns that unusual prices were charged at the start and end of a model's life cycle. Thus, the price movement of continuing varieties appears to be a flawed proxy for the pure price component of the difference between old and replacement varieties. A class mean imputation is mainly considered as a means of quality adjustment where there is a relatively high rate of frequent replacements, such as different models of automobiles launched each year.

6.98 The class mean method is a form of targeted imputation, for the treatment of replacements for permanently missing varieties, in which only the price changes of "comparable" replacements are used to impute the overlap price. The comparable replacements are being limited to those that have exactly the same price-determining characteristics, or those varieties with replacements that have been declared comparable after review or have already been quality-adjusted through one of the "explicit" methods, as described in paragraphs 6.120–6.188. For example, when the arrival of a new model of a particular kind of motor vehicle forces price collectors to find replacements, some of the replacements will be of comparable quality, and others comparable with explicit quality adjustments, but imputed prices for an overlap month will be needed for the remaining ones. Class mean imputation calculates imputed price relatives using only the prices of comparable and, where appropriate, explicitly quality-adjusted varieties or models. In general, it does not use the prices of the varieties or models that were not replaced, because these are likely to be different from those of the new models. The prices of old models tend to fall as they become obsolete, while the new models (represented by the replacements) tend to have a higher price before falling.

6.99 Class mean imputations rely on other explicit quality adjustments and comparable replacements. The other explicit quality adjustments may be from an available option or feature prices and may be limited in nature, covering only some of the differences in product attributes, available for only a small proportion of unrepresentative model or variety changes, and the availability of comparable replacements may be limited. Given a substantial churn in the market and difficulties with such imputations and estimates, an alternative recommended approach is that of hedonic indices, as outlined in paragraphs 6.140–6.176.

6.100 In some cases, sufficiently large samples of comparable substitutes or directly quality-adjusted varieties are unavailable, or the quality adjustments and selection of comparable varieties are not deemed sufficiently reliable. In that case, a targeted imputation may be considered. The targeted mean is less ambitious because it seeks only to capture price changes of similar varieties, irrespective of their point in the life cycle. However, it is an improvement on the overall mean imputation, as long as sufficiently large sample sizes are used.

Assumptions and Concerns on the Use of the Overlap Method

6.101 The accuracy of the estimates obtained with the overlap method depends on the validity of its underlying assumptions, as described in this section. If p_m^t and p_m^{t+1} denote the prices of an old variety m in periods t and $t + 1$, and p_n^{t+2} denote the price of a new replacement variety n in period $t + 2$, an overlap can be made by imputing a price for the new replacement variety in period $t + 1$, p_n^{t+1*}. In the case where variety n replaces m and is of a different quality, the measured price index between periods t and $t + 2$ (shown by the right-hand side expression in equation 6.7 is the price change of the old to new variety between these two

134

periods, multiplied by (adjusted for) the price overlap for m to n in period $t + 1$ which the method implicitly takes to be a measure of the quality difference.

6.102 A *forward* imputation is used in the previous example for the price of the old variety in period $t + 1$. Equation 6.7 also shows for this forward imputation that the overlap method depends on the validity of the relative difference in the prices of the old and new varieties in period $t + 1$ as a measure of the quality difference between the varieties, and on the reliability of the imputed price, designated with an asterisk (*), of the old variety m in period $t + 1$, p_m^{t+1*}, as an estimate of p_m^{t+1}.

$$I^{t,t+2} = \frac{p_m^{t+1*}}{p_m^t} \times \frac{p_n^{t+2}}{p_n^{t+1}} = \frac{p_n^{t+2}}{p_m^t} \times \frac{p_m^{t+1*}}{p_n^{t+1}} \qquad (6.7)$$

6.103. Table 6.6 shows an old model m permanently missing from March ($t + 1$) and replaced by a new model n in April ($t + 2$), with an overlap in March ($t + 1$). The price index for February (t) to April ($t + 2$), obtained using the overlap method, is given by the first expression in equation 6.7 as the product of the old model's price change between February and March and the new model's price change between March and April. This is equivalent to the second expression in equation 6.7, which is a direct price comparison between the new and old models between February and April with a quality adjustment as the value of the relative prices in the overlap month March ($t + 1$). The overlap method implicitly values the quality difference as the ratio of the two prices in the overlap period:

$$I^{\text{Feb,Apr}} = \frac{p_m^{\text{Mar*}}}{p_m^{\text{Feb}}} \times \frac{p_n^{\text{Apr}}}{p_n^{\text{Mar}}} = \frac{p_n^{\text{Apr}}}{p_m^{\text{Feb}}} \times \frac{p_m^{\text{Mar*}}}{p_n^{\text{Mar}}}$$
$$= \frac{30}{28} \times \frac{38}{35} = \frac{38}{28} \times \frac{30}{35} = 1.163 \qquad (6.8)$$

6.104 Moreover, for a longer-term price comparison, say January to June, the valuation of the quality difference remains as the price ratio in March, the time of the splice:

$$I^{\text{Feb,Apr}} = \frac{p_m^{\text{Feb}}}{p_m^{\text{Jan}}} \times \frac{p_m^{\text{Mar*}}}{p_m^{\text{Feb}}} \times \left(\frac{p_n^{\text{Apr}}}{p_n^{\text{Mar}}} \times \frac{p_n^{\text{May}}}{p_n^{\text{Apr}}} \times \frac{p_n^{\text{Jun}}}{p_n^{\text{May}}} \right)$$
$$= \frac{p_n^{\text{Jun}}}{p_m^{\text{Jan}}} \times \frac{p_m^{\text{Mar*}}}{p_n^{\text{Mar}}} = \frac{41}{25} \times \frac{30}{35} = 1.4057 \qquad (6.9)$$

6.105 The price of a missing variety is, by definition, not usually observed at the same time period as the price of the replacement variety, since the decision to replace the variety is only made after it has disappeared. Additionally, the list of specifications is not always comprehensive, since the main aim is to identify the variety in the outlet rather than to compare the varieties. However, the replacement variety may have been sold in the previous period and outlet staff may have a record of its price.

6.106 The underlying assumption in this case is that the quality difference in any period is equal to the price difference at the *time of the splice*. The *timing* of the switch from the old variety m to the new variety n is thus crucial. Unfortunately, price collectors usually maintain a variety until the decision to replace it is taken, so the switch may take place at an unusual period of pricing, near the end of variety m's life cycle and the start of variety n's life cycle. This analysis is more formally given in Annex 6.1.

6.107 Relative prices may not always reflect quality differences. For example, a new replacement model or brand of an improved quality may be stocked and sold at the same price as the old model, and the outlet competes in the market in part by changing the quality of what is sold, as opposed to the price. In other cases, retailers may reflect unusual pricing policies aimed at minority segments of the market. For example, the ratio of prices in an overlap period of a generic and a branded pharmaceutical drug may reflect the needs of two different market segments, rather than quality. The overlap method can be used with a careful choice of the overlap period. If possible, the overlap period should be a period before the use of the replacement, since in such periods the pricing may reflect a strategy to drop the old model to make way for the new one.

6.108 The overlap method is based on the law of one price: when a price difference is observed, it must arise from a difference in quality or similar 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 that markets are made up of rational producers and consumers. However, as outlined in paragraph 6.74, there are many reasons why identical varieties can be sold at different prices, and the law of one price does not hold in practice. These reasons include lack of information caused by search costs, price discrimination, and the existence of parallel markets.

6.109 The overlap method is commonly used as a default procedure for introducing replacements when varieties are permanently missing. NSOs that rebase the CPI less frequently may experience sample deterioration with many varieties becoming permanently missing and, without replacements, the sample becoming increasingly composed of imputed prices. Replacements serve to maintain the sample composition and to update the representativity of the varieties being priced.

6.110 In some cases, only noncomparable replacements (that is, those of a different quality) of varieties with missing prices are available. In these cases, an explicit adjustment to the price to account for the quality difference should be made to compare the prices of the noncomparable

Table 6.6   Introducing a Noncomparable Replacement via an Overlap

Model	January	February	March	April	May	June
Old (*m*)	25	28	**30**			
New (*n*)			35	38	40	41

Values in bold are imputed.

Table 6.6A   Introducing a Noncomparable Replacement to Illustrate Link-to-Show-No-Change

Model	January	February	March	April	May	June
Old (*m*)	25	28	**28**			
New (*n*)			35	38	40	41

Values in bold are imputed.

135

replacements in one period with the price of the now missing varieties in a previous period. The widely used overlap method has a major advantage of not requiring an explicit quality adjustment. Explicit quality adjustments (described in paragraphs 6.119–6.187) are more resource-intensive than the implicit overlap method. Furthermore, it is recommended that the overlap method be automated with information from the price collector, reported to the head office, and used in a computational routine. The recommendation is that in accepting the use of the overlap method, the CPI compiler should ensure that relative prices at the time of the overlap reflect quality differences. For example, a much-improved new model of a smartphone may be launched at the same price as the old model, and the relative price would not reflect the differences in quality. If the relative price does not reflect quality differences and resources allow, an explicit quality-adjustment method should be used.

6.111   The overlap method is implicitly employed when samples of varieties are rotated. That is, the old sample of varieties 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 variety-to-variety level—differences in price levels at a common point in time accurately reflect differences in qualities.

6.112   The bias in using the overlap method within an elementary aggregate depends on (1) the ratio of missing to total observations, and (2) the difference between the mean of price changes for existing varieties and the mean of quality-adjusted replacement price changes. As noted previously in paragraph 6.91, imputations can be made either forward or backward. The bias decreases as either of these terms decrease. A formal analysis is given in Annex 6.2.

Link-to-Show-No-Price-Change

6.113   Returning to the example in Table 6.6, reproduced in Table 6.6A, where the new replacement variety is noncomparable (that is, of a different quality), the link-to-show-no-price-change method imputes the price in March for the old model m to be the same as its February price, 28. The new variety is linked-in to show no price change in the period of replacement: February to March. This method is used for the treatment of a noncomparable replacement variety and should not be confused with carrying forward a previous period price for the treatment of temporarily missing prices (described in paragraph 6.65). This method biases the index downward when prices are rising and biases them upward when (true, quality-adjusted) prices are falling. The link-to-show-no-price-change method attributes the price difference between the new and old model in February to quality difference. The new model is of a better quality, valued at being worth an additional 7, from 28 to 35. The (quality-adjusted) price is therefore constant between February and March. The February to April price change is

$$\frac{28}{28} \times \frac{38}{35} = 1.0857,$$ an 8.57 percent increase in price.

In equation 6.6

$$p_n^{t+1} = p_n^{t*}\, p_m^{t+1*} = p_m^{t} \text{ and } I^{t,t+2} = \frac{p_m^{t}}{p_m^{t-1}} \times \frac{p_n^{t+2}}{p_n^{t+1}} = \frac{p_n^{t+2}}{p_n^{t}} \quad (6.10)$$

That is, a (quality-adjusted) price change between February (period t) and April (period $t+2$) is measured as that for the new variety between March ($t+1$) and April ($t+2$).

6.114   The bias is perpetuated through subsequent periods of measurement. For example, the February to June index would still have its price change measured as the product of month-on-month price changes which would include the February to March link that would show no price change.

6.115   The link-to-show-no-price-change method is not recommended. As with carryforward, the method is particularly harmful since it can be readily incorporated into a regular automatic compilation routine and not noticed: the price of the replacement is automatically imputed to form the overlap price and the index compiled.

Explicit Methods of Quality Adjustment

6.116   The methods described previously do not rely on explicit information on the value of the change in quality. This section discusses methods that rely on obtaining an explicit valuation of the quality difference: quantity adjustment; differences in production or option costs; and the hedonic approach.

Quantity Adjustment

6.117   Quantity adjustment is one of the most straightforward explicit quality-adjustment methods. It is applicable when the size of the replacement variety differs from that of the available variety, and any change in quantity is considered a change in quality. In some situations, there is a readily available quantity measure that can be used to compare the varieties. Examples are the number of units in a package (for example, paper plates or vitamin pills) or the size or weight of a container (for example, kilogram of flour or liter of cooking oil). Quantity adjustment to prices can be accomplished by scaling the price of the old or new variety by the ratio of quantities. The index calculation system may do this scaling adjustment automatically, by converting all prices in the category to a price per unit of size, weight, or number. For example, if the weight of a candy bar is 450 grams in the current period and 500 grams in the previous period, but the price remains unchanged, an adjustment is needed so that the index reflects the implicit price increase.

6.118   The specification of a variety is often to a specific size, for example, a one-kilogram packet of flour. If only two-kilogram packets are now sold in a specific outlet, the price collector should choose a representative two-kilogram packet, but mark the new specification as such and, after confirmation by the head office, prices should continue to be collected for the two-kilogram packet. In this example, an adjustment would be needed to ensure that the index reflects only pure price change and not changes resulting from differences in quality (that is, the change in size). This is particularly important where price variances are computed, or use is made of a Dutot price index number formula (that is, an elementary index formula sensitive to the homogeneity of the varieties used).[4]

---

[4]A Dutot elementary price index number formula is a ratio of arithmetic means of matched prices. More weight is given to price changes with higher prices in the reference period. Thus, if the price of a container in the price reference period doubles, while others remain the same, the implicit weight given to its price change will also double (see Chapter 8).

6.119 When replacing a variety, changes in the size of different varieties sold can be dealt similarly; however, there are some caveats. For example, in the pharmaceutical context, 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, for example, to a larger size container, and a *unit* price decrease of 2 percent accompanies this change, then it should not be regarded as a price fall of 2 percent 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 recommendation is not to automatically interpret unit price changes arising from packaging size changes as pure price changes, if contrary information is available.

6.120 Consider another example shown in Table 6.7: 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 price may have increased by $[(2.25/1.5) \times 100 = 150]$ 50 percent but the quality-adjusted price (that is, price adjusted by size) has remained

Table 6.7   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
0.50	1.50	3	1.50	3
0.75	2.25	3	2.25	3

constant $[(2.25/0.75) = (1.5/0.5) = 3]$; three per kilogram. The approach can be outlined in a more elaborate manner as illustrated by Figure 6.1. 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 price. The symbol Δ denotes a change. The slope of the line is β which is Δprice/Dsize = (2.25 − 1.5)/(0.75 − 0.50) = 3 (that is, 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, to make it equivalent to the new bag, n, is

$$\hat{p}_n^{t-1} = p_m^{t-1} + \beta \Delta size = 1.5 + 3(0.75 - 0.5) = 2.25 \quad (6.11)$$

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

$$\frac{p_n^t}{\hat{p}_n^{t-1}} = \frac{2.25}{2.25} = 1.00 \quad (6.12)$$

6.121 The approach is outlined in this form so that it can be seen as a special case of the hedonic approach (discussed in paragraphs 6.140–6.176), where the price is related to a number of quality characteristics of which size may be only one.

6.122 Now assume that the 0.5-kilogram bag was not available (missing) and a 0.25-kilogram replacement packet was used, priced at 0.75, as shown by the continuation to the coordinate (0.75, 0.25) of the unbroken line in Figure 6.1 and in Table 6.7; the quality-adjusted prices would again not change. However, if 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 6.7 and in Figure 6.1 (including

Figure 6.1   Quality Adjustments for Different Sized Varieties
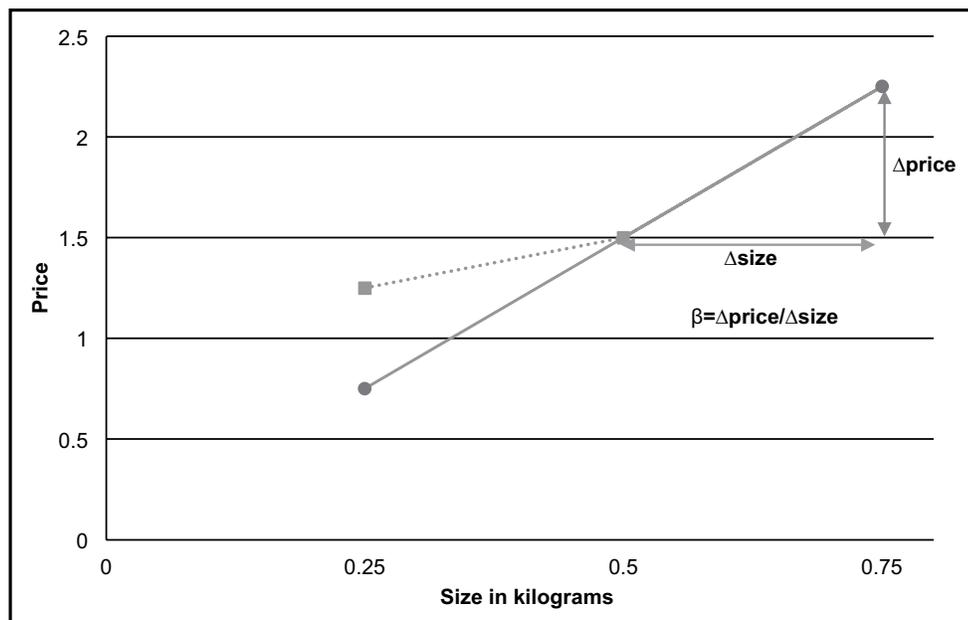
the broken line), then the measure of quality-adjusted price change would depend on whether the 0.5-kilogram bag was replaced by the 0.25-kilogram one (a 67 percent increase) or the 0.75-kilogram one (no change). This result is not satisfactory because the choice of replacement size is arbitrary. The rationale behind the quality-adjustment process is to separate pure price change from changes caused by differences in quality (in this case, quantity changes).

Differences in Feature/Option Costs

6.123   Consider an example of the *price* of an option being used to adjust for quality. Let the prices for a variety in periods $t-1$ and t be 10,000 and 10,500, respectively, but assume the price in period t is for the variety with a new feature, as standard, that previously in period $t-1$ had to be purchased as an "option" for an additional 300. Then between periods $t-1$ and t, the price change including the feature in both periods would be 10,500/10,300 = 1.01942 or 1.942 percent.

6.124   Option costs are useful in situations in which the old and new varieties differ by quantifiable characteristics that can be valued in monetary terms by reference to market prices. The valuation of a quantifiable product feature may be readily available from the comparison of different product prices. This is especially convenient because some goods and services sold on the internet can be identified by their brands and price-determining characteristics.

6.125   Consider the addition of a feature to a product, for example, an automatic icemaker in the door of a refrigerator, and refrigerators for a particular brand may be sold as standard or with a door-installed automatic icemaker. 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 icemaker. The cost of the option is thus known from before and a continuing series can be developed by simply adjusting the old price in the price reference period to include the option price. However, this process may have problems. First, the cost of producing an option as standard may be lower than when it was an option, and this saving may be passed on, at least in part, to the consumer. The option cost method would thus understate the price increase. Further, by including an option as standard, the consumer's valuation of it may fall since buyers cannot refuse it, and 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 they are difficult to quantify.

6.126   Quality differences are not necessarily positive, for example an airline may charge for a second piece of baggage when previously it did not. Again, there will be an option price available for the additional piece of baggage so that the price of like—two pieces of baggage—is compared with like.

6.127   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 an individual option/feature. The comparison is $\hat{p}_n^t / p_m^{t-1}$ 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 memory (RAM) of a computer when a specific model of computer is replaced by a model that is identical except for the amount of RAM it possesses. For example, the webpages of sellers of laptops allow buyers to customize their purchase, an extra four gigabytes of RAM, from eight to twelve gigabytes, for a specific brand and model of a laptop may cost an additional ¥70. Consider that the standard laptop used for CPI measurement has eight gigabytes of memory and costs ¥900 and is not available in the next period. The new standard model in period t has 12 gigabytes but costs the same ¥900, p_n^t. To compare the (constant-quality) price of the new model with the old model in $t-1$, the latter should have its price adjusted to include an extra four gigabytes of RAM. The price of an additional gigabyte of memory for this brand/model in period $t-1$ is 70/4 = 17.5, and its quality-adjusted price in period $t-1$ is $p_m^{t-1} = 900 + 17.5 \times (12 - 8) = 970$. The period t (unchanged) price of ¥900 is now compared with its comparable period $t-1$ price to yield a constant-quality price change of 900/970 = 0.9278, which is a price fall of 7.22 percent, while the package price is constant.

6.128   The previous description of the calculation is more complex than required: the adjustment is to simply add 70 to the old price, 70 + 900 = 970. However, it serves to demonstrate some limitations of this approach as a special case of the hedonic method, as outlined in the section on hedonic prices.

6.129   This calculation conveniently makes the quality adjustment to the old model's price in period $t-1$ so that the new model's price in future months can be directly compared with the *quality-adjusted* old price for the life of the new specification. However, the required information on the value of the option cost (an extra four gigabytes) may only be available in period t and not be applicable to a period $t-1$ adjustment. NSOs should ideally keep a record of, for example, web customizations of specified varieties along with comparable/noncomparable replacements especially for products with a high degree of technical change and turnover of models and maintain good relations with outlet staff.

6.130   In the previous example, if the relationship between price and RAM is linear, the previous formulation is appropriate. Many webpages give the price of additional RAM as being independent of other features of PCs, and a linear adjustment is appropriate. A linear formulation values the worth of an additional fixed amount of RAM to be the same, irrespective of the amount of RAM the computer possesses or a number of other features.

6.131   The relationship between price and the product features may be nonlinear. Denote the price-determining characteristics as z and assume there are k of them. The change in z is intended to reflect the service flow, but the nonlinearity in the price–z relationship may reflect consumer's decreasing marginal utility to the scale of the provision. In the previous example, the price a customer is willing to pay per gigabyte falls as increasing amounts of gigabytes are purchased. For some features, there will be economies of scale: supplying much more of a feature makes the price fall, possibly substantially; while for others, it may become technically difficult, and more expensive, to compress

138

higher amounts of a feature into the available space. The data should reveal some of this relationship, and caution is needed against applying linear relationships outside of the range in which they are warranted. Further, data should give some insight into the required adjustments for such nonlinear relationships, though this may be better estimated using a regression formulation and nonlinear specification, as considered in the section on hedonic prices.

6.132   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 estimates of the change in price resulting from a unit change in the option or size: the β slope estimates. In the case of the quantity adjustment, this was taken from a variety 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.

6.133   This same reasoning applies to option costs. For example, two varieties are identical except for a single feature. Their difference in price 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 a computer. 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 varieties and not all combinations of these may exist as varieties in the market in one period. Furthermore, the combinations existing in the second period being compared may be quite different from those in the first. Considering these aspects leads to a more general framework, known as the hedonic approach.

Differences in Production Costs

6.134   An alternative approach to quality adjustment is to adjust the price of an old variety by an amount equal to the resource costs of the additional features of the new variety. An important source of such data is the manufacturers. In this approach, the NSO can ask the manufacturer to provide data on direct and indirect production costs for the embodied quality change which would include research and development (R&D) costs, assembly and installation associated with the change, the manufacturer's established markup, the retail margin, and associated indirect taxes. This method is similar to estimating market-equivalent option prices in the absence of market prices. 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. Additionally, it can only be successfully implemented if there is good communication between manufacturers and the NSO staff. It is particularly suitable when the quality adjustments are also being undertaken to calculate the producer price index and export

and import price indices. As an example of the practical use of this approach, an NSO uses production cost estimates to value quality adjustments arising in from model changes in new vehicles. Allowable product changes for the purpose of quality adjustments include reliability, durability, safety, fuel economy, maneuverability, speed, acceleration/deceleration, carrying capacity, and related changes and additional parts required to accommodate the principal change in a component. Only price-determining characteristics are included, and any characteristics or features that do not affect or impact the price are excluded. Also excluded for the CPI, unlike the producer price index, are changes mandated by the government that provide no direct benefit to the purchaser, including modifications to meet air pollution standards. When a new model of a specified automobile is introduced, its changes in quality components are identified, valued, and added to the price of the old model so that the price of the old and new models can be compared on a like-to-like basis.[5]

6.135   A critical feature of this method is its reliance on estimates of the retail margin for the new components. With the option cost approach, a consumer's valuation of the new feature was available. If only production cost data are available, estimates of the retail markup must account for the (average) age of the models under consideration. Markups will decrease as models come to the end of their life cycles. Therefore, retail markups based on models at the start of their life cycle should not be applied to the production costs of models during their life cycle, and particularly at the end. Moreover, estimates of the retail margin of a component may well not be available. A pragmatic practice in one NSO is to use the proportionate retail markup on the vehicle. The proportionate retail markup is calculated based on the price charged by the manufacturer to the dealer for the identical vehicle and the manufacturer's suggested retail delivered price for the equipped vehicle.

Hedonic Approach: Patching

6.136   The hedonic approach is an extension of the two preceding approaches: the feature/option cost and the production cost approach. First, in the hedonic approach, the change in price arising from a unit change in quality (that is, the quantity or option/feature) is 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 determine price, rather than just the quantity or option/feature adjustment.

6.137   The hedonic approach is particularly useful when the market does not reveal the price of the quality characteristics required for the adjustment. Markets reveal prices of varieties, not quality characteristics, and it is useful to consider varieties as tied bundles of characteristics. A sufficiently large data set of varieties with their characteristics and sufficient variability in the mix of characteristics between the varieties allows the hedonic regression to provide estimates of the implicit prices of the characteristics. For example, the

---

[5]An illustration of the procedure used in the United States is available at: https://www.bls.gov/cpi/quality-adjustment/new-vehicles.pdf.

139

Figure 6.2   Scatter Diagram of Price against Capacity: Washing Machine Data
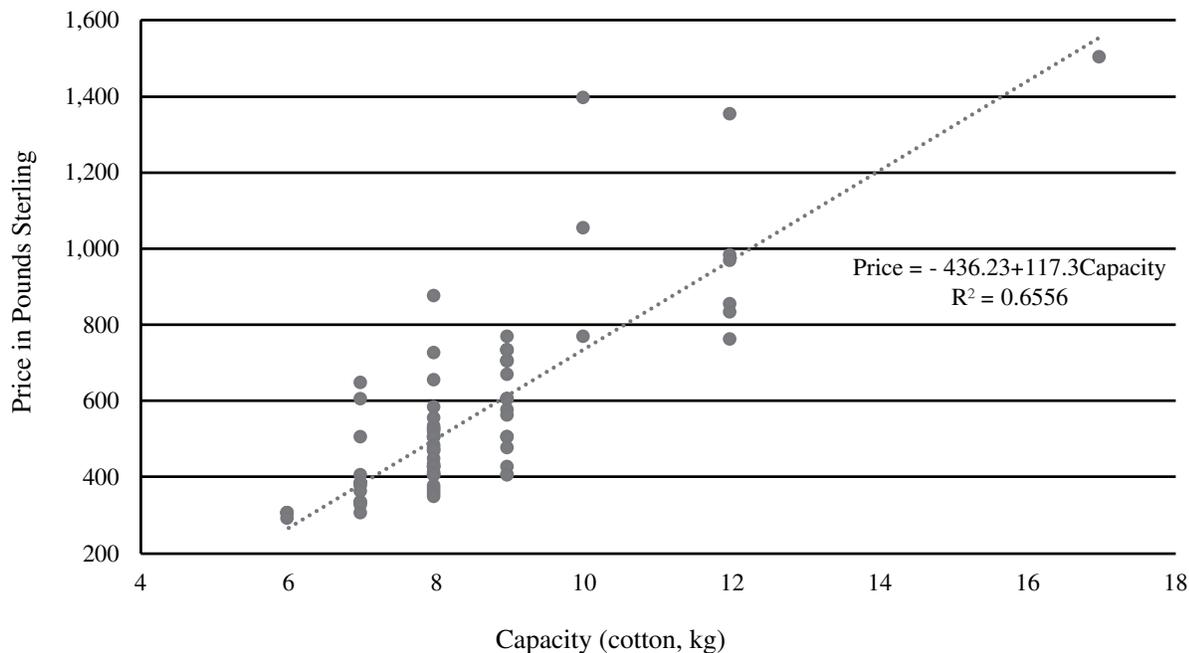
price of (clothes) washing machines will be listed, though a new (replacement) model for a brand may have a (cotton) capacity load size not previously available, say 12 kilograms, instead of the preceding model's 10 kilograms. To make an explicit quality adjustment, the price of the additional two kilograms is required. The regression approach using a data set of many models' prices and characteristics can estimate the price of additional kilogram of capacity from data for models of washing machines on their price, capacity, year (age of model), color, running cost, and so forth.

6.138   Under the MMM, each price collector needs to select a representative variety, record its price and specifications, and reprice the same variety in subsequent periods. The extension required in the hedonic approach is that the prices and price-determining characteristics should be collected for all, or a large sample of models. The method is particularly suitable when there are no immediately apparent comparable replacements and the noncomparable ones vary in their characteristics over more than one variable. A new model of car, household appliance, computer or related hardware and software, or telecommunications equipment, can differ from the old model in many respects, yet there is a single price for each new and old model. This approach is particularly necessary when there is a frequent turnover of models in the market, where new models with quite different values for their characteristics are frequently replacing old ones.

6.139   The requirement that data be collected on the prices and specifications of a large sample, if not all models, is not as demanding as it might appear. Extensive data on prices and characteristics of models of consumer goods and services are generally readily available on websites (for example, many comparing prices and salient characteristics),

can be copied with relative ease, and automated using web scraping. Such detailed information is also available as scanner data (see Chapter 10).

6.140.   Figure 6.2 is a scatter diagram relating the price (in pounds sterling, £) to the (cotton) capacity (kilogram) of models of washing machines sold in one country (the data are from a well-known consumer magazine). It is apparent that washing machines with larger capacities have higher prices—a positive relationship. It is also apparent from Figure 6.2 that there are several models of washing machine with the same capacity but quite different prices, resulting from the fact that other features differ. For example, 12-kilogram capacity machine prices range from £754 to £1,349.

6.141   To estimate the value given to additional units of capacity, an estimate of the slope of the line that best fits the data is required. The equation of a straight line is $Price = \beta_0 + \beta_1 z_1$.

6.142   The slope $\hat{\beta}_1$ is a measure of the change in *Price* that arises from a one-unit change in the characteristic, z_1, *Capacity*. The ^ (hat) above denotes that it is estimated from the data. The estimated slope is from the equation of a line that best fits the data (that is, that best represents the underlying pattern of the relationship). In Figure 6.2, the equation of the line that best fits the data was derived using an ordinary least squares (OLS) regression. The intercept and slope of the line that best fits the data are estimated as ones that *minimize the sum of the squared differences* between the individual prices and their counterpart prices predicted by the line: the least squares criterion. Tools for regression analysis are available on standard statistical and econometric software, as

140

Table 6.8   Estimated (Linear) Equation of Price against Capacity: Washing Machine Data

	Coefficients	Standard Error	t Stat	P-Value
Intercept	−436.229	86.310	−5.054	3.09E−06
Capacity	117.298	9.949	11.789	1.43937E−18

Table 6.9   Illustrative Hedonic Regression Estimates for Washing Machines

	Linear			Log-Linear		
	Coefficient	Std. Error	*p-Value*	Coefficient	Std. Error	*p-Value*
(Intercept)	−206.939	112.3	0.06986+	5.217	0.17	< 2E−16***
Age	−1.579	2.1	0.44971	−0.006	0.00	0.054218+
Capacity	81.024	13.2	5.32E−08***	0.108	0.02	4.86E−07***
Warranty	−138.651	48.7	0.00592**	−0.265	0.07	0.000466***
Steel	144.036	74.0	0.05608+	0.266	0.11	0.017575*
Energy Cost	10.103	3.4	0.00430**	0.019	0.01	0.000353***
LG	−115.816	43.9	0.01044*	−0.220	0.06	0.001167**
Steam	191.196	92.3	0.04233*	0.257	0.14	0.063056+
Hyg_AllergyP	63.627	40.2	0.11842	0.153	0.06	0.012198*
LED Display	49.409	52.9	0.35391	0.166	0.08	0.036833*
$\bar{R}^2$	0.701			0.721		
F-Statistic	20.25			22.22		
***P*-Value**	1.06E−15			2.20E−16		

***,**,*, and + denote statistically significant at 0.1, 1, 5, and 10 percent levels, respectively.

well as spreadsheets.[6] The estimated (linear) equation in this instance is presented in Table 6.8.

$$\text{Price} = -436.229 + 117.298 \quad \text{Capacity} \quad \bar{R}^2 = 0.65 \quad (6.13)$$

6.143   Formula 6.13 is the estimated regression equation of *Price* on *Capacity*; although there are many other price-determining variables, this regression equation only includes *Capacity*, for illustration. In Table 6.9, the regression model is expanded to include other variables.

6.144   The coefficient on *Capacity* is the estimated slope of the line: the change in price (£117.30) resulting from a one-kilogram change in *Capacity*. This can be used to estimate quality-adjusted price changes for washing machines of different capacities. The value of $\bar{R}^2$ (that is, the adjusted coefficient of determination) is 0.65; this indicates that 65 percent of price variation is explained by variation in *Capacity*. A *t*-statistic to test the null hypothesis of the coefficient being zero was found to be 11.789: recourse to standard tables on *t*-statistics found the null hypothesis was rejected with a *p*-value of 1.43937E−18. The fact that the estimated coefficient differs from zero cannot be attributed to sampling errors at this level of significance. There is a very low probability that the test has wrongly rejected the null hypothesis.

6.145   Hedonic regressions should generally be conducted using a semilogarithmic formulation. The focus is thus on the semilogarithmic form. The estimated (semilogarithmic) regression equation in this instance is

$$\log(\text{Price}) = 4.776 + 0.174 \quad \text{Capacity} \quad \bar{R}^2 = 0.61 \quad (6.14)$$

6.146   The coefficient of 0.174 has a useful direct interpretation: when multiplied by 100, it is the percentage change in price arising from a one unit (kilogram) change in capacity. There is an estimated 17.4 percent change in price for each additional kilogram of capacity.

6.147   The range of prices for a given capacity was noted to be substantial which suggests that other quality characteristics may be involved. Table 6.9 provides the results of a regression equation that relates price to a number of quality characteristics as listed in the Column A.[7] While the results are given for both linear and semilogarithmic regression specifications, the focus here is on the latter functional form.

6.148   A semilogarithmic hedonic multiple regression model is given by

$$\ln p = \beta_0 + z_1\beta_1 + z_2\beta_2 + z_3\beta_3 + \ldots z_n\beta_n + \varepsilon \qquad (6.15)$$

where ε is an error term assumed to have the usual properties to satisfy OLS assumptions (see paragraph 6.171). For this semilogarithmic form, logarithms are taken only of the left-hand side variable (that is, *Price*). Each of the *z* characteristics enters the regression without having logarithms taken. This has the advantage of allowing dummy variables for having or not having a feature included on the right-hand side.

---

[6]The illustrative empirical work in this section was undertaken using R, though it would be equally applicable with any such standard statistical software including EViews, SAS, and STATA.

[7]These include Age (months since model was launched); Capacity (for cottons, kilogram); Warranty (if five years, benchmark: two years); Steel (stainless steel outer); (Annual) Energy Cost, in £; LG (manufactured, benchmarked on Samsung); Steam (wash/refresh); Hygiene/Allergy program; LED display (benchmarked on LCD); and price for 74 models as advertised on November 2017 in *Which?* for three up-market brands: Brand A, Brand B, and Brand C. A much larger general regression model with more variables was first estimated but reduced to this smaller specific model using standard econometric principles and practice. The White (studentized) Breusch–Pagan test for homoscedastic residuals was not rejected at conventional significance levels with a *p*-value of 0.2197.

Such dummy variables assume the value of one if the variety has the feature and zero otherwise. The taking of logarithms of the first equation in 6.15 allows it to be transformed in the second equation to a linear form, and a conventional OLS estimator can then be used to yield estimates of the logarithms of the coefficients. These are given as the coefficients for the semilogarithmic model in Table 6.9. The estimated coefficients in Table 6.9 are based on a multiple regression model: for example, for *Capacity*, the estimated coefficient of 0.108 is of the effect of a unit change in capacity on price, *having controlled for the effect of other variables in the equation*. The scatter diagram in Figure 6.2 clearly shows the inadequacy of relying on a single price-determining variable and this approach can be justified because it addresses this issue. The preceding estimated coefficient of 0.174 was based on only one variable and is different from this improved result.

6.149   When dummy variables are used, the coefficients, when multiplied by 100, are estimates of the percentage change in price, given by $\left(e^{\beta_1} - 1\right) \times 100$. For example, from Table 6.9, Brand A models have a $\left(e^{-0.219743} - 1\right) \times 100 = 19.73$ percent lower price than their benchmarked Brand B counterpart, having controlled for other differences in their price-determining characteristics as specified in the regression equation.[8]

6.150   The value $\overline{R}^2 = 0.721$ is the proportion of variation in (the logarithm of) price explained by the estimated equation.[9] A high value of $\overline{R}^2$ can be misleading for the purpose of quality adjustment, although such values indicate that the explanatory variables account for much of the price variation over a relatively large number of varieties in the period concerned. This, of course, is not the same as implying a high degree of prediction for an adjustment to a replacement variety 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 variety whose price is to be predicted are from the means of the sample. The more unusual the variety, 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. A high $\overline{R}^2$ is a necessary, though not sufficient, condition for the use of hedonic adjustments.

The Interpretation of Estimated Hedonic Coefficients

6.151   Concerning the interpretation of the coefficients from hedonic regressions, there used to be a mistaken perception that these represented estimates of user value as opposed to resource cost. The former is the relevant concept in constructing a CPI, while for producer price index compilation it is the latter. Yet hedonic coefficients may reflect both user value and resource cost, both supply and demand influences. There is an identification problem, as referred in econometrics; the observed data do not permit the estimation of the underlying demand and supply parameters. What is being estimated is the actual point 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.

6.152   In many cases, the implicit quality adjustment to prices arising from the use of the overlap method may be inappropriate because the implicit assumptions are unlikely to be valid (as described in paragraphs 6.104–6.112). In such instances, the practical needs of reliable economic statistics require explicit quality adjustments. However, the use of the hedonic approach may only be justified when the weight, churn, and extent of the quality adjustment is substantial, due to the cost of implementing the method.

6.153   The proper use of hedonic regressions requires an examination of the coefficients of the estimated equations to ensure their plausibility. 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, make it unlikely that "reasonable" estimates will arise from such regressions. For example, a firm may cut a profit margin pertaining to a characteristic for reasons related to long-term strategic plans; this may yield a coefficient on a desirable characteristic that may even be negative. This situation does not invalidate 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. Over time, individual coefficients can show quite sensible patterns. Unreasonable coefficients on estimated equations are the exception and should be treated with some caution. Second, the CPI compiler should rely more on an estimated equation whose coefficients make sense, and which makes good predictions, than on 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 using, for example, variance inflation factors, to see if this is the case.

The Implementation of a Hedonic Quality Adjustment

6.154   The implementation of hedonic methods to estimate quality adjustments for matched noncomparable replacements can take two forms. The first form is referred to as "patching": undertaking a quality adjustment to the price of the old model to make it comparable with the new model. For many varieties, this can be seen as a one-off process for individual varieties within the lifetime of updating a sample. The second form is the more comprehensive process for rapidly changing high-technology products with substantial changes in quality within relatively short periods.

6.155   "Patching" is the term used in this Manual for introducing noncomparable replacements (that is, replacements of a different quality), using hedonic regression estimates.

---

[8]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. For the Brand A coefficient, the standard error from Table 6.9 is 0.064687; its variance is $(0.064687)^2 = 0.00418$: the adjustment is to add 0.00418/2 to −0.219743, giving −0.21765; a lower price of $-\left(e^{-0.21765} - 1\right) \times 100 = 19.56$ percent; see Triplett (2006) for further details.

[9]More formally, where p_i^t is the price of washing machine i in period t, $\overline{R}^2$ is 1 minus the ratio of the variance of the residuals, $\sum_{i=1}^{N}\left(p_i^t - \hat{p}_i^t\right)^2 / N$, of the equation to the variance of prices, $\sum_{i=1}^{N}\left(p_i^t - \hat{p}_i^t\right)^2 / N$. 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.

142

Table 6.10A   Hedonic Regression Imputation of New Variety's Price

Variety/Period	t	$t+1$	$t+2$	$t+3$	$t+4$
l	p_l^t	p_l^{t+1}	p_l^{t+2}	p_l^{t+3}	p_l^{t+4}
m	p_m^t	p_m^{t+1}	p_m^{t+2}		
n			$\breve{p}_n^{t+2}$	p_n^{t+3}	p_n^{t+4}

Consider varieties l, m, and n in Table 6.10A where variety l is available in all periods, the "old" variety m is only available in periods t, $t+1$, and $t+2$, and the replacement variety n is only available in period $t+3$ and subsequently. The varieties are defined by their z quality characteristics; for example, for variety m, in period t these are z_m^t and the price of variety m is p_m^t. The example assumes that there is no problem with comparing the prices of matched variety l with characteristics Z_1, as they have the same quality characteristics, but there is a problem when comparing varieties m and n. Variety m's replacement n is noncomparable, so p_m^{t+2} cannot be directly compared with p_m^{t+3}. An imputed price is required in order to have prices for both the old and new varieties in the same period. This could be achieved by imputing the price of the new variety n in period $t+2$ to form an overlap in this period with the actual price of the old variety m, in that period, as illustrated in Table 6.5C. This is a *backward* imputation. In this case, as illustrated in Table 6.10A, the overlap period is period $t+2$. However, variety n does not have a recorded price in period $t+2$, and it may not have been sold then. The backward hedonic *imputation* approach would predict the price of variety n in period $t+2$ using a hedonic regression estimated in period $t+2$ and the characteristics of the new variety n, taken from period $t+3$ (that is, the predicted price of variety n in period $t+2$, $\hat{p}_n^{t+2}$ —the hat over the price, $\hat{p}$, denotes a predicted value from the regression). The predicted prices are for the characteristics of the replacement variety n. This is an estimate of how much the price for the characteristics of the new replacement variety would have been if it had been sold in period $t+2$.

6.156   Where data are not available to support the monthly estimation of regression coefficients, as described in the previous paragraph, an alternative approach would be the hedonic quality-adjustment method.[10]

6.157   For short-term comparisons, an overlap method is used with a price relative for $t+2$ compared with $t+1$ given by p_m^{t+2} / p_m^{t+1}, and for $t+3$ compared with $t+2$ given by $p_n^{t+3} / \breve{p}_n^{t+2}$, and subsequently, without the need for an imputation, by p_n^{t+4} / p_n^{t+3}.

6.158   The simple example outlined before using data on washing machines is used here to illustrate the methodology. Assume that the linear regression equation 6.13 was estimated using period $t+2$ data, the old model m had a capacity of 10 kilogram, and the new model n in period

$t+3$ had a capacity of 12 kilograms. In this case, model n's price in period $t+2$ would be the predicted price, $\breve{p} = 117.3 \times 12 - 436.23 = 971.37$. The ratio of the actual price of model m in period $t+2$, for example, £750, to the predicted price in period $t+2$, is the quality adjustment shown for the overlap method in equation 6.6, though for period $t+2$ in this example, $\dfrac{p_m^{t+2}}{p_n^{t+2^*}}$, that is, $\dfrac{750}{971.37} = 0.7721$. The models are not comparable. The new model in period $t+2$ is more expensive even when its superior quality, its capacity, has been considered.

6.159   Given the availability of an estimate of the worth of an extra unit of capacity, an alternative approach would be to simply add 2×117.3 to the price of m in period $t+2$, rather than use predicted prices. Such use of individual coefficients is not recommended. In practice, a hedonic regression will include several explanatory price-determining variables that may be linearly related and thus not strictly independent. For example, larger (higher-capacity) washing machines may also have higher spin speeds or be more likely to have a steam feature. The estimated coefficient of each such multicollinear variable would be imprecise, though the predicted price of a regression equation that includes them would be unbiased.

6.160   With the option cost example, the quality adjustment might be for a single characteristic and an explicit valuation of the price of further units of this characteristic (for example, a gigabyte of storage for a computer, available from another source). Hedonic regressions are used where the market does not reveal the implicit shadow prices of individual characteristics; these shadow prices have to be estimated from price data for many varieties with differing bundled sets of characteristics.

6.161   The hedonic method makes use of short-term month-on-month comparisons. Predicting the price of variety n in period $t+2$, if it was sold then, is only for this one-off period as the new variety replaces the old, with a quality adjustment. Variety n's characteristics are held constant for month-on-month comparisons from $t+2$ onward, and variety m's characteristics are held constant for month-on-month comparisons from period t up to, and including, period $t+2$.

6.162   Alternatively, a forward imputation might have been used, a procedure similar to that adopted in Table 6.5C. The price of variety m might be predicted from a hedonic regression run on period $t+3$ data, $\hat{p}_m^{t+3}$. As with the preceding methodology, a predicted price is only required for the overlap period, after which the replacement variety forms the continuing index. It is not obvious which of the two approaches, predicting prices for m or n, is preferred. Resources permitting, a geometric mean of the two would be defensible, as would a clear rule from the outset as to the method applied based on some retrospective research on the outcome of using either method for particular product groups.

6.163   Table 6.5C shows that the backward and forward imputation methods yield the same result when the imputation is based, for both methods, on the price movements of varieties available in all periods. However, in this case, the backward prediction is based on a hedonic regression run in period $t+2$ and the forward imputation on a hedonic regression run in period $t+3$. The practical advantage of

---

[10]Jack Triplett. 2006. *Handbook on Hedonic Indexes and Quality Adjustment in Price Indexes*. Paris: OECD. https://www.oecd-ilibrary.org/science-and-technology/handbook-on-hedonic-indexes-and-quality-adjustments-in-price-indexes_9789264028159-en.

143

running a hedonic regression in a prior period argues for a *backward* imputation, as in Table 6.10A, as the most feasible procedure.

6.164 A refinement to these approaches is to use predicted values, for varieties *m* and *n,* in the overlap period, p_m^{t+3} / p_m^{t+2}. For this purpose, consider a misspecification problem in the hedonic equation, for example, there may be an interaction effect between a brand dummy and a characteristic. Having a characteristic for a particular brand may be priced higher than all other brands, say a 5 percent premium. The use of $\hat{p}_m^{t+3} / \hat{p}_m^{t+2}$ would be misleading since the actual price in the denominator would incorporate the premium, while the one predicted from the hedonic regression 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 omitted variable (interaction term) bias. The dual imputation approach is preferred whenever there are concerns about the suitability of the regression equation's specification to fully model prices, as would generally be the case.

6.165 A further approach would be to not use a replacement variety. Variety *m*'s characteristics would be held constant in the comparison from period *t* + 2 onward. However, this would require a hedonic regression being run for each subsequent period, $\hat{p}_m^{t+3}$, $\hat{p}_m^{t+4}$. It would also lead to a continuing degradation of the sample as an obsolete old variety *m* would have its characteristics repeatedly priced into the future, rather than being replaced by a new variety. For this reason, this method is not recommended.

6.166 In the previous examples, short-term price comparisons are used and are preferable to long-term ones. A long-term equivalent of Table 6.10A is shown in Table 6.10C. A predicted price for any replacement variety *n* in its month of introduction is estimated for the reference period *t* using a hedonic regression based on that period's data. The regression is estimated using period *t* prices and characteristics, but the predicted prices are for the characteristics of the replacement variety *n* in *t* + 3 and subsequently. It is an estimate of what the characteristics of the new replacement variety would have been priced at had it been sold in period *t*.

6.167 The long-term method has the significant advantage of only requiring a hedonic regression to be estimated in the single reference period. For periods *t* + 3 and *t* + 4, the price relatives are $p_n^{t+3} / \hat{p}_n^t$ and $p_n^{t+4} / \hat{p}_n^t$, respectively. However, as time passes, such comparisons become less meaningful. For example, comparing the actual price this month of a model of a laptop with one predicted 18 months ago using the hedonic approach, would estimate market valuations of each characteristic which is then applied to the characteristic set of a laptop sold now. Indeed, the need

for a double imputation becomes more important as time passes by, yet a double imputation requires monthly estimation of hedonic regressions that hinder the advantage of this approach. If hedonic regressions are to be used on this long-term basis, it is important that the regressions are reestimated regularly at a rate that will depend on the rate of the technological innovations, and changes in consumer preference specific to that product. For example, it may be that consumer's valuations of characteristics of washing machines, including spin speed, front-loaders, capacity, or number and types of wash programs, are fairly constant over time, even if the technology is changing rapidly. Frequent, say monthly, updating of estimated hedonic regression equations is not required. Prior empirical studies on the stability over time in hedonic characteristics would be valuable in this respect. As a general principle, short-term hedonic imputations are preferred to long-term ones.

Limitations of the Hedonic Approach

6.168 The limitations and challenges of implementing the hedonic approach should be considered by the NSO:

(1) First, the hedonic approach requires statistical expertise for the estimation and maintenance of the hedonic regression equations. The availability of user-friendly statistical/econometric software with regression tools makes this less problematic. Yet staff must possess sufficient expertise and understanding of statistical regression methodology applied to hedonic regression equations, and the interpretation of the results and diagnostic statistics of regression models.

- Statistical and econometric software carry a range of diagnostic tests to help judge if the final formulation of the model is satisfactory. These include $\bar{R}^2$ as a measure of the overall explanatory power of the equation, and F-test and t-test statistics that test whether the differences between the estimated coefficients of the explanatory (price-determining) variables are jointly and individually different from zero at specified levels of statistical significance. These statistics make use of the errors from the estimated regression equation.

- The regression equation can be used to predict prices for each variety by inserting the values of the characteristics of the varieties against the estimated coefficients of the explanatory variables. The differences between the actual prices and these predicted results are the residuals. Statistical/econometric software calculate predicted values and residuals as a routine. A hedonic regression equation estimated using OLS requires assumptions on the nature of the distribution

Table 6.10B Hedonic Regression Imputation of Old Variety's Price

Variety/Period	t	$t+1$	$t+2$	$t+3$	$t+4$
l	p_l^t	p_l^{t+1}	p_l^{t+2}	p_l^{t+3}	p_l^{t+4}
m	p_m^t	p_m^{t+1}	p_m^{t+2}	$\breve{p}_m^{t+3}$	
n				p_n^{t+3}	p_m^{t+4}

Table 6.10C Hedonic Regression Imputation of New Variety's Price

Variety/Period	t	$t+1$	$t+2$	$t+3$	$t+4$
l	p_l^t	p_l^{t+1}	p_l^{t+2}	p_l^{t+3}	p_l^{t+4}
m	p_m^t	p_m^{t+1}	p_m^{t+2}		
n	$\breve{p}_n^t$			p_n^{t+3}	p_n^{t+4}

144

of the residual errors. These include: (1) the error term has a constant variance—if this assumption is violated, the errors are heteroscedastic; consequently, standard tests of statistical significance can be biased and unreliable; (2) that explanatory variable(s) are not correlated with the error term, they are not endogenous—this is particularly important when explanatory price-determining characteristics are omitted from the hedonic regression: if an omitted variable is correlated with an included one, the estimated coefficient on the included one is biased; and (3) the price-determining explanatory independent variables are not truly independent, but correlated with each other—multicollinearity; the coefficient estimates and their tests become sensitive to change in the model and data. While the estimated coefficients are imprecise, the predicted prices in a hedonic regression would be unbiased.

- A full account of all OLS assumptions, consequences, means of detection of violation, and treatment, that may involve use of an alternative to OLS estimators, can be found in any introductory econometrics/statistical text. Modern software provides the appropriate tests for, and means of overcoming, violations of these assumptions and thus, validation of the hedonic model used. It is recommended that the NSO develops and publishes detailed metadata on the hedonic regression model used and its supporting diagnostic statistics to demonstrate the validity of the model and satisfy the need for transparency.

(2) Second, the estimated coefficients require regular updating. Consider that the predicted price is for the new model in a reference period, as presented in Table 6.10C. Although it might seem that there is no need to update the estimated coefficients each period, the valuation of characteristics in the price reference period may be quite out of line with their valuation in the new period. For example, quite dramatic falls in the price of storage and processing speed of computers, among other attributes, make the valuation of additional GBs of a new model, introduced a few years after the hedonic regression was estimated, a less meaningful exercise. Continuing to use the coefficients from some far-off period to adjust prices in the current period is similar to using out-of-date reference period weights. The comparison may be well defined but have little meaning. There is a need to update the hedonic regression estimates if they are considered out of date, because of changing tastes or technology, and to 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 instability in the parameter estimates of the hedonic regression over time.

(3) 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 group, or webpage, and then used to adjust noncomparable prices for varieties sold in quite different outlets, there must at least be an intuition that the marginal price differences for characteristics are similar between the outlets. A similar principle applies for the brands of varieties used in the sample for the hedonic regression. It should be borne in mind that high $\bar{R}^2$ 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 varieties and brands. They are not a measure of the prediction error for a particular variety, sold in a specific outlet, of a given brand in a subsequent period, though they can be an important part of this.

(4) Fourth, the functional form and choice of variables to include in the model should be considered. Simple functional forms generally work well, though there is a class of more complex flexible-functional forms. These include linear, semilogarithmic (logarithm of the left-hand side), and double-logarithmic (logarithms of both sides) forms. Semilogarithmic models are often employed since many of the price-determining explanatory variables are binary, 1 or 0, depending on whether or not a model has a particular feature (dummy variables). The specification of a model should include all price-determining characteristics. 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 analyzing 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.

(5) Fifth, the resources requirements for hedonic regression should be considered. Hedonic regressions require data on prices and price-determining characteristics for varieties (models) sold. Extensive data sets may be readily available on the internet or from scanner data, containing all pertinent price-determining characteristics, either from the websites of individual retailers or specialist websites comparing prices and features of laptops, household appliances, and many other such goods and services. For example, the data used for the previous example of (patched) hedonic explicit quality adjustments for washing machines was taken from a website and was copied and pasted relatively quickly. Web scraping software can reduce even this workload substantially.

(6) Finally, while data and software may not be problematic, NSO staff resources will be required in devising the specification, estimation, and validation of the estimated hedonic model for each product. Such hedonic models should be estimated regularly prior to their use in the CPI and the results made available as part of the detailed metadata for the purpose of transparency and feedback. In this regard, the resource requirements can be substantial compared with an implicit overlap method. At least at first, hedonic methods should be applied only to products with a relatively high weight and profile for which the implicit assumptions of alternative methods are found to be invalid and badly distort the results, especially if they provide a reputational risk to the NSO.

145

6.169 Hedonic methods may also improve quality-adjustment procedures in the CPI by indicating which product attributes do *not* appear to have material impacts on the prices. That is, if a replacement variety differs from the old variety only in characteristics that have been rejected as price-determining variables in a hedonic study, this would support a decision to treat the varieties as comparable. Care has to be exercised in such analysis because a feature of multicollinearity in regression estimates is an imprecision of the estimated parameter estimates. This may give rise to statistical tests that do not reject null hypotheses that are false. However, econometric/statistical software provides the tools to explore the nature and extent of multicollinearity; these include variance inflation factors. The results from variance inflation factors provide valuable information on the nature and extent to which different explanatory variables (characteristics) are interrelated and this in turn can help in the selection of replacement varieties. The results from hedonic regressions thus have a role to play in identifying price-determining characteristics and may be useful in the design of quality checklists in price collection.

## Choice between Quality-Adjustment Methods

6.170 The choice of the method to be used for quality adjustments is not straightforward. The CPI compiler must consider the technology and market for each product and devise appropriate methods, considering that the methods selected for one product area might not 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 product may lead to less resource-intensive methods for others. The methods adopted for individual product groups 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 in this chapter. A good understanding of the methods, and their implicit and explicit assumptions, is essential to the choice of an appropriate method.

6.171 Figure 6.3 provides a guide to the decision-making process. Assume that the MMM is being used. If the variety is matched for repricing in a subsequent period, there is no change in the specifications and no quality adjustment is required. This is the simplest procedure. However, there is a caveat: if the variety belongs to a product group where model replacement is rapid, and replacements are noncomparable, the matched sample may become unrepresentative of the universe of transactions. Continued long-term matching would deplete the sample. This is a matter for the frequent rebasing and maintenance of the sample (see Chapter 7).

6.172 Consider a variety found to be *temporarily missing*. If it was a seasonal product, its treatment would follow the principles and practices outlined in Chapter 11. If it was temporarily missing but not a seasonal product, a price imputation is required, and if subsequently determined to be permanently missing (for example, either from information from outlet staff or use of a three-month rule) a replacement needs to be found. *Overall or targeted price imputations* for temporarily missing prices may be used; the carryforward method is not recommended except for controlled or regulated prices.

Figure 6.3 Guide to Treatment of Missing Prices
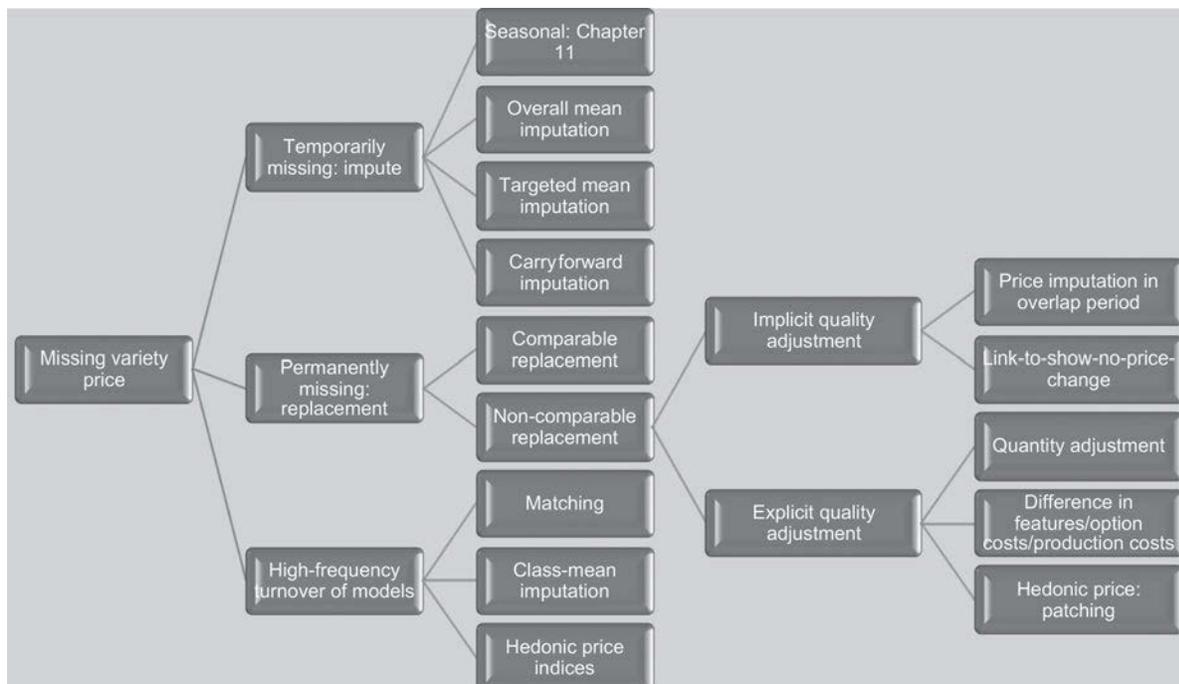
6.173   For *permanently missing* varieties, the selection of a comparable variety is preferred, as is the use of its price as a *comparable replacement* price which is then directly compared with the preceding variety price. This direct price comparison would require that none of the price difference between the comparable replacement and the previous variety is attributable to quality, and confidence that all price-determining factors are included in the specification. In practice, varieties may be considered comparable if there are limited price-determining differences, as might be the case with styling, color, even some more substantial technical changes, including performance and reliability, that may not be immediately apparent to the consumer. A decision on the comparability of a replacement must be made by CPI staff with appropriate information on product differences supplied by the price collector. A comparable replacement variety should also be representative and account for a reasonable proportion of sales. Caution is required when replacing near obsolete varieties 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 in paragraphs 6.182–6.222 and in Chapter 7, including early substitutions before pricing strategies become dissimilar. With comparable replacements, the price of the old variety is directly compared with the price of the comparable replacement in the next period.

6.174   Figure 6.3 considers the case where only *noncomparable replacements* are available. If explicit estimates of the price dimension of the quality differences are unavailable, and no replacement varieties are deemed comparable, *implicit estimates* might be used. One such method is the continued use of *imputations* as they are applied to temporarily missing varieties. Such use is *not recommended* as a default procedure.

6.175   The use of imputations has advantages resource-wise, as it is relatively easy to employ and requires no judgment (unless it is a targeted mean imputation) 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. The bias from using imputations for permanently missing variety prices is directly related to the proportion of missing varieties and the difference between quality-adjusted prices of available matched varieties and the quality-adjusted prices of unavailable ones. The nature and extent of the bias depends on whether short-term or long-term imputations are being used (the former being preferred) and on market conditions. Imputation, in practical terms, produces the same result as deletion of the variety for an elementary aggregate. The inclusion of imputed prices may give the illusion of larger sample sizes. Imputation should by no means be the overall catch-all strategy, and NSOs are strongly advised against its use as a default device that may lead to serious sample degradation.

6.176   Imputations can be used to extend the period of search for a replacement, though the absence of the old variety and the unavailability of a replacement should indicate to the CPI compiler that the weight for that variety might be better attributed to a quite different variety. Such changes naturally take place on updating an index, as described in Chapter 7.

6.177   If the old and replacement varieties 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 varieties in a period in which they both exist is assumed to be attributable to quality. This *overlap method*, in replacing the old variety 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 varieties are taken. The assumption of relative prices equating quality differences at the time of the splice is unlikely to hold if the old and replacement varieties 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 variety to clear inventories, and price skimming of market segments that will purchase new varieties at relatively high prices. As with comparable replacements, early substitutions are advised so that the overlap is at a time when varieties are at similar stages in their life cycles. It may well be the case, however, that overlap prices are unavailable. In such cases, a range of imputation approaches is available to estimate an overlap price.

6.178   The quality differences between the replacement and missing variety may be explicitly quantified. *Explicit estimates* of quality differences are generally considered to be more reliable, although they are also more resource-intensive, at least initially. Once an appropriate methodology has been developed, it can often be easily replicated. General guidelines are more difficult to provide as the choice depends on the factors already discussed in this chapter, 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. Estimates based on objective data are preferred. A relatively straightforward quality adjustment is when the quantity differs. The standardization of quantity units sold across outlets, for example, to price per kilogram, is relatively straightforward, though a change in the quantity of a variety included in the price—a quantity adjustment—may be more complicated than expected.

6.179   The replacement variety may differ from the old one for *having a different characteristic*. Often it is the *price collector* who is best placed to provide an estimate of the price difference in quality of a noncomparable replacement. For example, if a specified brand of a bottle of tomato ketchup used for pricing is missing in the current period, and a noncomparable replacement of the same brand is available, though the bottle has been restyled to now stand on its head, and the label has been reversed. The price collector might note that other brands have both sizes sold with a 25 percent price margin for the new one. The price collector in selecting a noncomparable replacement might also provide the basis for the head office staff to make an explicit quality adjustment. The head office staff might also make *use of the internet* to identify the percentage markup for a quality characteristic, for example, for additional memory for a computer or Bluetooth technology in an automobile. The *option cost* approach is applicable when a new feature is first sold as an option and then becomes a standard component included in the basic price. This requires that the old and new varieties differ by easily identifiable characteristics that are or have been separately priced as options. The use of production cost estimates critically relies on the availability of suitable estimates for the price-cost margin.

147

6.180   The use of *hedonic regressions for patching* price changes because of quality differences 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 with regard to 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 should be collected from the prices of varieties 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 a variety 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 variety 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 noncomparable replacements, though not a frequent churn, and where the differences between the old and new varieties can be well defined by its characteristics.

6.181   As previously discussed in this chapter, the use of the *link-to-show-no-price-change* method for permanently missing and the *carryforward* method for temporarily missing variety prices are not generally advised for making quality adjustment and imputations.

6.182   While Figure 6.3 is appropriate for the treatment of temporarily and permanently missing prices in the routine compilation of a CPI, there is a context in which a quite different strategy is required. The context is where there is a rapid turnover or "churn" in the models or varieties sold. For example, television sets are sold by several manufacturers each having a range of models with different features. Over time many new phases of technological development have occurred including the cathode ray tube, color televisions, wireless remotes, plasma, LCD televisions, digital, high definition, larger screens, smart functions, 3D, LEDs, ultra HD resolution, OLED, and roll-up OLED. New features and restyling extend the life cycle of each model in each phase. As with automobiles, computers, computer-related hardware and software, telecommunications equipment, or household appliances, the product market is characterized by different manufacturers producing several varieties (models) each of different quality, such as screen size for a computer or television set, aimed at different segments of the market. These will, over time, usually have a rapid turnover in their quality characteristics. The previously outlined methods, if applied to these markets, may lead to a biased CPI. Figure 6.3 notes that *matching, class mean imputations,* and *hedonic price indices* may be used, though there may be severe bias in the use of the former. The next section considers CPI measurement for these product markets.

High Technology and Other Sectors with a Rapid Turnover of Models

6.183   The measurement of price changes of varieties unaffected by quality changes is primarily achieved by matching models; however, when the matching breaks down the implicit or explicit methods can be used. But what should be done in the case of industries where the matching breaks down on a regular basis because of the high turnover in new models of different qualities than 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 varieties consumed and a static universe of the varieties selected for repricing. For example, if the sample is initiated in December, by the subsequent May, for a long-term price comparison, the static universe will be matching prices of those varieties available in the static universe in both December and May but will omit the unmatched new varieties introduced in January, February, March, April, and May, and the unmatched old ones available in December but unavailable in May. For December to May cumulative month-on-month short-term comparisons, similar considerations apply. Although there will be improved imputations for temporarily missing variety prices and an improved timelier introduction of replacements, the replacements only draw from the dynamic universe of new models on a one-on-one basis. This example refers to a December to January matched price comparison. For many countries, matching may effectively continue for many years until the CPI is updated leaving an extremely degraded sample. Two empirical questions indicate 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 old varieties likely to have quality-adjusted prices that substantially differ from those of the matched varieties in the current and the reference periods?

6.184   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 reference period. It may be that the models that are disappearing have relatively low prices, while the entrants have relatively high ones. By ignoring these prices, a bias, known as sampling bias, is being introduced. Using old low-priced varieties and ignoring new high-priced ones has the effect of biasing the index downward. For some products, the new variety may be introduced at a relatively low price though the old one may continue at a relatively high price, serving a minority segment of the market. 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 varieties. Some strategies for the introduction of new models, and implications for CPI measurement, are considered in Annex 6.3.

6.185   This sampling bias exists for most products. However, the concern here is with product markets where the NSOs are finding the frequency of new variety introductions and old variety obsolescence sufficiently high that they may have little confidence in their results. Three procedures will be considered: an extensive use of the matched-model (overlap) technique, class mean imputation, and the use of hedonic price indices (as opposed to the partial, hedonic patching discussed in paragraphs 6.136–6.150).

Matching and the Overlap Method for Markets with Rapid Turnover of Models

6.186   The first approach to address markets with rapid turnover of models is simply a more extensive use of the overlap approach outlined previously for permanently missing prices. In this case, it is adopted for permanently missing varieties that occur frequently, as is usual for changes in models of electronic goods and automobiles. Matching prices of a few representative varieties becomes a less feasible approach in this context. A backward imputation is illustrated in Table 6.11 though, as outlined in the illustration for Table 6.5C, a forward imputation can be equally justified and both methods provide the same results when the imputations are based on the price changes of overlapping matched samples.

6.187   Considering model 1 in Table 6.11, in November there is no overlap price for the new model 1R, so its price is imputed "backward" by using the ratio of geometric means of the December to November prices but only including those for which matched models exist, that is, models 2R, 4R, and 5. These are all constant-quality price comparisons; of like with like. As shown in expression 6.13, its imputed price is $0.974 \times 30 = 29.2$:

$$\frac{(37 \times 30 \times 29)^{\frac{1}{3}}}{(40 \times 30 \times 29)^{\frac{1}{3}}} = \left(\frac{37}{40} \times \frac{30}{30} \times \frac{29}{29}\right)^{\frac{1}{3}} = 0.974$$

6.188   The imputed price for the replacement model 2R in June is based on the price changes of matched models 1, 3, and 5 for June and July, that is, $\dfrac{(25 \times 30 \times 29)^{\frac{1}{3}}}{(25 \times 30 \times 29)^{\frac{1}{3}}} = 1.00$

and its imputed price: $1.00 \times 30 = 30$. The imputed prices for

3R in November and 4R in June are 32.15 (rounded to 32.2, for simplicity of exposition) and 30.0, respectively.

6.189   The overall price relatives for each model, and their linked-in replacements, can now be computed as the product of short-term month-on-month price changes, that is, model 1 and its replacement 1R, for January to January, using the overlap month of November:

$$\frac{p_1^{\text{Feb}}}{p_1^{\text{Jan}}} x \frac{p_1^{\text{Mar}}}{p_1^{\text{Feb}}} x \frac{p_1^{\text{Apr}}}{p_1^{\text{Mar}}} \frac{p_1^{\text{May}}}{p_1^{\text{Apr}}} x \frac{p_1^{\text{Jun}}}{p_1^{\text{May}}} x \frac{p_1^{\text{Jul}}}{p_1^{\text{Jun}}}$$

$$x \frac{p_1^{\text{Aug}}}{p_1^{\text{Jul}}} x \frac{p_1^{\text{Sep}}}{p_1^{\text{Aug}}} x \frac{p_1^{\text{Oct}}}{p_1^{\text{Sep}}} x \frac{p_1^{\text{Nov}}}{p_1^{\text{Oct}}} x \frac{p_1^{\text{Dec}}}{p_{1R_{\text{imp}}}^{\text{Nov}}} x \frac{p_{1R}^{\text{Jan}}}{p_{1R}^{\text{Dec}}}$$

$$= \frac{p_1^{\text{Nov}}}{p_1^{\text{Jan}}} x \frac{p_{1R}^{\text{Dec}}}{p_{1R_{\text{imp}}}^{\text{Nov}}} x \frac{p_{1R}^{\text{Jan}}}{p_{1R}^{\text{Dec}}} \qquad (6.16a)$$

$$\frac{25}{25} \times \frac{25}{25} \times \frac{25}{25} \times \frac{25}{25} \times \frac{25}{25} \times \frac{25}{25} \times \frac{25}{25}$$

$$\times \frac{25}{25} \times \frac{25}{25} \times \frac{20}{25} \times \frac{30}{29.2} \times \frac{30}{30}$$

$$= \frac{20}{25} \times \frac{30}{29.2} \times \frac{30}{30} = 0.8219 \qquad (6.16b)$$

Table 6.11   Illustration of Rapid Model Turnover

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Jan.
1	25	25	25	25	25	25	25	25	25	25	20		
1R											**29.2**	30	30
2	35	35	35	35	35	32							
2R							**37**	37	37	37	37	40	40
3	30	30	30	30	30	30	30	30	30	30	27		
3R											**32.2**	33	33
4	29	29	29	29	29	28							
4R							**30**	30	30	30	30	30	30
5	29	29	29	29	29	29	29	29	29	29	29	29	29

Values in bold are imputed.

6.190   The price relative for model 1 from January in the preceding year to January in the current year shows a $(1 - 0.8219) \times 100 = 17.81$ percent price decrease. It is clear from Table 6.11 that the price for model 1 has been constant up to October, and there was a price fall in November, but this was to clear the market for the replacement. There should be a November to December price increase for the old model 1 to the new replacement 1R that reflects that part of the price difference between model 1 and model 1R was not due to quality differences. But the imputation is based on the constant price movements of models 4R and 5 and a coincidental price increase in model 2R; it assumes that short-term price movements of matched pairs will proxy the price change of model 1. However, in this context, the constant price changes of matched models are an inappropriate proxy badly biasing the measured price change downward. At fault are first, the use of the unrepresentative price for model 1 at the end of its life cycle in November, and second, the inappropriate imputation for the replacement variety.

6.191   The January to January price decreases for models 2, 3, and 4, using replacements, and 5 are, respectively, 1.2, 6.3, and 3.4 percent and no change for model 5. With a substantial churn of models, possibly more frequently than annual, the bias from using overlaps can be substantial.

6.192   As an advantage, the method is simply an extension of the linking-in of new products and can be readily applied by an NSO, especially one with limited resources. However, this approach biases the CPI because it bases the imputations on price changes of matched varieties not subject to the price changes that occur on the replacement of a model.

6.193   The overlap method may be subject to bias if applied to where there is substantial churn in the product market and an active policy by the supplier of introducing upgraded replacement models. The nature and extent of the bias depends on the pricing strategy. Table 6.11 illustrates a policy of lower pricing at the end of the life cycle and a higher price at the start. Importantly, the example has no price change for other matched models, from which the imputation was drawn, and thus a biased imputation at this critical overlap period. The bias is substantial and downward. Alternative pricing strategies are given in Annex 6.3 along with their implication for bias from using the overlap matching. It has been demonstrated in empirical studies that the method can introduce substantial bias under quite reasonable conditions. The nature and extent of the bias depend on business pricing strategies that may change over time and are unpredictable, which is a concern for CPI compilation in this important product area. The overlap method is thus

149

not recommended for product markets with a high rate of model churn.

Use of a Class Mean Imputation

6.194   The previous example showed that an imputation based on price movements of other matched models not at the end of their life cycle could introduce a bias. An alternative, though more resource-intensive method, is to base the imputations not on price changes of matched varieties but to use, where possible, explicit quality adjustments for linked-in noncomparable replacement. For example, internet webpages of prices of similar products may show the difference in characteristics and prices of the old models and the replacement models. The replacement may simply have a higher value of some performance characteristic or feature for which the price is available as an option. If enough explicit quality adjustments can be made, imputations might be better made on the basis of only those models that have had explicit quality adjustments to their price.

6.195   The nature of the high frequency of replacements makes the procedure resource-intensive and, in some instances, not viable because of the absence of explicit information on prices of features or options. However, if enough models have an explicit adjustment, an average of their price change could be used to impute the price change of other models being replaced. This is the basis of using class mean imputations. The method requires care that the linking-in of replacement models does not take place at the end of the model's life cycle when pricing might be abnormally low for a variety that relatively few consumers are purchasing. This not only has a detrimental effect on the quality-adjustment methodology, but also on the representativity of the models upon which the prices change measurement is based.

6.196   The class mean imputation method was outlined in paragraphs 6.96–6.100. It is similar in procedure to the overall mean imputation and is a form of targeted imputation. The "target" is measured price changes of replacements for permanently missing products. Only the price changes of "comparable" replacements are used to impute the overlap price, the replacements being limited to those that have exactly the same price-determining characteristics, or those varieties with replacements that have been declared comparable after review and have already been quality-adjusted through one of the "explicit" methods. For example, when the arrival of a new model of a particular kind of automobile forces price collectors to find replacements, some of the replacements will be of comparable quality, while others can be made comparable with explicit quality adjustments, but the remaining ones will need imputed prices for an overlap month

Hedonic Price Indices

6.197   It is important to distinguish between the use of hedonic regressions for patching and their use as *hedonic price indices,* which are measures of quality-adjusted price changes. Patching adjusts individual item prices for quality differences when a noncomparable replacement is used while hedonic price indices are measures of quality-adjusted price changes. Hedonic price indices are suitable when the pace and scale of replacements of varieties are substantial because, first, an extensive use of these overlap quality

adjustments may lead to bias and, second, the sampling will be from a static 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 in the hedonic regression. Note that all the indices described in the following text use a fresh sample of the data available in each period. If there is a new variety in a period, it is included in the data set and its quality differences controlled for by the regression. Similarly, if old varieties drop out, they are still included in the data for the indices in the periods in which they exist. Paragraphs 6.110–6.115 stress the need for caution in the use of hedonic regressions for quality adjustments.

6.198   Consider a price comparison between two adjacent time periods, periods t and $t + 1$. The models sampled do not have to be matched, they may simply be all recorded models sold in the two periods, and they comprise a different mix of qualities. The hedonic formulation regresses the price of model i, p_i, on the $k = 2, \dots, K$ characteristics of the varieties 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+1} being 1 in period $t + 1$, zero otherwise.

The Time Dummy Variable Approach

6.199   A single hedonic regression equation is estimated with observations across models over adjacent time periods, including the reference period 0 and a subsequent period t. The logarithm of prices of individual models is regressed on their characteristics and a dummy variable for time, taking the values of $D_i^1 = 1$ if the model is sold in period 1 and 0 otherwise. A log-linear specification is given by

$$\ln p_i^t = \ln \beta_0^t + \sum_{k=1}^{K} z_{k,i}^t \ln \beta_k^t + \ln \varepsilon_i^t \qquad (6.17)$$

6.200   The δ^t are estimates of the *proportionate* change in price arising from a change between the excluded reference period $t = 0$ and successive periods $t = 1, T$ having controlled for changes in the quality characteristics via the term $\sum_{k=1}^{K} Z_{k,i}^{0,t} x \widehat{ln \beta_k}$.

6.201   In principle, the index $100 \times \exp\left(\hat{\delta}^t\right)$ requires an adjustment for it to be a consistent (and almost unbiased) approximation of the proportionate impact of the time dummy variable.[11] In practice, the adjustment usually has little effect.

6.202   The method implicitly restricts the coefficients on the quality characteristics to be constant over time: for example, for an adjacent period January and February regression, for $k = 1, \dots, K$ characteristics and where period 0 and t are January and February, respectively, $\beta_k = \beta_k^{\mathrm{Jan}} = \beta_k^{\mathrm{Feb}}$. The (relative) valuation of a characteristic, for example, for a washing machine with an additional 100 revolutions per minute spin speed, is the same in January as in February.

----

[11] See Triplett (2006) for more details.

The index, $100 \times \exp\left(\hat{\delta}^{\text{Feb}}\right)$, is an estimate of the quality-adjusted price change for February (January = 100).

The Characteristics/Repricing Approach

6.203 A hedonic regression is run to determine the price-determining characteristics of models, for example in a reference period 0. The average model in period 0 can then be defined as a tied bundle of the averages of each price-determining characteristic. In the previous example for washing machines, these would include: spin speed: 1,375 revolutions per minute; capacity (cotton load): 8.5 kilogram; annual energy cost: £36.5; steam facility: 4 percent; Brand A: 15 percent; warranty period: 5.4 years; and run-time (cotton): 18.8 minutes. These are the $\overline{Z}_k$ averages for each of the k price-determining characteristics.

6.204 The average values of each characteristic are held constant in each period but valued in turn using period 0 and period t hedonic regressions. One form of the (average) *characteristics* approach is as a measure of the price change of a set of average period 0 characteristics valued first, at period t hedonic valuations, and second, at period 0 hedonic valuations. A ratio of the results is a constant (period 0 characteristics) quality price index. The numerator, the period t hedonic valuation, provides an answer to a counterfactual question: what would be the estimated transaction price of a model with period 0 average characteristics, were it on the market in period t?

6.205 A constant-quality *hedonic geometric mean characteristics* price index from a log-linear hedonic regression equation is a ratio of geometric means with average characteristics held constant in the *reference* period 0, $\overline{z}_k^0$:

$$I_{\text{HGMC}:\overline{z}^0}^{0:t} = \frac{\prod_{k=0}^{K}\left(\hat{\beta}_k^t\right)^{\overline{z}_k^0}}{\prod_{k=0}^{K}\left(\hat{\beta}_k^0\right)^{\overline{z}_k^0}} = \frac{\exp\left(\sum_{k=0}^{K}\overline{z}_k^0 \ln \hat{\beta}_k^t\right)}{\exp\left(\sum_{k=0}^{K}\overline{z}_k^0 \ln \hat{\beta}_k^0\right)} \quad (6.18)$$

where $\overline{z}_k^0 = \frac{1}{N^0}\sum_{i\in N^0} z_{i,k}^0$.

6.206 Equation 6.18 holds the (quality) characteristics constant in period 0, though a similar index could be equally justified by valuing in each period a constant period t average quality set:

$$....I_{\text{HGMC}:\overline{z}^t}^{0:t} = \frac{\prod_{k=0}^{K}\left(\hat{\beta}_k^t\right)^{\overline{z}_k^t}}{\prod_{k=0}^{K}\left(\hat{\beta}_k^0\right)^{\overline{z}_k^t}} = \frac{\exp\left(\sum_{k=0}^{K}\overline{z}_k^t \ln \hat{\beta}_k^t\right)}{\exp\left(\sum_{k=0}^{K}\overline{z}_k^t \ln \hat{\beta}_k^0\right)} \quad (6.19)$$

where $\overline{z}_k^t = \frac{1}{N^t}\sum_{i\in N^t} z_{i,k}^t$.

$N^0 = N^t$ are the number of matched observations (varieties) in the sample. Neither a period 0 constant-characteristics index nor a period t constant-characteristic quantity basket can be considered to be superior, both acting as bounds for their theoretical counterparts. Some average or compromise

solution is required. An index making symmetric use of period 0 and period t characteristics values is intuitive:

$$I_{\text{HGMC}:\overline{z}^\tau}^{0:t} = \frac{\prod_{k=0}^{K}\left(\hat{\beta}_k^t\right)^{\overline{z}_k^\tau}}{\prod_{k=0}^{K}\left(\hat{\beta}_k^0\right)^{\overline{z}_k^\tau}} = \frac{\exp\left(\sum_{k=0}^{K}\overline{z}_k^\tau \ln \hat{\beta}_k^t\right)}{\exp\left(\sum_{k=0}^{K}\overline{z}_k^\tau \ln \hat{\beta}_k^0\right)} \quad (6.20)$$

where $\overline{z}_k^\tau = \left(\overline{z}_k^0 + \overline{z}_k^t\right)/2$.

6.207 Equations 6.18, 6.19, and 6.20 all use predicted prices in both the denominator and numerator. This follows the recommendation in the following text to use dual imputations. However, this method also entails running hedonic regressions in each period. Yet a fortuitous result is that a feature of the OLS estimator is the mean of actual prices being equal to the mean of predicted prices:

$$\frac{1}{N^0}\sum_{i\in N^0}\ln \hat{p}_{i|z_i^0}^0 = \frac{1}{N^0}\sum_{i\in N^0}\ln p_i^0 \text{ and } \frac{1}{N^t}\sum_{i\in N^t}\ln \hat{p}_{i|z_i^t}^t = \frac{1}{N^t}\sum_{i\in N^t}\ln p_i^t$$

Thus, while the numerator of equation 6.18 and denominators of equation 6.19 must be counterfactual—the valuing period 0 (t) average characteristics at period t (0) prices—the denominator of equation 6.18 and numerator of 6.19 can use actual prices, since the means are the same for an OLS estimator. This leads to the important results that equation 6.19 does not require a hedonic regression to be estimated in every current period t, only in the price reference period 0. This is an important result since it aids the practical work of compilers who do not have to estimate a hedonic regression equation in each period, but maybe once every one or two years, depending on the amount of churn in the market and shifting technologies and preferences. The hedonic indices from one regression can be chained to its preceding hedonic indices, and so forth, using successive multiplication.

The Hedonic Imputation Approach

6.208 In contrast to the characteristics approach, the *imputation* approach works at the level of individual varieties/models, rather than the average values of their characteristics. The rationale for the imputation approach lies in the MMM. Consider a set of models transacted in period 0. The objective is to compare their period 0 prices with the prices of the same matched models in period t. In this way, there is no contamination of the measure of price change by changes in the quality mix of models transacted. However, for goods and services with a high model turnover, not all the period 0 models were sold in period t—there is no corresponding period t price in many cases. The solution—in the numerator of equation 6.21—is to predict the period t price of each i period 0 model, $p_{i|z_i^0}^t$.

6.209 A constant-quality *hedonic geometric mean imputation* price index from a log-linear hedonic regression equation is a ratio of geometric means with characteristics held constant in the *reference* period 0, $\overline{z}_k^0$:

$$I_{\text{HGMI}:z_i^0}^{0:t} = \frac{\prod_{i\in N^0}\left(\hat{p}_{i|z_i^0}^t\right)^{\frac{1}{N^0}}}{\prod_{i\in N^0}\left(\hat{p}_{i|z_i^0}^0\right)^{\frac{1}{N^0}}} = \frac{\exp\left(\frac{1}{N^0}\sum_{i\in N^0}\ln \hat{p}_{i|z_i^0}^t\right)}{\exp\left(\frac{1}{N^0}\sum_{i\in N^0}\ln \hat{p}_{i|z_i^0}^0\right)} \quad (6.21)$$

151

6.210   Alternatively, the value in the numerator of equation 6.21 is the geometric mean of the period t price for the price-determining characteristics in that period, $Z_{i,k}^t$. This is compared, in the denominator, with the geometric mean of the period 0 predicted price of the same period t price-determining characteristics, $Z_{i,k}^t$. For each model, the quantities of characteristics are held constant in period t, $Z_{i,k}^t$; only the characteristic prices change:

$$I_{HGMI:z_i^t}^{0:t} = \frac{\prod_{i \in N^t} \left(\hat{p}_{i|z_i^t}^t\right) \frac{1}{N^t}}{\prod_{i \in N^t} \left(\hat{p}_{i|z_i^t}^0\right) \frac{1}{N^t}} = \frac{\exp\left(\frac{1}{N^t} \sum_{i \in N^t} \ln \hat{p}_{i|z_i^t}^t\right)}{\exp\left(\frac{1}{N^t} \sum_{i \in N^t} \ln \hat{p}_{i|z_i^t}^0\right)} \quad (6.22)$$

6.211   As with the characteristics approach, a compromise solution of whether period 0 or period t constant characteristics should be used, is to apply an average of the two. However, as with the characteristics approach, equation 6.22 has the advantage of only requiring a single hedonic regression to be estimated in the price reference period 0. If this is used, the regression should be reestimated every year or so, the frequency being determined by the turnover of products.

6.212   The three approaches have different, yet valid, intuitions. As long as the functional form of the aggregator is aligned to the hedonic regression in the manner shown in Table 6.12, the imputation and characteristics approaches yield the same result. This consolidation not only markedly narrows down the choice between approaches, but also validates the measure as one resulting from quite different intuitions.

6.213   For a log-linear functional form of a hedonic regression, the requirements are that (1) for the characteristics approach, $\bar{z}_k^0$ and $\bar{z}_k^t$ are arithmetic means of characteristic's values, the right-hand side of the hedonic regression, and (2) for the imputation approach, the ratio of average predicted prices is a ratio of geometric means, the left-hand side of the regression.

6.214   The important feature of the hedonic indices is that they require no matching of individual models in the periods compared. Matching is required so that the price of a model in period 0 can be compared with that in period t, without a concern that the price change is affected by changes in quality. Such matching restricts the sample and, importantly in this context of a high level of churn in models and where prices change when models change, can lead to bias. This was illustrated in Table 6.11. The price comparison of matched models effectively removes from the sample price changes in the important period of a price comparison when models change. The imputation for November to December for model 1 in Table 6.11 is based on matched

prices only. Hedonic indices adjust for quality change not by any meticulous or time-consuming matching and, for that matter, identification of replacements, but by applying a hedonic regression to value constant-quality characteristics.

6.215   Hedonic indices use data on matched and unmatched observations and, again importantly, can naturally be applied to large monthly data sets, such as scanner and web-scraped data, as opposed to a small sample of what may have been in some long-past reference period, a representative variety.

6.216   An advantage of the imputation approach over the dummy variable approach is that explicit weighting systems can be more readily, accurately, and intuitively applied at this elementary level. For example, equation 6.22 may be defined for models i over a set of models of television sets sold in period t. The formula gives equal weight to each model sold. A major improvement would be to apply to each model's quality-adjusted price change the weight of that price change, that is, the individual model's share of transaction expenditure values, for example from scanner data. Silver (2018) outlines the methodology for the imputation approach, in the context of house price indices, to include quasi-superlative and superlative formulations. The weighted imputation approach also has a correspondence to a weighted characteristics approach, and the more intuitive application of weights, if formulated as in Table 6.12.

6.217   Hedonic indices are particularly well suited for large data sets, as web-scraped or scanner data (see Chapter 10), for which there is no matching of varieties. It is at initiation that a price collector selects a representative variety and matches its characteristics in subsequent periods in order to track the price of this same variety. In doing so, the sample of prices collected is highly restricted to what may be a single price. With hedonic indices, it is the varying values of the characteristics of the models that enable a constant-quality price change. There may be data sets in which accurately matched sampled prices form part of the sampled data. In such a case there would be no need for predicted prices to be used for constant-quality price change. The overall measure for this data set would contain: (1) actual price changes for the matched sample; (2) hedonic price changes for the period 0 models not sold in period t (as, for the hedonic imputation approach, in equation 6.21); and (3) hedonic price changes for the period t models not sold in period 0 (as, for the hedonic imputation approach, in equation 6.22). Each of these terms would be weighted by their relative expenditure shares, if available. It is from the measure of all three components that the difference between the MMM and hedonic indices becomes apparent.

The Difference between Hedonic Indices and Matched Indices

6.218   As already mentioned, an advantage of hedonic indices over matched comparisons was the inclusion by the former of unmatched data. Consider a data set of prices and characteristics over two successive time periods, periods 0 and t. Assume there are m matched models in both periods 0 and t, o old models in period 0, but disappearing thereafter, and n new models appearing in period t, and subsequently, as shown in Table 6.13.

Table 6.12   Equivalences of Hedonic Approaches

Hedonic Regression: Functional Form	Characteristics Approach: Form of Average of Characteristics	Imputation Approach: Form of Average of Predicted Prices
Linear	Arithmetic Mean	Arithmetic Mean
Log-Linear	Arithmetic Mean	Geometric Mean
Log–log	Geometric Mean	Geometric Mean

152

Table 6.13 Difference between Hedonic and Matched Indices

	Period 0	Period t
Matched Models (*m*)	m	m
Old Model (*o*)	o	
New Model (*n*)		n

6.219 A constant-quality, period 0 to t, price index, from a hedonic imputation approach, is made up of three terms:

- *The change in the geometric mean price of the* m *matched models*, with no need for quality adjustment because they are matched.

- *The change in the constant-quality geometric mean price of the old models with actual prices in period 0 and counterfactual ones in period* t. The counterfactual constant-quality price in period t has to be estimated since there is only a price in period 0. A prediction is required of what each old model's price in period 0 would have been had it been sold in period t. A period t hedonic regression is estimated, and a predicted price estimated for each model by inserting its period 0 characteristic Z_k^0 values into the right-hand side of the estimated regression equation. A geometric mean is compiled of these predicted values, $\prod_{i \in o} \left(\hat{p}^t_{i|z_i^0} \right)^{\frac{1}{No}}$, and compared with the period 0 geometric mean, $\prod_{i \in o} \left(p_t^0 \right)^{\frac{1}{No}}$, as in equation 6.19.

- *The change in the constant-quality geometric mean price of the new model in period* t. The counterfactual constant-quality price in period 0 has to be estimated since there is only a price in period t. A prediction is required of what each new model's price in period t would have been had it been sold in period 0. A period 0 hedonic regression is estimated, and a predicted price estimated for each model by inserting its period t characteristics Z_k^t values into the right-hand side of the estimated regression. A geometric mean is compiled of these predicted values, $\prod_{i \in n} \left(\hat{p}^0_{i|z_i^t} \right)^{\frac{1}{Nn}}$, and compared with the period t geometric mean, $\prod_{i \in n} \left(p_i^t \right)^{\frac{1}{No}}$, as in equation 6.20.

6.220 The overall index can be phrased as a weighted average of these three elements with the matched comparison having a weight of $2Nm/(2Nm + No + Nn)$, the old of $No/(2Nm + No + Nn)$, and the new, $Nn/(2Nm + No + Nn)$, though preferably the weights should be expenditure shares rather than the numbers of each model.

6.221 The MMM effectively ignores the last two elements of the bullet points in paragraph 6.222. This procedure would result in no bias if the imputed quality-adjusted price change of new and old varieties were the same as that for matched models. The MMM might be appropriate if the number of new and old models—or their expenditure weights—is small relative to matched models. This would be the case for the hedonic patching of permanently missing model prices outlined previously, but not for this context where there is a high and frequent turnover in models.

6.222 Even if the MMM is used with replacements, something of the dynamic universe of models is brought into the measure, but only insofar as there is one-on-one variety replacement. Furthermore, hedonic indices employ a consistent basis for the explicit quality adjustment for non-comparable replacements.

6.223 The deficiency of the MMM against a hedonic index has been shown previously with regard to the hedonic imputation approach, though similar considerations apply to a time dummy variable approach. Consider an adjacent period time dummy variable hedonic index of the form of equation 6.17, with the index change captured by the coefficient on the dummy variable for time. For example, a sample of models of washing machines for periods t and $t + 1$ would have in the regression the log of price on the left and price-determining characteristics on the right-hand side. On the right-hand side, a dummy variable would also denote whether the observation is drawn from period t or $t + 1$. The hedonic regression includes matched, new, and old models and the quality adjustment is achieved through the term $\sum_{k=1}^{K} z_{k,i}^t \ln \beta_k^t$ in equation 6.17. A matched-model measure of price change would again only measure the price change for the more limited sample of matched models but would not require a quality adjustment. The hedonic dummy variable approach, with its inclusion of unmatched old and new observations, will likely 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 varieties leaving and entering the sample and on the price changes of old and new varieties 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. Different market behavior will lead to different forms of bias (see Annex 6.3).

Key Recommendations

- All temporarily missing prices should be imputed using one of the imputation methods described in the chapter. Methods include overall mean and target mean imputations.

- The imputation of temporarily missing prices is especially important when using the short-term formulas—modified Lowe and modified Young. Imputations, which are self-correcting, avoid introducing any bias into the index.

- Imputations can be made either forward or backward. The results are equivalent, and the countries can choose which is most appropriate.

- The carryforward method should not be used, except for fixed or controlled prices. This method introduces a downward bias into the index.

- NSOs should define a period during which nonseasonal products can be considered temporarily missing. While this threshold varies from country to country, the most commonly used threshold is three months but can be longer.

- Permanently missing prices require a replacement variety.

- Quality change refers to changes in the price-determining characteristics when one variety replaces another. If these differences are judged to be comparable (that is, they are deemed to be similar), the price of the old and the new variety can be compared directly and any difference in price is reflected as price change. Should the differences be such that the old and the new variety are deemed to be noncomparable, a quality adjustment is needed. Quality adjustments ensure that the index reflects only pure price change and not changes because of differences in quality.

- Explicit or direct quality adjustments are preferred. They include quantity adjustment resulting from changes in size or quantity, changes in option costs, differences in production costs, and hedonics. Quantity adjustments are straightforward, and many countries apply this method for changes in size. The other explicit methods require data and experience making explicit quality adjustments.

- Implicit or indirect quality adjustments are the second-best approach; however, they could be preferred given a lack of data and expertise required for the explicit methods. Implicit methods include overlap pricing and imputation.

- The rapid turnover in the models sold of select products (for example, televisions, computers, telecommunications equipment, or appliances) requires a different strategy. Over time, these products usually have a rapid turnover in their quality characteristics. While the MMM, class mean imputations, and hedonic price indices may be used, the chapter notes that the MMM may lead to significant bias.

154

Annex 6.1

Overall Mean (or Targeted) Imputation

Consider $i = 1 \ldots m$ varieties in period t and p_i^t as the price of variety i in period t. All varieties continue into period $t + 1$ except for the single variety m which is replaced by variety n. p_n^{t+1} is the price of a replacement variety n in period $t + 1$. Now n replaces m but is of a different quality. There are $(m - 1)$ matched prices between periods t and $t + 1$ and a single replacement price. Let $A(z)$ be the quality adjustment to p_n^{t+1} which equates its quality services or utility to p_m^{t+1}, had it existed, 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 change of the $i = 1 \ldots m$ varieties, including the quality-adjusted price p_m^{*t+1}, given on the left-hand side of equation A6.1.1, must equal the average price change from just using the overall mean of the rest of the $i = 1 \ldots m - 1$ varieties, on the right-hand side of equation A6.1.1. 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 given is based on Triplett (2006), though a similar geometric one can be readily formulated. The equation for one unavailable variety is given by

$$\frac{1}{m}\left[\frac{p_m^{*t+1}}{p_m^t} + \sum_{i=1}^{m-1}\frac{p_i^{t+1}}{p_i^t}\right] = \left[\frac{1}{(m-1)}\sum_{i=1}^{m-1}\frac{p_i^{t+1}}{p_i^t}\right] + Q \quad \text{(A6.1.1)}$$

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

and for x unavailable varieties by

$$Q = \frac{1}{m}\sum_{i=m-x+1}^{m}\frac{p_m^{*t+1}}{p_m^t} - \frac{x}{m(m-x)}\sum_{i=1}^{m-x}\frac{p_i^{t+1}}{p_i^t} \quad \text{(A6.1.2)}$$

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

$$r_1\left[\sum_{i=p}^{m-x} p_i^{t+1} / p_i^t\right] \div (m-x) \quad \text{and}$$

$$r_2 = \left[\sum_{i=m-x+1}^{m} p_i^{*t+1} / p_i^t\right] \div x \quad \text{(A6.1.3)}$$

then the bias of the arithmetic mean of ratios from substituting equation A6.1.3 in equation A6.1.2 is

$$Q = \frac{x}{m}(r_2 - r_1) \quad \text{(A6.1.4)}$$

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 varieties 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 varieties 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 $m = 3$ varieties, each with a price of 100 in period t. Let the $t + 1$ prices be 120 for two varieties, but assume the third, that is, $x = 1$, is unavailable and is replaced by a variety with a price of 140, of which 20 is attributable to quality differences. Then the arithmetic bias as given in equations A6.1.3 and A6.1.4, where $x = 1$ and $m = 3$, is

$$\frac{1}{3}\left[(-20+140)/100 - \left(\frac{120}{100} + \frac{120}{100}\right)/2 = 0\right] \quad \text{(A6.1.5)}$$

If 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 the 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.

The analysis is framed with regard to a short-term price change framework. That is, the short-term 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-term imputation where a base price is compared with prices in subsequent months, and where the implicit assumptions are more restrictive.

Table A6.1 provides an illustration in which the (mean) price change of varieties that continue to exist, r_1, can vary for values between 1.0 and 1.5, corresponding to a variation between no price change and a 50 percent increase. The (mean) price change of the quality-adjusted new varieties compared with the varieties they are replacing is assumed not to change (that is, $r_2 = 1.00$). The bias is given for ratios of missing values of 0.01, 0.05, 0.1, 0.25, and 0.5, for both arithmetic means and geometric means. For example, if 50 percent of price quotes are missing and the missing quality-adjusted prices do not change, but the prices of existing varieties increase by 5 percent ($r_1 = 1.05$), then the bias for the geometric mean is represented by the proportional factor 0.9759 (that is, instead of 1.05), the index would be $0.9759 \times 1.05 = 1.0247$. For an arithmetic mean, the bias is -0.025; instead of 1.05, it should be 1.025.

Equation A6.1.4 shows that the ratio x/m and the difference between r_1 and r_2 determine the bias. Table A6.1 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 percent for existing varieties translates to an index change of 3.73 and 3.75 percent for the geometric and arithmetic formulations, respectively, when $r_2 = 1.00$ (that is, when quality-adjusted prices of unavailable varieties are constant). Instead of being 1.0373 or 1.0375, ignoring the unavailable varieties would give a result of 1.05. Even with 10 percent missing ($x/m = 0.1$), an inflation rate of 5 percent for existing varieties translates to 4.45 and 4.5 percent for the geometric and arithmetic formulations, respectively, when $r_2 = 1.00$.

155

Table A6.1   Example—Bias from Implicit Quality Adjustment When the (Mean) Price Change of Quality-Adjusted New Varieties Compared with the Varieties They Are Replacing Is Assumed Not to Change ($r_2 = 1.00$)

r_1	Geometric Mean					Arithmetic Mean				
	Ratio of Missing Varieties, x/m					Ratio of Missing Varieties, x/m				
	0.01	0.05	0.10	0.25	0.50	0.01	0.05	0.10	0.25	0.50
1	1.00000	1.00000	1.00000	1.00000	1.00000	0.00000	0.00000	0.00000	0.00000	0.00000
1.01	0.99990	0.99950	0.99901	0.99752	0.99504	−0.00010	−0.00050	−0.00100	−0.00250	−0.00500
1.02	0.99980	0.99901	0.99802	0.99506	0.99015	−0.00020	−0.00100	−0.00200	−0.00500	−0.01000
1.03	0.99970	0.99852	0.99705	0.99264	0.98533	−0.00030	−0.00150	−0.00300	−0.00750	−0.01500
1.04	0.99961	0.99804	0.99609	0.99024	0.98058	−0.00040	−0.00200	−0.00400	−0.01000	−0.02000
1.05	0.99951	0.99756	0.99513	0.98788	0.97590	−0.00050	−0.00250	−0.00500	−0.01250	−0.02500
1.1	0.99905	0.99525	0.99051	0.97645	0.95346	−0.00100	−0.00500	−0.01000	−0.02500	−0.05000
1.15	0.99860	0.99304	0.98612	0.96566	0.93251	−0.00150	−0.00750	−0.01500	−0.03750	−0.07500
1.2	0.99818	0.99093	0.98193	0.95544	0.91287	−0.00200	−0.01000	−0.02000	−0.05000	−0.10000
1.3	0.99738	0.98697	0.97411	0.93651	0.87706	−0.00300	−0.01500	−0.03000	−0.07500	−0.15000
1.5	0.99595	0.97993	0.96027	0.90360	0.81650	−0.00500	−0.02500	−0.05000	−0.12500	−0.25000

Considering a fairly low ratio of x/m, for example, 0.05, then even when $r_2 = 1.00$ and $r_1 = 1.20$, Table A6.1 shows that the corrected rates of inflation should be 18.9 percent and 19 percent 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 variety and the old variety 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 A6.1. The latter would have a lower mean, rendering comparisons of bias meaningless.

An awareness of the market conditions relating to the products concerned is instructive in understanding likely differences between r_1 and r_2. The concern is when prices vary over the life cycle of the varieties. For instance, at the introduction of a new model, the price change may be quite different from price changes of other existing varieties. Thus, assumptions of similar price changes, even with quality adjustment, might be inappropriate. For example, if 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. Or if new clothing enters the market at relatively high quality-adjusted prices, while old, end-of-season, or out-of-style clothes are being discounted, there will be bias, as r_1 differs from r_2.

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). In addition, consumer marketers are taught to plan the marketing mix for the life cycle of varieties. 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 behavior. Thus, the quality-adjusted price change of an old variety compared with a new replacement variety may be higher than price changes of other varieties in the product group. After the introduction of the new variety, its prices may fall relative to others in the group. There may be no law of one price change for differentiated varieties within a market.

There is thus little in economic or marketing theory to support any expectation of similar (quality-adjusted) price changes for new and replacement varieties, as compared to other varieties 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 A6.1 provides guidance here. The second is the expected difference between r_1 and r_2. It is clear from the previous 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 varieties expected to have similar price changes. The selection of such varieties, however, should take account of the need to include a sufficiently large sample so that the estimate is not subject to undue sampling error.

The manner in which these calculations are undertaken is also worth considering. In its simplest form, the pro forma setting for the calculations (for example, on a spreadsheet), would usually have each variety description and its prices recorded on a monthly basis. The imputed prices of the missing varieties 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."

The method described previously is an illustration of a short-term imputation. As discussed in Chapter 8, there is a strong case for using short-term imputations as against long-term ones.

156

Annex 6.2

Quality Adjustment Using a Replacement and Price Overlap

Consider p_m^t as the price of variety m in period t, p_n^{t+1}, the price of a replacement variety n in period $t + 1$; n replaces m but is of a different quality. Let there be overlap prices for m and n in period t and let $A(z^{t+1})$ be the quality adjustment to p_n^{t+1} which equates its quality to p_m^{t+1} such that the quality-adjusted price $p_m^{*t+1} = A(z^{t+1})p_n^{t+1}$. Now the quality adjustment to prices in period $t + 1$ is defined as $p_m^{*t+1} = A(z^{t+1})p_n^{t+1}$ which is the adjustment to p_n in period $t + 1$ that equates it 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:

$$\left(p_m^{*t+1} / p_m^{t+1} \right) \qquad \text{(A6.2.1)}$$

The overlap formulation equals this when:

$$\frac{p^{*t+1}}{p_m^{t-1}} = A\left(z^{t+1}\right)\frac{p_n^{t+1}}{p_m^{t-1}} = \frac{p_n^{t+1}}{p_n^t} \times \frac{p_m^t}{p_m^{t-1}} \qquad \text{(A6.2.2)}$$

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

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

But what if the assumption does not hold? What if the relative prices in period t, $P^t = p_m^t \big/ p_n^t$, do not equal $A(z^t)$ in some future period, say $A\left(z^{t+i}\right) = \alpha_i R^t$? 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 variety n is effectively being compared with itself:

$$\frac{p_m^{*t+4}}{p_m^{t-1}} \Big/ \frac{p_m^{*t+3}}{p_m^{t-1}} = \frac{\alpha R^t}{\alpha R^t}\frac{p_n^{t+4}}{p_n^{t+3}} = \frac{p_n^{t+4}}{p_n^{t+3}} \qquad \text{(A6.2.3)}$$

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

$$\frac{p_m^{*t+4}}{p_m^{t-1}} \Big/ \frac{p_m^{*t+3}}{p_m^{t-1}} = \frac{\alpha_4}{\alpha_3}\frac{p_n^{t+4}}{p_n^{t+3}} \qquad \text{(A6.2.4)}$$

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

Annex 6.3

The Nature and Extent of the Index Number Bias If Only Matched Varieties Are Used

Sample degradation and differences in the (quality-adjusted) prices of unmatched new, unmatched old, and matched models can lead to bias in matched-model price indices that take into account only models available in two consecutive periods. The nature and extent of such bias depends on the frequency with which manufacturers turn over their models and the pricing strategy retailers employ over the life cycle of the models. For example, assuming that matched prices go down, if the quality-adjusted prices of new unmatched models in period $t = 2$ are higher than their matched counterparts in period 2, and if the quality-adjusted prices of old unmatched models in period 1 are lower than their matched counterparts in period 1, then there will be a *larger* fall in the matched-model index between periods 1 and 2 compared with a hedonic index that uses all of the data. Similarly, if the quality-adjusted prices of unmatched new models are *below* matched ones in period 2, and the quality-adjusted prices of old models *above* matched ones in period 1, there will be a *smaller* fall in the matched-model index compared with a hedonic index that uses all of the data. The nature and extent of the matched models index bias thus depends on the pricing strategy adopted for new and old models. Indeed, if hedonic-adjusted prices are consistently above or below average prices for unmatched new *and* old models some of the bias will cancel.

The case for old unmatched models having below average quality-adjusted prices is based on an inventory-clearing argument. For an old model near or at the end of its life cycle retailers want to clear out the remaining inventory from both their warehouses and store shelves so they have room to stock and display the replacement model. They do not wish the old model to cannibalize some of the sales of the new model which may have a higher price (profit) margin. The extent of any such cannibalization will depend on the cross-price elasticities between the new and old models. This inventory-clearing argument is noted in quadrant IV of Figure A6.2.

New and old models may coexist for some time and existing models having their prices *increased* following the introduction of a new model is of interest. Such a pricing strategy can be used by a multiproduct firm that anticipates the introduction of a new model. A *multiproduct* monopolist can increase the prices of existing models because some of the demand for existing models that would usually be lost to competitors, because of the price increase, will now not go to the competitor's products but will go instead to the firm's new model. The new model will cannibalize some of the existing model's sales that would otherwise be lost because of the price increase in the existing model (Figure A6.2, quadrant II). For example, the price of old, branded pharmaceutical drugs can increase after the expiration of a patent and introduction of new generic models. This is due to price discrimination with some market segments remaining with particularly strong preferences for the old models willing to pay higher prices (Figure A6.2, quadrant II). However, it has been argued that any such effect may be outweighed by the need to cut the prices of the existing models to prevent the existing model's sales cannibalizing sales of the new, more profitable model (Figure A6.2, quadrant IV). A study of computer processors and disk drives found that with the introduction of products embodying new technologies the prices for older products decline rapidly to permit the older technology to compete with the newer one for a limited time, but the old technology is eventually driven out.

New models may have *above* average quality-adjusted prices in their period of introduction because firms "price skim" market segments willing to pay a premium for the new model over and above that due to its improved quality (Figure A6.2, quadrant I). Indeed, marketing texts advocate price skimming as one of two "new product" pricing strategies. The alternative strategy is "market-penetration" pricing for which a low initial price is set for a new model to attract a large number of buyers quickly to win market share and take advantage of falling costs because of scale economies. Such pricing may initially be possible because the new model is based on new, lower cost components that can provide a feature set that is comparable to existing models, but at a lower price point. In either event, quality-adjusted prices of new models may have *below* average prices (Figure A6.2, quadrant III).

Figure A6.2 summarizes these positions. The combination of above average prices for new models in quadrant

---

Figure A6.2   Matched-Model Price Index Bias and Pricing Strategies

	New Unmatched Models	Old Unmatched Models	Pricing	Matched-Model Price Index Bias
Quality-Adjusted Prices **above** Matched-Model Prices	Market Skimming	Multiproduct Monopoly Pricing Strategy; Price Discrimination to Segments with Sticky Downward Pricing; Old Technologies Reduce Prices	I and IV	Upward
	[quadrant I]	*[quadrant II]*	II and III	Downward
Quality-Adjusted Prices **below** Matched-Model Prices	Market-Penetration Pricing; Low Unit Costs, New Producing Technology	Inventory Clearing	I and II	Countervailing
	[quadrant III]	*[quadrant IV]*	III and IV	Countervailing

159

I and below average prices for old models in quadrant IV leads to an overall net downward bias of the matched-model price index. Similarly, pricing in quadrants II and III leads to matched-model indices that are biased upward. However, pricing in quadrants I and II leads to an indeterminate bias, with countervailing effects from the new and old above average priced models. The outcome from pricing in quadrants III and IV is also indeterminate, with countervailing effects from the new and old below average priced models.

It is possible to say something about the likely pricing strategies of different products. Consider the case of digital cameras compared to film-based cameras. Given the current differential in product costs for the two technologies, and where the two categories are in their respective life cycles, relatively greater effort is likely to be placed on R&D in new models that reduce unit costs for digital cameras compared to R&D in film-based camera models that reduce unit costs. Products in a mature stage of their category life cycle, where R&D development is relatively small and product enhancing, as opposed to cost reducing, may be more likely to have above average quality-adjusted prices for new models.