CHAPTER 7: TREATMENT OF QUALITY CHANGE

A. Introduction

A.1 Why quality change is an issue

1. When routinely compiling a CPI, specific varieties of commodities in the index regularly appear and disappear. New goods and services can appear because technical progress makes production of new varieties possible. Even without technical progress in the supplying activity, however, products previously feasible, but not produced, may emerge because the technology of the using activity or the tastes of the final consumer have shifted. Existing varieties often decrease in importance or disappear from the market altogether as new varieties appear. Moreover, the priced set of products often is a small sample of the full range of products that exist at any given time. Products in the sample may appear and outlets in the sample.

2. This chapter covers how to deal with the problem of continuous change in the assortment of transactions whose prices make up CPIs. The overarching principle for designing methods to deal with variety turnover is that, at the most detailed level, the prices of items between any two periods may be directly compared only if the items are *essentially the same*. Violating this principle would mean that a given monthly price ratio measures not only the change in price, but also the value of the qualitative difference between two items. This contaminates the estimate of relative price change with an element, quality, that measures relative volume rather than price. It degrades the accuracy of the price index formed with the price ratios or relatives for the specific transactions.

3. What does "essentially the same" mean in practical terms? For measurement purposes, a product equates to a *complete description* of price-determining characteristics. It may, for example, be the case that a good can be bought with a warranty or an extra option for the same price. They are not the same good and the price of the option may decrease over time while that of the warranty increase. They have to have the same product description. The form of this description often is simply text. It also can be highly structured, however. In *structured product descriptions*, the product's characteristics are specific levels of indicators for several dimensions that are known to affect the average transaction price.[1] Each set of these indicators' levels frames a specific product. Examples of these dimensions are the horsepower of an automobile, the speed of a computer, or the species of a piece of fruit. Examples of product-determining levels or specific settings of these respective dimensions are 325 horsepower, 2 gigahertz, or flame red grape. Another set of products for n automobiles, computers, or fruit would be described by the characteristics levels 110 horsepower, 3 gigahertz, or Thompson green grape.

[1] See Chapter 6 on structured product descriptions, also termed checklists by some statistical agencies.

4. For price measurement purposes, the comparative quality of a product comprises its description and price. Distinct descriptions represent different qualities of products, to the extent that they contain different levels of characteristics that affect the average price of transactions of things with that description in a given month. When comparing descriptions, the practice of price statistics thus judges quality by price. If products with two distinct descriptions are transacted at the same time, the description with the higher price must be the higher quality. This corresponds to what is called a higher revealed preference or value in use of the product (demand side), as well as higher content in the input needed to make the product (supply side). For index compilers, then, quality is an ordinal concept, comprising the set of complete product descriptions ordered by price for a given month.

5. When an existing variety of a product disappears and a new one appears, a new description manifests itself as well. The new description is different from the descriptions of existing products because the level of at least one characteristic in the description has changed. The difference in the characteristic explains the difference in price compared with varieties already available. For example, a new variety of computer emerges with a processor speed of 3 gigahertz instead of 2, and it has, say, a $325 premium over 2 gigahertz computers already available. Thus, the value of the additional gigahertz of speed is $325 and the new computer is, by implication, of higher quality than the old one.

6. When a variety of a of a product is no longer sold, there are a number of options open to the price statistician. As will be outlined below, these include the use of a comparable replacement; a non-comparable replacement, with an explicit quality adjustment to one of the prices to make them comparable, or an adjustment based on their relative prices in some overlap period; or more simply, the dropping of the variety from the sample until a new set of commodities are selected on rebasing or sample rotation. In each case a quality adjustment is being undertaken, be it implicit or explicit. It is sometimes said by statistical offices that they do not undertake quality adjustments in their CPI compilation. Some adjustments are always made; they are implicit in the practice of treating missing varieties. A number of empirical studies for CPIs and PPIs have found the choice of method for dealing with such missing values can matter substantially (Armknecht and Weyback, 1989; Dulberger, 1989; Lowe, 1995; and Moulton and Moses, 1997). This chapter also is a guide to selecting methods based on the measurement circumstances. The purpose of this chapter is to outline their nature so that the methods chosen can be the appropriate ones.

A.2 Why the matched-models method may fail

7. The matched-models approach to variety turnover described in Section A.1 is subject to three broad sources of error: (i) missing products, (ii) sample space change (sampling issues), and (iii) new products. *Missing products* are concerned with the solution of the problem of ‘what to do when a product is no longer produced or purchased’ by means of replacement varieties or imputations. In solving the problem the methods attempt to preserve the original matched sample with one-on-one replacements. For *sample space changes* the issue of concern is with incorporating the price changes of unmatched old and new models which are no longer available or introduced outside of the matched sample. However, when

a very *new product* arrives which cannot by definition be easily linked to existing ones, a different problem arises.

A.2.1 Missing products

8. For each outlet, compilers measure the long-run price change for a product by comparing the price of the product in the current period, usually month, with its corresponding price in the price reference period. This reference period would be the month when the products in the index entered the sample. When a product is missing, it may be because it has been discontinued, or it may not be available to the same specification—its quality has changed. We thus encounter the first potential source of error in the matched-models method. There are several specific contexts for this. It may be a seasonal product, or the product may be a custom-made good or service supplied each time to a customer's specification, or it may be that it is in short supply, or simply no longer demanded and supplied.. There are four main approaches for dealing with missing products:

- Approach 1: The price change of the discontinued product may be *imputed by the aggregate price change of a group of other products* whose price evolution compilers judge to be similar to that of the missing product. Such imputations should only be undertaken for short periods.

- Approach 2: A *replacement product may be selected, comparable in quality* to the missing product, and its price used directly to form a price relative.

- Approach 3: The replacement may be deemed non-comparable with the missing product, but prices of both the missing and replacement products may be available in an *overlap period* before the product was missing. Compilers use the price difference in this overlap period to quality-adjust the replacement product's price until there are at least two observations on the replacement product.

- Approach 4: The price of a non-comparable replacement may be used with an *explicit adjustment for the quality difference* to extract the pure price change.

9. This chapter discusses these four approaches to quality adjustment in some detail along with the assumptions they imply. Because the prices of the unavailable products are not measured by definition, the veracity of some of the maintained assumptions about their price changes, had they been available, is difficult to establish. Nevertheless, the objective of each of the methods is to produce matched comparisons of the prices of products: to compare like with like from month to month. When products are replaced with new ones of a different quality, then a quality-adjusted price is required to produce a match. 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.

A.2.2 Sampling issues

10. Sampling issues comprise four main areas of concern. First, samples lose relevance. A given set of matched models or products is likely to become increasingly unrepresentative of the population of transactions over time. It may be that the prices of old products being dropped are relatively low and the prices of new ones relatively high, and their prices are different, even after quality adjustment (Silver and Heravi, 2002). For strategic reasons, firms may wish to dump old models, to make way for the introduction of new models priced relatively high. Ignoring such unmatched old and new models in CPI measurement will bias the indexes downward (see Section G.2.3 in this chapter). Ironically, the matched-models method compilers employ to ensure constant quality may itself lead to bias, by ignoring such unmatched models, especially if used with an infrequently updated product sample. (See also Koskimäki and Vartia, 2001, for an example.)

11. Second, because of the additional resources required to make quality adjustments to prices, it may be in the interests of the price collectors, and indeed fall within their guidelines, to avoid making non-comparable replacements and quality adjustments. They keep with their products until they are no longer available—that is, continue to monitor old products with limited sales. Such products may exhibit unusual price changes as they near the end of their life cycle. These unusual price changes arise because marketing strategies typically identify gains to be made from different pricing strategies at different times in the life cycle of products, particularly at the introduction and end of the cycle (Parker, 1992). Yet their weight in the index, which is based on their sales share when they were sampled, would remain constant in the index and probably would be too high at the end of the life cycle. Further, new and, therefore, unmatched products with possibly large sales would be ignored. Undue weight would be given to the unusual price changes of matched products at the end of their life cycle. This issue again is resolved by more frequent sample reselection of products within a given sample of outlets.

12. Third, the methodology for selecting replacement products advises price collectors to choose comparable replacements to avoid the need for explicit quality adjustments to prices. Obsolete products are by their nature at the end of their cycle, and replacements, to be comparable, must also be near or at the end of their cycles. Obsolete products with unusual price changes at the end of their cycle may be replaced by other obsolete products with unusual price changes. This compounds the problem of unrepresentative samples and continues to bias the index against technically superior products delivering cheaper service flows.

13. Finally, the sampling problem with the matching procedure occurs when the price collector continues to report prices of products until replacements are forced, that is, until the products are no longer available, but has instructions to replace them with popular products. This improves the coverage and representativity of the sample. But the wide disparity between the characteristics of the old, obsolete products and new, popular ones makes accurate quality adjustment more difficult. The (quality-adjusted) price changes of very old and very new products may not be similar as required by the imputation methods under

approach 1. The differences in quality are likely beyond what can be attributed to price differences in some overlap period under approach 3, since one product is in the last stages of its life cycle and the other in its first. Further, the technical differences between the products 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 under approach 4. By implication, many of the methods of dealing with quality adjustment for unavailable products will work better if the switch to a replacement product is made sooner rather than later. Sampling issues thus are closely linked to quality adjustment methods. This will be taken up in Chapter 8, in the section on product selection and the need for an integrated approach to dealing with both representativity and quality-adjusted prices.

A.2.3 New products

14. The third potential source of error is distinguishing between new products and quality changes in old ones, also covered in Chapter 8. When a truly new product is introduced, there are at least two reasons why early sales are at high prices that later fall, often precipitously: capacity limitations and market imperfections. Both of these may be present shortly after introduction of a new product because there are few suppliers for it.

15. Early in the product life cycle, production processes may have limited capacity; therefore, producers find themselves operating at relatively high and increasing marginal costs of production. Marginal costs of operation tend to decline as more producers enter the market or as existing producers redesign and upgrade production facilities for higher volume. Both of these bring operating levels back from high marginal cost, near full capacity levels.

16. With or without early capacity constraints, the small number of suppliers early in the life cycle allows what economists call *market imperfections* to arise. In an imperfectly competitive market, the producer can charge a monopoly price higher than the marginal cost of production. A segment of the consumer market may put a premium on buying the latest model. As more competitors enter the market for the new good or service, the monopoly power of early sellers decreases and the price tends to drop toward marginal cost.

17. The initially high price at introduction and its full subsequent decline would not be brought into the index fully by the usual methods. Compilers commonly either wait until the index is rebased or until a product in the sample becomes unavailable to seek a replacement product and admit the possibility of detecting a new good. After capacity constraints or monopoly profits diminish, subsequent price changes may show little difference from other broadly similar products. Standard approaches thus wait too long to pick up these early downtrends in the prices of new goods.

18. At the extreme, capturing the initial price decline requires a comparison between the first observed price and a hypothetical price for the period before its introduction. The hypothetical price would be the price below which there would be no positive market

equilibrium quantity bought and sold.[2] Again, frequent resampling offers the possibility of catching new goods early in the product cycle when their prices are high and market share relatively low, thereby capturing early price declines as producers relieve capacity constraints and new entrants compete market imperfections away.

19. Finally, it is important to emphasize that there is not only a price decline but also a market share increase in the stylized product life cycle. Frequent resampling and focused scanning for new products should be at least somewhat effective in capturing the price declines in early product cycles. Compilers face a potentially serious problem, however, if they have no market share information to go with the prices. The stylized facts of the product cycle are that a new product comes in at a high price and a low market share. The price then declines and market share increases. Both prices and market share then stabilize for a period, until a successor product emerges at a high price and low market share and then begins to take market share from the now mature existing product. Early and normally large price declines for new products thus should figure into the elementary aggregate price index at relatively low weight, while later and normally smaller price declines figure in at successively higher weight. Without current market share data, early price declines may well be overemphasized and the growth in the price index for the elementary aggregate underestimated.

A.3 Temporarily missing products

20. Products that are *temporarily* missing are not available and thus not priced in the month in question, but are expected to be priced in subsequent months. The lack of availability may be because, for example, inventories are insufficient to meet demand, or material inputs are seasonal, as is the case with some fruits and vegetables for food canning. There may also be shortages. The treatment of seasonal goods is the subject of chapter 22. Such goods may go missing in some months, but are distinguished by the fact that they are expected to reappear, at similar level of quality, in the next season.

21. Standard good survey management practice requires that seasonal products be separately identified by the price collector as "temporarily missing" or "seasonal," so compilers can remain alert to the product's reappearance later in the year. Principles and methods for dealing with such products are outlined in Armknecht and Maitland-Smith (1999), Feenstra and Diewert (2001), and Chapter 22.

A.4 Outline for the remainder of the chapter

22. Section B.1 first considers further what is meant by quality change and then considers conceptual issues for the valuation of quality differences. The meaning of quality change requires a conceptual and theoretical platform so that adjustments to prices for quality

[2]This hypothetical price is the *reservation price*, for the CPI. It is the highest notional price at which the quantity demanded would have been zero. The user's reservation price thus will be *higher* than the first observed price. (Hicks, 1940).

differences are made against a well-considered framework. Section B.2 examines quality-adjustment techniques in a national accounting context. Readers interested only in methods of quality adjustment will find them in Sections C through G. Section C provides an overview of the methods available for dealing with unavailable price observations. Methods for quality-adjusting prices are classified into two types: *implicit* and *explicit* adjustments, covered in greater depth in Sections D and E, respectively. Section F considers how to choose among methods of quality adjustment.

23. The implicit and explicit adjustment methods to be used when matching fails—when there are missing, matched models—are first outlined. However, where products are experiencing rapid technological change, these methods may be unsuitable. The use of imputations and patching in of quality-adjusted replacement prices when the matching fails is appropriate when failed matches are the exception. But in high-technology product markets likely to experience rapid turnover of models, they are the rule. Section G considers alternative methods using chained or hedonic frameworks to meet the needs of rapidly changing production and purchasing portfolios. Section H examines frequent resampling as an intermediary, and for imputation a more appropriate, approach. Chapter 22 discusses issues relating to seasonal products in more detail.

B. What is meant by quality change

Before turning to the methods of quality adjustment, the nature of quality changes is first discussed along with a brief outline of the conceptual basis for the indices. Sections C to G include material on the methods price statisticians might employ in dealing with the problem of quality adjustment. In choosing between, and applying, some of these methods arguments may be made as to the conceptual basis as to how quality should be valued. This in turn requires a conceptual basis for the indexes, which again, in turn, should be one that is not at odds with the national accounts system, of which trade price deflators are part. Thus this digression into conceptual issues, though readers interested in specific adjustment methods may move to section C.

B.1 Nature of quality change

24. Bodé and van Dalen (2001) undertook an extensive study of the prices of new automobiles in the Netherlands between 1990 and 1999. The average price increase per car over this period was around 20 percent, but the mix of average quality characteristics changed at the same time. For example, the horsepower (HP) of new cars increased on average from 79 to 92 HP; the average efficiency of fuel consumption improved from 9.3 to 8.4 litres/100km; the share of cars with fuel injection went from 51 percent to 91 percent; the share of cars with power steering went from 27 percent to 94 percent; and the share of cars with airbags went from 6 percent to 91 percent. There were similar increases for central locking, tinted glass, and many more features.

25. Standard price index practice matches the prices of a sample of models in, for example, January with the same models in subsequent months. This holds the characteristics mix

constant to keep quality differences from contaminating the estimate of price change. However, as considered later in this chapter, the resulting sample of matched models (products) is one that gives less weight (if any) to models subsequently introduced. Yet the later models benefit from more recent technological developments and may have different price changes given the quality of services they provide. One approach to correct for such quality changes using the whole sample of both new and existing models is a dummy variable hedonic regression (see Section G.2.1). Bodé and van Dalen (2001), using a variety of formulations of hedonic regressions, found the quality-corrected prices of these new automobiles to be about constant over this period. In this case, the value of the quality improvements explained the entire nominal price increase.

26. Recorded changes in prices are the outcome of shifts in both demand and supply. Chapter 21 explains that these shifts arise from a number of sources, including environmental changes; changes in users' technology, tastes, and preferences; and changes in producers' technology. More formally, the observed data on prices are the loci of the intersections of the demand curves of different final users with varying tastes or intermediate users with possibly varying technologies, and the supply curves of different producers with possibly varying technologies. Separately identifying the effects of changes in environment, technology, and tastes and preferences on the spectrum of product characteristics present in markets at any given time is conceptually and empirically difficult. Fortunately, as Bodé and van Dalen and others demonstrate, compilers do not have to separately identify these effects to produce a good price index in the face of quality change. They need only identify their overall impact.

27. Our concern is not just with the changing mix of the observed characteristics of products. There is the practical problem of not always being able to observe or quantify characteristics, such as style, reliability, ease of use, and safety. The *System of National Accounts 1993* (*1993 SNA*, Chapter 16) notes factors other than changes in physical characteristics that improve quality. These include

> Transporting a good to a location in which it is in greater demand is a process of production in its own right in which the good is transformed into a higher quality good. [Paragraph 16.107]

The same good provided at a more convenient location may command a higher price and be of higher quality. Further, different times of the day or periods of the year may also give rise to quality differences:

> For example, electricity or transport provided at peak times must be treated as being of higher quality than the same amount of electricity or transport provided at off-peak times. The fact that peaks exist shows that purchasers or users attach greater utility to the services at these times, while the marginal costs of production are usually higher at peak times…. [Paragraph 16.108]

28. Other differences, including the conditions of sale and circumstances or environment in which the goods or services are supplied or delivered, can make an important contribution to differences in quality. An outlet, for example, may attract customers by providing better

service; more credit opportunities; shorter waiting lines; and better car parking. These sorts of benefits may well be price-determining. If so, they belong among the characteristics in the product's structured definition.

29. There is a very strong likelihood some price-determining characteristics will be unmeasured in any quality adjustment situation. Compilers cannot produce timely statistics if they are perpetually seeking more characteristics data to produce a still better quality adjustment. How many characteristics data are enough? Characteristics data are sufficient when products are described completely enough. Products are described completely enough when there is low variability of prices over transactions with that description in any given month. If we use characteristics from a structured product description to estimate a hedonic regression model, as did Bodé and van Dalen, the model will fit well only if the structured descriptions are reasonably complete. The first criterion for sufficiency of structured characteristics data, then, is a good fit to a hedonic model. If there is a good fit using a set of objective characteristics, there may be still other characteristics such as style and reliability not yet included in the structured description and thus unmeasured, but they cannot contribute much more to the fit of the model. A second, qualitative criterion is that the included characteristics be meaningful to the participants in the market for the product.

B.2 Conceptual issues

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

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

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

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

33. Defining quality change in terms of its effect on utility is of obvious benefit to the economic approach to index numbers (Chapter 21). Fixler & Zieschang (1992), Feenstra (1995), Triplett (1987) and Diewert (2003a) have developed theoretical frameworks for COLIs akin to those defined in Chapter 21, but which also incorporate goods and services whose quality changes. Silver & Heravi (2001a and 2003) and Kokoski *et al.* (1999) have undertaken empirical studies based on these frameworks for comparisons over time and between geographical areas, respectively. The use of utility as a guide towards understanding quality adjustments to prices is not, however, confined to the economic theory of cost-of-living indices (Chapter 21). Consumer price indices based on a fixed basket concept have the pragmatic need to adjust for quality differences when an item is unavailable, and there is nothing in the definition of a fixed basket index that precludes differences in utility being used as a guideline. If item A is better than its old version, item B, it is because it delivers something more to the consumer who is willing to pay more. That "thing" is called utility.

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

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

of a characteristic is the appropriate quality adjustment for the output price index, and its user value is the quality adjustment for the COLI index or input index."

35. In Chapter 21 the problems of empirically identifying, using hedonic regressions, a valuation of quality characteristics in such clear terms is taken up and in Chapter 21 Section B.6 a formulation from a pure user cost approach is derived.

C. An Introduction to Methods of Quality Adjustment When Matched Items Are Unavailable

C.1 Introduction

36. It may be apparent from the preceding text that quality adjustments to prices are not going to be a simple issue or involve routine mechanical methods whereby a single methodology will be applied to prices in all products groups to yield adjustments. A number of alternative approaches will be suggested, and some will be more appropriate than others for specific items regardless of their product group. An understanding of the technological features of the producing industry, the product market, and alternative data sources will be required for the successful implementation of a quality-adjustment program. Specific attention must be devoted to product areas with relatively high weights and where large proportions of products are turned over. Some of the methods are not straightforward and require some expertise, although methods learned and used on some products may be applicable elsewhere. The issue of quality adjustment is met by developing a gradual approach on an product-by-product basis. It is emphasized that such concerns should not be used as reasons to obviate the estimation of quality-adjusted prices. The practice of statistical agencies in dealing with missing products, even if it is to ignore them, implicitly involves a quality adjustment, and the form of the implicit one undertaken may not be the most appropriate one and may even be misleading. The extent of quality changes and the pace of technological change require that appropriate methods be used.

37. To measure aggregate price changes, a representative sample of products are selected along with a sample of producing/distributive firms along with a host of details that define each *price*, including details on the conditions of the sale where relevant. This is to establish an insight into the price basis of the product. This is then followed by a periodic survey for which the firms report prices (reprice the product) each month for these selected products. They do so to the same specifications, that is, on the same price basis. The detailed specifications are included on the repricing form each month as a prompt to ensure that the price basis has remained the same. Price collectors must be aware of the need to report the details of any change in the price basis; confusion may lead to biased results. It must be borne in mind that firms have no incentive to report such changes since this will invariably involve additional work in costing the change. Attention should also be devoted to ensuring that the description of the price basis contains all pertinent, price-determining elements. If an element is excluded, any change is much less likely to be reported. In both of these cases, the quality change would be invisible to the price measurement process.

C.2 Methods for making quality adjustments

38. When a product is missing in a month for reasons other than being off-season or off-cycle, the replacement may be of a different quality—the price basis may have changed, and like may be no longer compared with like. A number of approaches exist for dealing with such situations and are well documented for the CPI, as outlined in Turvey and others (1989); Moulton and Moses (1997); Armknecht, Lane, and Stewart (1997); Moulton, LaFleur, and Moses (1998); and Triplett (2002). Though the terms differ among authors and statistical agencies, they include

- *Imputation*—When no information is available to allow reasonable estimates to be made of the effect on price of a quality change. The price change of all products—or of more or less similar products—are assumed to be the same as that for the missing product.

- *Overlap*—Used when no information is available to allow reasonable estimates to be made of the effect on price of a quality change, but a replacement product exists in the same period as the old product. The price difference between the old product and its replacement in the same overlap period is then used as a measure of the quality difference.

- *Direct comparison*—If another product is directly comparable, that is, so similar it has more or less the same quality characteristics as the missing one, its price replaces the unavailable price. Any difference in price level between the new and old is assumed to be because of price changes and not quality differences.

- *Explicit quality adjustment*—When there is a substantial difference in the quality of the old and replacement products, estimates of the effect of quality differences on prices are made to enable quality-adjusted price comparisons.

39. Before outlining and evaluating these methods, one should say something about the extent of the problem. This situation arises when the product is unavailable. It is not just a problem when *comparable* products are unavailable, for the judgment as to what is and what is not comparable itself requires an estimate of quality differences. Part of the purpose of a statistical meta-information system for statistical offices (outlined in Chapter 8) is to identify and monitor the sectors that are prone to such replacements and determine whether the replacements used really are comparable.

40. Quality adjustment methods for prices are generally classified into the implicit or imputed (indirect) methods explained in Section D (the differences in terminology are notorious in this area) and explicit (direct) methods explained in Section E. Both decompose the price change between the old product and its replacement into quality and pure price changes. However, in the latter, an explicit estimate is made of the quality difference, usually on the basis of external information. The pure price effect is identified as a remainder. For implicit adjustments, a measurement technique is used to compare the old product with the replacement, so that the extent of the quality and pure price change is implicitly determined

by the assumptions of the method. The accuracy of the method relies on the veracity of the assumptions, not the quality of the explicit estimate. In Sections D and E, the following methods are considered in detail:

Implicit methods:

- Overlap;
- Overall-mean/targeted mean imputation;
- Class-mean imputation;
- Comparable replacement;
- Linked to show no price change; and
- Carryfoward.

Explicit methods:

- Expert judgment;
- Quantity adjustment;
- Differences in production/option costs; and
- Hedonic approach.

C.3 Some points

C.3.1 Additive versus multiplicative

Table 7.1. Estimating a Quality-Adjusted Price

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

41. The quality adjustments to prices may be undertaken by either adding a fixed amount or multiplying by a ratio. For example, where m is the old product and n its replacement for a comparison over periods t, $t+1$, $t+2$, the use of the overlap method in period $t+1$ required the ratio p_n^{t+1}/p_m^{t+1} to be used as a measure of the relative quality difference between the old item and its replacement. This ratio could then be *multiplied* by the price of the old item in period t, p_m^t to obtain the quality-adjusted prices p_m^{*t} shown in Table 7.1. Such multiplicative formulations are generally advised because the adjustment is invariant to the absolute value of the price. It would be otherwise possible for the absolute value of the change in specifications to exceed the value of the product in some earlier or—with technological advances—later period. Yet for some products, the worth of the

constituent parts is not in proportion to their price. Instead, they have their own intrinsic, absolute, additive worth, which remains constant over time. Producers selling over the Internet may, for example, include postage, which in some instances may re-main the same irrespective of what is happening to price. If postage is subsequently excluded from the price, the fall in quality should be valued as a fixed sum.

C.3.2 Base- versus current-period adjustment

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

C.3.3 Long-run versus short-run comparisons

43. Much of the analysis of quality adjustments in this *Manual* has been undertaken by comparing prices between two periods (for example, periods 0 and 1). For long-run comparisons, suppose the base period is taken as period *t* and the index is compiled by comparing prices in period *t* first with *t* + 1, then with *t* + 2, then with *t* + 3, etc. The short-run framework allows long-run comparisons—say, between periods *t* and *t* + 3—to be built as a sequence of links joined by successive multiplication—say, period *t* with *t* + 2 and period *t* + 2 with *t* + 3. This can also be done by chaining period *t* with *t* + 1, *t* + 1 with *t* + 2, and *t* + 2 with *t*+3. In Section H, the advantages of the short-run framework for imputations are outlined. In Section G.3, chained indices are considered for industries experiencing a rapid turnover in products. These quality adjustment methods are now examined in turn, and in Section F, the choice of method is discussed.

C.3.4 Statistical metadata

44. In Sections D and E, implicit and explicit methods of quality adjustments to prices are discussed. In Section F, the choice between these methods is examined. Any consideration of the veracity of these methods, resource implications, and the choice between them needs to be informed by appropriate information on an product-by-product basis. Section C of Chapter 8 considers information requirements for a strategy for such quality adjustment which makes use of a statistical metadata system.

D. Implicit Methods

D.1 Overlap method

45. Consider an example where the items are sampled in January and prices are compared over the remaining months of the year. Matched comparisons are undertaken between the January prices and their counterparts in successive months. Five products are assumed to be sold in January with prices p_1^1, p_2^1, p_5^1, p_6^1, and p_8^1 (Table 7.2, part a). Two types of similar products are sold in the product group concerned, A and B. An index of the elementary level is required for the overall price change of these two product types. At this level of aggregation, the weights can be ignored assuming only one quote is taken on each product. A price index for February compared with January = 100.0 is straightforward in that prices of products 1, 2, 5, 6, and 8 are used and compared only by way of the geometric mean of price ratios, known as the Jevons index (which is equivalent to the ratio of the geometric mean in February over the geometric mean in January—see Chapter 20). In March, the prices for products 2 and 6—one of type A and one of type B—are missing.

46. In Table 7.2, the lower part (b) is a numerical counterpart of the upper part (a), further illustrating the calculations. The overlap method requires prices of the old and replacement products to be available in the same period. In Table 7.2(a), product 2 has no price quote for March. Its new replacement is, for example, product 4. The overlap method simply measures the ratio of the prices of the old and replacement product in an overlap period. In this example, the period is February, and the old and replacement products are products 2 and 4, respectively. This is taken to be an indicator of their quality differences. The two approaches outlined in Section C.3.2 are apparent: either to insert a quality-adjusted price in January for product 4 and continue to use the replacement product 4 series, or continue the product 2 series by patching in quality-adjusted product 4 prices. Both yield the same answer. Consider the former. For a Jevons geometric mean from January to March *for product type A only*, assuming equal weights of unity

$$(7.1)\ P_J(p^1, p^3) = \left[p_1^3 / p_1^1 \times p_4^3 / \left(\left(p_4^2 / p_2^2 \right) \times p_2^1 \right) \right]^{1/2}$$

$$= [\ 6/4 \times 8/\ ((7.5\ /\ 6) \times 5)]^{1/2}$$

$$= 1.386.$$

47. Note that the comparisons are long-run ones, that is, they are between January and the month in question. The short-run (modified) Laspeyres framework provides a basis for short-run changes based on data in each current month and the immediately preceding one. In Table 7.2(a) and (b), the comparison for product type A would first be undertaken between January and February using products 1 and 2. The result would be multiplied by the comparison between February and March using items 1 and 4. Still, this implicitly uses the differences in prices in the overlap in February between items 2 and 4 as a measure of this quality difference. It yields the same result as before:

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

The advantage of recording price changes for, say, January to October in terms of January to September and September to October is that it allows the compiler to compare immediate month-on-month price changes for data editing purposes. Moreover, it has quite specific advantages for the use of imputations as discussed in Sections D.2 and D.3 for which different results arise for the long- and short-run methods. A fuller discussion of the long-run and short-run frameworks is undertaken in Section H.

Table 7.2. Example of Overlap Method of Quality Adjustment

(a) General Illustration

Product Type	**Item**	**January**	**February**	**March**	**April**
A	1	p_1^1	p_1^2	p_1^3	p_1^4
	2	p_2^1	p_2^2		
	3			p_3^3	p_3^4
	4		p_4^2	p_4^3	p_4^4
B	5	p_5^1	p_5^2	p_5^3	p_5^4
	6	p_6^1	p_6^2		
	7			p_7^3	p_7^4
	8	p_8^1	p_8^2	p_8^3	p_8^4

(b) Numerical Illustration				
Product Type	**Item**	**January**	**February**	**March**
A	1	4.00	5.00	6.00
	2	5.00	6.00	
	2. overlap			**6.90**
	2. imputation			**6.56**
	2.trgeted imputation			**7.20**
	2. comparable replacement			**6.50**
	3			6.50
	4		7.50	8.00
B	5	10.00	11.00	12.00
	6	12.00	12.00	
	6. imputation			**13.13**
	6. targeted imputation			**12.53**
	7			14.00
	8	10.00	10.00	10.00

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

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

$$= \frac{p_n^{t+1}}{p_m^{t-1}} \times \frac{p_m^t}{p_n^t}.$$

49. The quality adjustment to prices in period *t* + 1 is defined as before, $p_m^{*t+1} = A(z^{t+1})\, p_n^{t+1}$, which is the adjustment to p_n in period *t* + 1, which equates its value 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:

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

The overlap formulation equals this when

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

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

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

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

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

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

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

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

Note that the quality difference here is not related to the technical specifications or resource costs but to the relative price purchasers pay.

51. Relative prices may also reflect unusual pricing policies aimed at minority segments of the market. In the example of pharmaceutical drugs (Berndt, Ling, and Kyle, 2003), the overlapping prices of a generic and a name brand product were argued to be reflective of the needs of two different market segments. The overlap method can be used with a judicious choice of the overlap period. It should be a period before the use of the replacement, since in such periods the pricing may reflect a strategy to dump the old model to make way for the new one.

52. The overlap method is implicitly employed when samples of products are rotated, meaning that the old sample of products 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 product-to-item level—that differences in price levels at a common point in time accurately reflect differences in qualities.

53. The overlap method has at its roots a basis in the law of one price. The law states that when a price difference is observed, it must be the result of some difference in physical quality or some such factor 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 difference would not persist given markets made up of rational producers and consumers. However, *1993 SNA* (Chapter 16) notes three reasons why this might fail:

- First, purchasers may not be properly informed about existing price differences and may therefore inadvertently buy at higher prices. While they may be expected to search out for the lowest prices, costs are incurred in the process.

- Secondly, purchasers may not be free to choose the price at which they purchase because the seller may be in a position to charge different prices to different categories of purchasers for identical goods and services sold under exactly the same circumstances—in other words, to practice price discrimination.

- Thirdly, buyers may be unable to buy as much as they would like at a lower price because there is insufficient supply available at that price. This situation typically occurs when there are two parallel markets. There may be a primary, or official, market in which the quantities sold, and the prices at which they are sold are subject to government or official control, while there may be a secondary market—a free market or unofficial market—whose existence may or may not be recognized officially.

54. There is extensive literature in economics dealing with theory and evidence of price dispersion and its persistence, even when quality differences have been accounted for. The differences can be substantial: Yoskowitz's (2002) study for raw water found one supplier discriminating against a private customer, charging $500 per acre foot (AF) while a municipality was charged $20 per AF, though there was some evidence of arbitrage and learning. It is not the role of this *Manual* to examine such theories and evidence, so readers are referred to the following studies: Stigler (1961) and Lach (2002) on search cost theory; Sheshinski and Weiss (1977) and Ball and Mankiw (1994) on menu cost theory; and Friedman (1977) and Silver and Ioannidis (2001) on signal extraction models.

D.2 Overall mean/targeted mean Imputation

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

56. In the example in Table 7.2(b), the January to February comparison for both product types is based on products 1, 2, 5, 6, and 8. For March compared with January—weights all equal to unity—the product 2 and product 6 prices are imputed using the short-run price change for February (p^2) compared with March (p^3) based on products 1, 5, and 8. Since different formulas are used for elementary aggregation, the calculation for the three main formulas are illustrated here (see Chapter 20 for choice of formulas). The geometric mean of the price ratios—the Jevons index—is

$$(7.7)\ P_J\left(p^2, p^3\right) = \left[\prod_{i=1}^{n} p_i^3 / p_i^2\right]^{1/3}$$

$$= \left[\left(p_1^3/p_1^2\right) \times \left(p_5^3/p_5^2\right) \times \left(p_8^3/p_8^2\right)\right]^{1/3}$$

$$= \left[(6/5) \times (12/11) \times (10/10)\right]^{1/3}$$

= 1.0939, or a 9.39 percent increase.

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

$$(7.8)\ P_D(p^2, p^3) = \sum_{i=1}^{N} p_i^3 / N / \sum_{i=1}^{N} p_i^2 / N$$

$$= \left(p_1^3 + p_5^3 + p_8^3\right)/3 \ \div \ \left(p_1^2 + p_5^2 + p_8^2\right)/3$$

= (6 + 12 + 10) / (5 + 11 + 10) = 1.0769,

or a 7.69 percent increase.

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

$$(7.9)\ P_C(P^3, P^2) = \sum_{n=1}^{N} (p_n^3 / p_n^2)/N$$

$$= \left[\left(p_1^3 / p_1^2\right) \ + \ (p_5^3 / p_5^2) \ + \ (p_8^3 / p_8^2)\right]/3$$

= [(6/5 + 12/11 + 10/10)] / 3 = 1.09697,

or a 9.697 percent increase.

57. In practice the imputed figure would be entered onto the data sheet. Table 7.2(b) has the overall mean imputation in March for products 2 and 6, using the Jevons index, as 1.0939 x 6 = 6.563 and 1.0939 x 12 = 13.127, respectively, (bold type). It should be noted that the Dutot index is in this instance lower that the Jevons index, a result not expected from the relationships established in Chapter 20. The relationship in Chapter 20 assumed the variance in prices would increase over time whereas in Table 7.2(b), it decreases for the three products. The arithmetic mean of price relatives—the Carli index—equally weights each price change, but the ratio of arithmetic means—the Dutot index—weights price changes according to the prices of the product in the base period relative to the sum of the base-period prices. Item 1 has a relatively low price, and thus weight, in the base period 1 of 4, but this product has the highest price increase, one of 6/5. Therefore, the Dutot index is lower than the Carli index.

58. As noted above, it is also possible to refine the imputation method by targeting the imputation: including the weight for the unavailable products in groupings likely to experience similar price changes—say, by product type. Any stratification system used in the selection of

outlets and product varieties would facilitate this. For example, in Table 7.2(b) assume that the price change of the missing product 2 in March is more likely to follow price changes of product 1, and product 6 is more likely to experience price changes similar to products 5 and 8. For March compared with February, with weights all equal to unity, the geometric mean of price ratios (Jevons) is

$$(7.10)\ P_J(p^2,p^3)=\prod_{n=1}^{N}(p_n^3/p_n^2)^{1/N}$$
$$=\left[\left(p_1^3/p_1^2\right)^2\times\left(p_5^3/p_5^2\times p_8^3/p_8^2\right)^{3/2}\right]^{1/5}$$
$$=\left[\left(6/5\right)^2\times\left(12/11\ \times\ 10/10\right)^{3/2}\right]^{1/5}$$
$$=1.1041.$$

Note the weights used: for product type A, the single price represents 2 prices; for product type B, the prices represent three or 3/2 = 1.5 each.

59. The ratio of average (mean) prices—the Dutot index—is

$$(7.11)\ P_D(p^2,p^3)=(\sum_{n=1}^{N}p_n^3/N)/(\sum_{n=1}^{N}p_n^2/N)$$
$$=\left[\frac{\left(2p_1^3+1.5p_5^3+1.5p_8^3\right)/5}{\left(2p_1^2+1.5p_5^2+1.5p_8^2\right)/5}\right]$$
$$=\left[\frac{\left(2\times6\ +\ 1.5\times12\ +\ 1.5\times10\right)/5}{\left(2\times5\ +\ 1.5\times11\ +\ 1.5\times10\right)/5}\right]$$
$$=1.0843.$$

60. The average (mean) of price ratios—the Carli index—is:

$$(7.12)\ P_C(p^2,p^3)=\sum_{i=1}^{N}(p_i^3/p_i^2)/N$$
$$=\frac{2}{5}\left(p_1^3/p_1^2\right)\ +\ \frac{3}{5}\left[\left(p_5^3/p_5^2\right)\ +\ \left(p_8^3/p_8^2\right)/2\right]$$
$$=\frac{2}{5}(6/5)\ +\ \frac{3}{5}\left[(12/11)\ +\ (10/10)/2\right]$$
$$=\ 1.1073$$

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

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

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

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

and for *x* unavailable products by

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

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

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

$$r_2 = \left[\sum_{i=1}^{x} p_i^{*t+1} / p_i^t\right] \div x,$$

then the ratio of arithmetic mean biases from substituting equation (7.16) into equation (7.15) is

(7.17) $Q = \frac{x}{m}\left(r_2 - r_1\right)$,

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 products 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 relies on a comparison between price changes for existing products and *quality-adjusted* price changes for the replacement/unavailable comparison. This is more likely to be justified than a comparison without the quality adjustment to prices. For example, let us say there were *m* = 3 products, each with a price of 100 in period *t*. Let the *t* + 1 prices be 120 for two products, but assume the third is unavailable, that is, *x* = 1, and is replaced by a product with a price of 140, of which 20 is the result of quality differences. Then the arithmetic bias as given in equations (7.16) and (7.17) where $x = 1$ and *m* = 3 is

$$\begin{aligned}&\tfrac{1}{3}\left[(-20 + 140)/\ 100\right]\\ &\quad -\tfrac{1}{3}\left[\left(\tfrac{120}{100} + \tfrac{120}{100}\right)/2\right]\\ &\quad = 0\end{aligned}$$

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

63. It is noted that the analysis here is framed in terms of a short-run price change framework. This means that the short-run price changes between two consecutive periods are used for the imputation. This is different from the long-run imputation, where a base-period price is compared with prices in subsequent months and where the implicit assumptions are more restrictive.

64. Table 7.3 provides an illustration whereby the (mean) price change of products that continue to exist, r_1, is allowed to vary for values between 1.00 and 1.50: no price change and a 50 percent increase. The (mean) price change of the *quality-adjusted* new products compared with the products they are replacing is assumed to not change, that is, r_2 = 1.00. The bias is given for ratios of missing

Table 7.3. Example of the Bias from Implicit Quality Adjustment for r_2 = 1.00

	Geometric mean Ratio of missing products, x/m					Arithmetic mean Ratio of missing products, x/m				
r_1	0.01	0.05	0.10	0.25	0.50	0.01	0.05	0.10	0.25	0.50
1.00	1	1	1	1	1	0	0	0	0	0
1.01	0.999901	0.999503	0.999005	0.997516	0.995037	–0.0001	–0.0005	–0.001	–0.0025	–0.005
1.02	0.999802	0.999010	0.998022	0.995062	0.990148	–0.0002	–0.0010	–0.002	–0.0050	–0.010
1.03	0.999704	0.998523	0.997048	0.992638	0.985329	–0.0003	–0.0015	–0.003	–0.0075	–0.015
1.04	0.999608	0.998041	0.996086	0.990243	0.980581	–0.0004	–0.0020	–0.004	–0.0100	–0.020
1.05	0.999512	0.997563	0.995133	0.987877	0.975900	–0.0005	–0.0025	–0.005	–0.0125	–0.025
1.10	0.999047	0.995246	0.990514	0.976454	0.953463	–0.0010	–0.0050	–0.010	–0.0250	–0.050
1.15	0.998603	0.993036	0.986121	0.965663	0.932505	–0.0015	–0.0075	–0.015	–0.0375	–0.075
1.20	0.998178	0.990925	0.981933	0.955443	0.912871	–0.0020	–0.0100	–0.020	–0.0500	–0.100
1.30	0.997380	0.986967	0.974105	0.936514	0.877058	–0.0030	–0.0150	–0.030	–0.0750	–0.150
1.50	0.995954	0.979931	0.960265	0.903602	0.816497	–0.0050	–0.0250	–0.050	–0.1250	–0.250

values of 0.01, 0.05, 0.10, 0.25, and 0.50, 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 products 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 should be 0.9759 × 1.05 = 1.0247. For an arithmetic mean, the bias is –0.025; instead of 1.05, it should be 1.025.

65. Equation (7.17) shows that the ratio *x*/*m* and the difference between r_1 and r_2 determine the bias. Table 7.3 shows that the bias can be quite substantial when *x*/*m* is relatively large. For example, when *x*/*m* = 0.25, an inflation rate of 5 percent for existing products translates to an index change of 3.73 percent and 3.75 percent for the geometric and arithmetic formulations, respectively, when r_2 = 1.00, that is, when quality-adjusted prices of unavailable products are constant. Instead of being 1.0373 or 1.0375, ignoring the unavailable products would give a result of 1.05. Even with 10 percent missing (*x*/*m* = 0.1) an inflation rate of 5 percent for existing products translates to 4.45 percent and 4.5 percent for the respective geometric and arithmetic formulations when r_2 = 1.00. However, consider a fairly low ratio of *x*/*m*, say, 0.05; then even when r_2 = 1.00 and r_1 = 1.20, Table 7.3 finds 1.89 percent and 1.9 percent corrected rates of inflation for the respective geometric and arithmetic formulations. 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 product and the old product *after adjusting for quality differences*. If r_1 and r_2 are the same, then there would be no bias from the method even if *x*/*m* = 0.9. There may, however, be more sampling error. It should be borne in mind that it is not appropriate to compare bias between the arithmetic and geometric means, at least in the form they take in Table 7.3. The latter would have a lower mean, rendering comparisons of bias meaningless.

66. An awareness of the market conditions relating to the commodities is instructive to any understanding of likely differences between r_1 and r_2. The concern here is when prices vary over the life cycle of the products. Thus, at the introduction of a new model, the price change may be quite different from price changes of other existing products. Assumptions of similar price changes, even when quality adjusted, might be inappropriate. Greenlees (2000) uses the example of personal computers: new computers enter the market at prices equal to or lower than prices of previous models but with greater speed and capability. An assumption that $r_1 = r_2$ could not be justified.

67. Some of this bias relates to the fact that markets are composed of different market segments of consumers and producers tailor their output to meet such needs. Indeed, the very training of marketers involves consideration of developing different market segments and ascribing to each segment appropriate *pricing*, *product quality*, *promotion*, and *place* (methods of distribution). This is known as the 4 Ps of the marketing mix (Kotler, 1991). In addition, marketers are taught to plan the marketing mix over the product's life cycle. Such planning would allow 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, whereby higher prices are charged to skim off the surplus from segment(s) of purchasers willing to pay more. The economic theory of price discrimination would also predict such behavior. Thus, the quality-adjusted price change of an old product compared with a new replacement product may be higher than price changes of other products in the product group. After the introduction of the new product, its prices may fall relative to others in the group. There may be no law of one price

change for differentiated products within a market. Berndt, Ling, and Kyle (2003) clearly showed how after patent expiration, the price of brand name prescription pharmaceuticals can increase with the entry of new generic pharmaceuticals at a lower price, particularly as loyal, less-price-sensitive customers maintain their allegiance to the brand name pharmaceuticals.

68. There is little in economic or marketing theory to support any expectation of similar (quality-adjusted) price changes for new and replacement products and other products 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 things matter in any decision to use the imputation approach. The first is the proportion of replacements and, Table 7.3 provides guidance here. The second is the expected difference between* r_1 *and* r_2, *and it is clear from the above discussion that there are markets in which they are unlikely to be similar.* This is not to say the method should not be used. It is a simple and expedient approach. Arguably what should not happen is that the method is used as a default process without any prior evaluation of expected price changes and the timing of the switch. Furthermore, attention should be directed to its targeted use, using products expected to have similar price changes. However, the selection of such products should also be based on the need to include a sufficiently large sample so that the estimate is not subject to undue sampling error.

69. Some mention should be made of the way these calculations are undertaken. A pro forma setting for the calculations—say, on a spreadsheet—would have each product description and its prices recorded on a (usually) monthly basis. The imputed prices of the missing products are inserted into the spreadsheet being highlighted as imputed. The reasons for highlighting such prices are (i) because they should not be used in subsequent imputations as if they were actual prices and (ii) the inclusion of imputed values may give the 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. It is stressed that this is an illustration of a *short-run* imputation, and, as will be discussed in Section H, there is a strong case for using *short-run* imputations against *long-run* ones.

D.3 Class mean imputation

70. The *class mean* (or *substitution relative*) method of implicit quality adjustment to prices as used in the U.S. CPI is discussed in Schultz (1995); Reinsdorf, Liegey, and Stewart (1996); Armknecht, Lane, and Stewart (1997); and Armknecht and Maitland-Smith (1999). It arose from concerns similar to those considered in Section D.2, namely that unusual price changes were found in the early introductory period when new models were being introduced, particularly for consumer durables. In their study of selected products, Moulton and Moses (1997), using U.S. CPI data for 1995, found the average pure price change to be only 0.12 percent for identical products being repriced (on a monthly or bimonthly basis). This is compared with an average of 2.51 percent for comparable substitutes—items judged equivalent to the products they replaced. The corresponding average price change for directly substituted quality-adjusted price changes was 2.66 percent. Therefore, the price movement of continuing products appears to be a flawed proxy for the pure price component of the difference between old and replacement items.

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

D.4 Comparable replacement

72. This is where the price collector makes a judgment that the replacement is of a similar quality to the old product and any price changes are untainted by quality changes. For product type A in Table 7.2(b), product 3 might be judged to be comparable to product 2 and its prices in subsequent months used to continue the series. In March the price of 6.5 would be used as the price in March for product 2, whose January to March price change would be $6.5/6 \times 100 = 1.0833$ or 8.33 percent. Lowe (1998) noted the common practice of television set manufacturers changing model numbers when there is a new production run, though nothing physically has changed, or when small changes take place in specifications, such as the type of remote controls or the number or placement of jacks. The method of comparable replacement relies on the efficacy of the price collectors and, in turn, on the adequacy of the specifications used as a description of the price basis. Statistical agencies may be rightly wary of sample sizes being worn down by dropping products using imputation and also of the resource-intensive explicit estimates outlined below. The use of repriced products of a comparable specification has much to commend it. If, however, the quality of products is improving, the preceding product will be inferior to the current ones. In addition, continually ignoring the 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 products are accepted in spite of quality differences, and the weight attached to them. Proposals in Chapter 8 to monitor types of quality adjustment methods by product area will provide a basis for a strategy for applying explicit adjustments where they are most needed.

D.5 Linked to show no price change

73. Linking attributes any price change between the replacement product in the *current* period and the old product in the preceding period to the change in quality. A replacement product 7 is selected, for example, in Table 7.2(b) from product type B for the missing March product 6. The replacement product 7 may be of a very different quality compared with product 6, with the price difference being quite large. The change in price is assumed to be due to a

change in quality. An estimate is made for p_7^2 by equating it to p_7^3 to show no change, that is, the assumed price of product 7 in February is 14 in Table 7.2(b). There is, therefore, assumed to be no price change over the period February to March for product 7. The January to March result for product 6 is (12/12) × (14/14) = 1.00, or no change. However, for the period March to April, the price of item 7 in March can be compared with the imputed p_7^2 for February and linked to the preceding results. So the January to April comparison is composed of the January to February comparison for product 6 and linked to (multiplied by) the February to April comparison for item 7. This linking is analogous to the procedures used for the chained and short-run framework discussed in Sections G.3 and H.3. The method is born out of circumstances where comparable replacement products are not available, and there are relatively large price differences between the old and replacement products, having significant differences in price base and quality. It is not possible to separate out how much of this difference is due to price changes and how much to quality changes, so the method attributes it all to quality and holds price constant. The method introduces a degree of undue price stability into the index. It may well be the case that the period of replacement is when substantial price changes are taking place, these changes being wrongly assigned to quality changes by this method. Article 5 of the European Commission (EC) Regulation No. 1749/96 requires member states to avoid such automatic linking. Such linking is equivalent to the assumption that the difference in price between two successive models is wholly attributed to a difference in quality (Eurostat, 2001, p. 125).

D.6 Carryforward

74. With this method, when a product becomes unavailable—say, in period $t + 2$—the price change calculation uses the old t price, carried forward as if there was no change. Thus, from Table 7.2(a) for product type A for the January to March Jevons and Dutot indices (Chapter 20, Section B)

(7.18) $P_J(p^1, p^3) = \left[\left(p_1^3 / p_1^1 \times p_2^2 / p_2^1\right)\right]^{1/2}$, and

$$P_D(p^1, p^3) = [(p_1^3 + p_2^2)/(p_1^1 + p_2^1)],$$

with p_2^2 filling in for the missing p_2^3. This introduces undue stability into the index, which is aggravated if the old price p_2^2 continues to be used to fill in the unobserved prices in subsequent periods. It introduces an inappropriate amount of stability into the index and may give a misleading impression of the active sample size. The practice of the carry-forward method is banned for harmonized CPIs under Article 6 of the EC Regulation No. 1749/96 for Harmonized Indices of Consumer Prices (Eurostat, 2001, p. 126). To use this method, an assumption is made that the price from this product type would not change. This method should be used only if it is fairly certain that there would be no price change.

E. Explicit Methods

75. All of the aforementioned methods do not rely on explicit information on the value of the change in quality, $A(z)$. Now methods that rely on obtaining an explicit valuation of the quality difference are discussed.

E.1 Expert judgment

76. Hoven (1999) describes comparable replacement as a special case of "subjective quality adjustment," because the determination of product equivalence is based on the judgment of the commodity specialist. It is important to mention this because an objection to subjective methods is the inability to provide results that can be independently replicated. Yet in comparable replacement, and for the selection of representative products, a subjective element is part of normal procedure. This is not, of course, a case for its proliferation.

77. Price collectors would generally not be well placed to quantify the price difference between a new and non-comparable replacement item that can be attributed to the quality change. Alternatively, the use of experts' views may be appropriate for highly complex products where alternative methods are not feasible. Experts, as noted above, should be directed to the nature of the estimate required as discussed in the conceptual section. More than one expert should be chosen, and, where possible, they should be from different backgrounds. Some indication of the interval in which their estimate should lie is also advisable. The well-used Delphi method (for example, see Czinkota, 1997) may be applicable. In this approach, a panel of experts work separately to avoid any bandwagon effect regarding their estimates. They are asked to provide an estimate of the average and range of likely responses. The median is taken of these estimates, and any estimate that is considered extreme is sent back to the expert concerned. The expert is asked to identify reasons for the difference. It may be that the particular expert has a useful perspective on the problem that the other experts had not considered. If the expert argues a case, the response is fed back to the panel members, who are asked if they wish to change their views. A new median is taken, and there are possible further iterations. It is time consuming and expensive but illustrates the care needed in such matters. However, if the adjustment is needed for a product area with a large weighting in a trade price index, and no other techniques are available, it is a possible alternative. In all of this guidelines are required as to the conceptual base for the valuation, as discussed in section B above and E.3 below.

E.2 Quantity adjustment

78. This is one of the most straightforward explicit adjustments to undertake and is applicable to products for which the replacement is of a different size than the available one. In some situations, there is a readily available quantity metric that can be used to compare the products. Examples are the number of units in a package (for example, paper plates or vitamin pills), the size or weight of a container or the size of sheets or towels. Quality adjustment to prices can be accomplished by scaling the price of the old or new product by the ratio of quantities. The index production system may do this scaling adjustment automatically by converting all prices in the category to a price per unit of size, weight, or number. Such scaling is most important. For example, it should not be the case that because a price collecter finds a soft drink is now only available in an outlet in 1 liter containers instead of the previously recorded 0.5-liter ones, its price has doubled.

79. There is, however, a second issue. It should be kept in mind that a pure price change is concerned with changes in the revenue received from the sale of the exact same products, produced under the exact same circumstances, and sold under the exact same terms. In the pharmaceutical context, for example, prices of bottles of pills of different sizes differ. A bottle of

100 pills, each pill having 50 milligrams of a drug, is not the same as a bottle of 50 pills of 100 milligrams each, even though both bottles contain 5,000 milligrams of the same drug. It may also be reasonable to decide that a bottle of aspirin, for example, containing 500 tablets may not have 10 times the quality of a 50-tablet bottle. If the smaller size is no longer available and 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 if there is a differential in the cost of producing and margin on selling the larger size of 2 percent or more. If, however, the price collector acknowledged that the change in packaging size for this product led to a 1 percent saving in resource costs (and margin) and prices of other such products without any quantity changes were also falling by 1 percent, then the pure price change would be a fall of 1 percent. In practice, the price collector may be able to make some rough estimates of the effect on the unit cost of the change in packaging size. However, it may well be that no such information is available, and the general policy is to not automatically interpret unit price changes arising from packaging size changes as pure price changes if contrary information exists.

80. Consider another example: a brand name bag of fertilizer of a specific type, previously available in a 0.5 kg. bag priced at 1.5 is replaced with a 0.75 kg. bag at 2.25. The main concern here is with rescaling the quantities as opposed to differential cost or margin adjustments. The method would use the relative quantities of fertilizer in each bag for the adjustment. The prices may have increased by [(2.25/1.5) × 100 = 150] 50 percent, but the quality (size)-adjusted prices have remained constant [(2.25/1.5) × (0.5/0.75) × 100 = 100].

81. The approach can be outlined in a more elaborate manner by referring to Figure 7.1. The concern here is with the part of the unbroken line between the price and quantity coordinates (1.5, 0.5) and (2.25, 0.75), both of which have *unit* prices of 3 (price = 1.5/0.5 and 2.25/0.75). There should be no change in quality-adjusted prices. The delta symbol (Δ) denotes a change. The slope of the line is β, which is ΔPrice/ΔSize = (2.25 – 1.5)/(0.75 – 0.50) = 3, that is, the change in price arising from a unit (kg.) change in size. The quality (size)-adjusted price in period $t-1$ of the old m bag is

$$(7.19)\quad \hat{p}_m^{t-1} = p_m^{t-1} + \beta \Delta size$$

$$= 1.5 + 3\,(0.75 - 0.5) = 2.25.$$

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

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

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

82. The method can be seen to be successful on intuitive grounds as long as the unit price of different-sized bags remains constant. If the switch was from a 0.5 kg. bag to a 0.25 kg. one priced at 0.75, as shown by the continuation of the unbroken line in Figure 7.1, to coordinate (0.75,0.25) quality-adjusted prices would again not change. However, assume the *unit* (kg.)

prices were 5, 3, and 3 for the 0.25, 0.5, and 0.75 kg. bags, respectively, as shown in the example below and in Figure 7.1 by the *broken* line. Then the measure of quality-adjusted price change would depend on whether the 0.5 kg. bag was replaced by the 0.25 kg. one (a 67 percent increase) or the 0.75 kg. one (no change). This is not satisfactory because the choice of replacement size is arbitrary. The rationale behind the quality adjustment process is to ask: does the difference in unit price in each case not arise from differences in unit costs of producing and unit margins on selling? Are consumers indifferent to the selection of a unit from one packet size as against another? If so, quantity adjustments are appropriate. It may be obvious from the nature of the product that a product packaged in a very small size with disproportionately high unit price has an unusually high profit margin or will have quite different unit production costs and an appropriate replacement for a large-sized product would not be this very small one. If so they should be treated as a different product.

Example of Quantity Adjustments

Size	First Price	First Unit Price	Second Price	Second Unit Price
0.25	0.75	3.00	1.25	5.00
0.50	1.50	3.00	1.50	3.00
0.75	2.25	3.00	2.25	3.00

E.3 Differences in production and option costs

83. A natural approach is to adjust the price of the old product by an amount equal to the costs of the additional features. This approach is associated with resource-cost valuations discussed in Section B.2. Yet Section B.2 advocated a user-value approach, the appropriate valuation being the change in production costs associated with a quality change plus any price-cost margin. This amounts to a comparison of relative prices using

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

and x is the cost or contribution to revenue of the additional features in period $t-1$. The price collector is a natural expert source of such information. Greenlees (2000) provides an example for new trucks and motor vehicles in the United States in 1999. Just before the annual model year introductions, Bureau of Labor Statistics (BLS) staff visit selected manufacturers to collect cost information. The data are used in the PPI and International Price Comparison programs, as well as in the CPI, and the information-gathering activity is a joint operation of the three programs. Allowable product changes for the purpose of quality adjustments include occupant safety enhancements, mechanical and electrical improvements to overall vehicle operation or efficiency, changes that affect length of service or need for repair, and changes affecting comfort or convenience.

84. As an example of option cost adjustments, assume the prices for a product in periods t and $t + 2$ were 10,000 and 10,500, respectively, but assume the price in period $t + 2$ is for the item with a new feature or option. Also, let the price of the additional feature in period $t + 2$ be

300. Then the price change would be 10,200/10,000 = 1.02, or 2 percent. The adjustment may take a multiplicative form (see Section A); the additional options are worth 300/10,500 = 0.028571 of the period $t + 2$ price. The adjusted price in period t is, therefore, 10,285.71 and the price change 10,500/10,285.71 = 1.020833, or about 2 percent. If in subsequent periods either of these elements change, then so too must $\hat{p}_n^{t-1}$ for those comparisons. Option cost is thus a method for use in stable markets with stable technologies. Alternatively, it may be preferable to estimate a one-off adjustment to the preceding base-period price and then compare all subsequent products with the new option to this estimate; that is, 10,500/10,300 = 1.019417, or approximately 2 percent.

85. In the example above, the prices available for the options were sales prices. Sometimes only production costs for options may be known. Production costs of options need to be upgraded to user values by adding price cost markups and indirect taxes. Often such data are available for only one period. If the markups are considered to be in the same proportion in subsequent periods, then there is no problem since the retail price changes would proxy the producer ones after adjustment for proportionate margins. However, if the average age or vintage of the products have changed, then they will be at different stages in their life cycles and may have different margins.

86. Consider the addition of a feature to a product. Chairs, for example, can be produced and sold as standard or with a lever mechanism for height adjustment. The specification may always have been the standard model, but this may no longer be in production. The new spec may be a model with height adjustment. The cost of the option is, therefore, known from before, and a continuing series can be developed by using equation (7.20) and simply adding the option cost back into the base period, old price. Even this process may have its problems. First, the cost of producing something as standard, since all new chairs now have the height adjuster, may be lower than when it was an option. The option cost method would thus understate a price increase. It may be that the manufacturer has an estimate of the effects of such economies of scale to allow for further adjustments. Triplett (2002) cites a study by Levy and others (1999) in which an automobile antitheft system was installed as standard, but disabled when the option was not required. It was seemingly cheaper to produce this way. Second, by including something as standard, the revenue received may be less for some sales than the marginal cost of producing it. The decision to include it as standard precludes buyers from refusing it. It may be that they will turn to other manufacturers who allow them to exclude the option, although it is unlikely that this will be the sole criterion for the purchase. 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 revenue purchasers accord it as standard. Third, the height adjuster may be valued at an additional amount x when sold separately. There is likely to be a segment of the market that particularly values price adjusters and is willing to spend the additional amount. However, when it is sold as standard, many of the purchasers will not value it so highly since these were the very ones who chose the standard chair. The overall user value would be less than x, although it is not immediately apparent how much less. Some statistical offices take one-half x as the adjustment. Some insight into the proportion of the market purchasing the standard products would help generate more precise estimates.

87. Option cost adjustments are similar to the quantity adjustments, with the exception that the additional quality feature of the replacement is not limited to size. The comparison is $p_n^t / \hat{p}_m^{t-1}$, where $\hat{p}_m^{t-1} = p_m^{t-1} + \beta\Delta z$ for an individual z characteristic where $\Delta z = \left(z_n^t - z_m^{t-1}\right)$. The characteristic may be the amount of RAM on a personal computer (PC) as a specific model is replaced by one that is identical except for amount of RAM. If the relationship between price and RAM is linear, this formulation is appropriate. On the web pages of many computer manufacturers, the price of additional RAM is independent of other features, and a linear adjustment is appropriate. Bear in mind that a linear formulation values the worth of a fixed additional amount of RAM as the same irrespective of the machine's total amount of RAM.

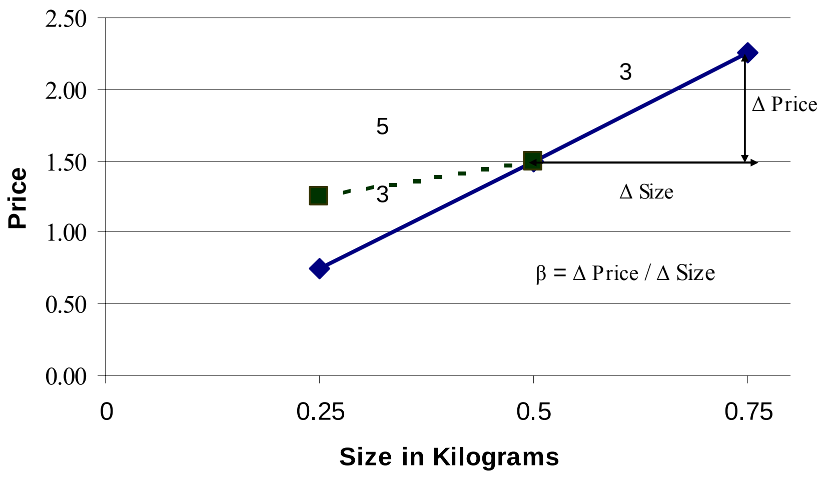
88. The relationship may be nonlinear. For example, for every additional 1 percent of *x*, *y* increases by 1.5 percent $(\beta = 1.015)$, in this case

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

for $p_n^t / \hat{p}_m^{t-1}$ as a measure of quality-adjusted price changes. Again, the *z* change may reflect the service flow, but the nonlinearity in the price–*z* relationship may reflect the increasing or decreasing utility to the scale of the provision. The characteristic may be priced at a higher rate in up-market models of the product versus down-market ones, that is, $\beta \geq 1$ in equation (7.21).

89. The similarity between the quantity adjustment and the option cost approach can be identified by simply considering Figure 7.1 with the *z* characteristic being the option horizontal axis. The similarity between the quantity adjustment and the option cost approach is apparent because both relate price to some dimension of quality: the size or the option. The option cost approach can be extended to more than one quality dimension. Both approaches rely on the acquisition of estimates of the change in price resulting from a change in the options or size: the β slope estimates. In the case of the quantity adjustment, this is taken from a product identical to the one being replaced except for the size. The β slope estimate in this case would be perfectly identified from the two pieces of data. It is as if changes in the other factors' quality were accounted for by the nature of the experiment; this is done by comparing prices of what is essentially the same thing except for change in quantity. There may be, for example, two items that are identical except for of a single feature. This allows the value of the feature to be determined. Yet sometimes the worth 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 one feature varying in a product: processor speed in a PC. It is not a straightforward matter to determine the value of an additional unit of speed. To complicate matters, there may be several quality dimensions to the items, and not all combinations of these may exist as items in the market in any one period. Furthermore, the combinations existing in the second period being compared may be quite different from those in the first. All of this leads to a more general framework.

Figure 7.1. Quality Adjustment for Different-Sized Items



E.4 Hedonic approach

E.4.1 Principles and method

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

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

92. Appendix 7.1 provides data taken from the U.K. Compaq and Dell websites in July 2000 on the prices and characteristics of 64 desktop PCs. Figure 7.2 is a scatter diagram constructed from these data relating the price (£) to the processing speed (MHz). It is apparent that PCs with higher speeds command higher prices—a positive relationship. Under the option cost framework described above, a switch from a 733 MHz PC to a 933 MHz one would involve a measure of the slope of the line between two unique points. The approach requires that there are 733 MHz and 933 MHz PCs that are otherwise identical. From Figure 7.2 and Appendix 7.1, it is apparent that in each instance there are several PCs with the same speed but different prices, owing to differences in other things. To estimate the required value given to additional units of speed, an estimate of the slope of the line that best fits the data is required. In Figure 7.1, the actual slope was used; for the data in Figure 7.2, an estimate of the slope needs to be de- rived from an estimate of the equation of the line that best fits the data, using ordinary least squares

Figure 7.2. Scatter Diagram of PC Prices

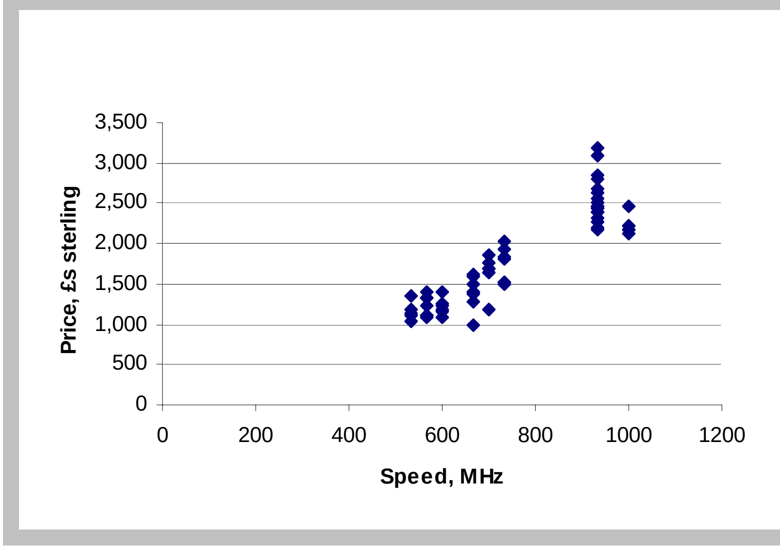


Table 7.4. Hedonic Regression Results for Dell and Compaq PCs

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

† Read-only memory.
Figures in brackets are *t*-statistics testing a null hypothesis of the coefficient being zero.
***, **, and * denote statistically significant at a 0.1 percent, 1 percent, and 5 percent level, respectively, tests being one-tailed.

(OLS) regression. Facilities for regression are available on standard statistical and econometric software, as well as spreadsheets. The estimated (linear) equation in this instance is

$$(7.22)\quad \hat{P}\text{rice} = -658.436 + 3.261\ \text{Speed}$$

$$\bar{R}^2 = 0.820 .$$

The coefficient on speed is the estimated slope of the line: the change in price (£3.261) resulting from a 1 MHz change in speed. This can be used to estimate quality-adjusted price changes for PCs of different speeds. The $\bar{R}^2$ finds that 82 percent of price variation is explained by variation in processing speed. A *t*-statistic to test the null hypothesis of the coefficient being zero was found to be 18.83; recourse to standard tables on *t*-statistics found the null hypothesis was rejected at a 1 percent level. The fact that the estimated coefficient differs from zero cannot be attributed to sampling errors at this level of significance. There is a probability of less than 1 percent that the test has wrongly rejected the null hypothesis. However, the range of prices for a given speed—933 MHz, for example—can be seen from Appendix 7.1 to be substantial. There is a price range of about £1,000, which suggests other quality characteristics may be involved. Table 7.4 provides the results of a regression equation that relates price to a number of quality characteristics using the data in Appendix 7.1. Such estimates can be provided by standard statistical and econometric software, as well as spreadsheets.

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

94. The brand variables are dummy intercepts taking values of 1 if, for example, it is a Dell computer and zero otherwise. While brands are not in themselves quality characteristics, they may be proxy variables for other factors such as after-service reliability. The inclusion of such brand dummies also reflects segmented markets as communities of buyers as discussed in Chapter 21, Appendix 21.1. Similar dummy variables were formed for other makes and models, including the Compaq Presario and Compaq Prosignia. The Compaq Deskpro, however, was omitted to form the benchmark against which other models are compared. The coefficient on Dell is an estimate of the difference between the worth of a Dell and a Compaq Deskpro, other variables being constant (that is, £1,330.78 cheaper). Similarly, an Intel Pentium III commands a premium estimated at £282.78 over an AMD Athlon.

95. The estimate for processor speed was based on data for Dell and Compaq PCs. If the adjustment for quality is between two Dell PCs, it might be argued that data on Compaq PCs should be ignored. Separate regressions could be estimated for each make, but this would severely restrict the sample size. Alternatively, an interaction term or slope dummy can be used for variables that are believed to have a distinctive brand-interaction effect. Take Dell × Speed, which takes the value of speed when the PC is a Dell and zero otherwise. The coefficient on this variable is 1.714 (see Table 7.4); it is an estimate of the additional (positive sign) price arising for a Dell PC over and above that already arising from the standard valuation of a 1 MHz increase in speed. For Dell PCs, it is 2.731 + 1.714 = £4.445. Therefore, if the replacement Dell PC is 200 MHz faster than the unavailable PC, the price adjustment to the unavailable PC is to add 200 × £4.445 = £889. Interactive terms for other variables can similarly be defined and used. The estimation of regression equations is easily undertaken using econometric or statistical software, or data analysis functions in spreadsheets. An understanding of the techniques is given

in many texts, including Kennedy (2003) and Maddala (1988). In Chapter 21, Appendix 21.1, econometric concerns particular to the estimation of hedonic regressions are discussed.

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

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

(7.23) $\text{Price} = \beta_0 \times \beta_1^{z_1} \times \beta_2^{z_2} \times \beta_3^{z_3} \times .... \times \beta_n^{z_n} \times \varepsilon$ or

$$\begin{aligned}\ln \text{Price} = \ln \beta_0 + z_1\beta_1 + z_2\beta_2 + z_3\beta_3 + ...\\ + z_n\beta_n + \ln\varepsilon.\end{aligned}$$

Note that this is a semi-logarithmic form; logarithms are taken of only the left-hand side variable, that is, price. Each of the *z* characteristics enter the regression without having logarithms taken. This has the advantage of allowing dummy variables for the possession or otherwise of a feature to be included on the right-hand side. Such dummy variables take the value of 1 if the product possesses the feature and zero otherwise, it not being possible to take a logarithm of the value zero. Issues on choice of functional form are discussed in more detail in Chapter 21.

98. The taking of logarithms in the first equation (7.23) allows it to be transformed in the second equation to a linear form. This allows the use of a conventional OLS estimator to yield estimates of the logarithm of the coefficients. These are given in column 3 of Table 7.4 and have a useful direct interpretation: if these coefficients are multiplied by 100, they are the percentage change in price arising from a 1-unit change in the explanatory variable. For processor speed, there is an estimated 0.1364 percent change in price for each additional MHz the replacement product has over and above the unavailable one. When dummy variables are used, the coefficients—when multiplied by 100—are estimates of the percentage change in price given by $\left(e^{\beta} - 1\right) 100$; for example, for a rewritable CD drive (CD-RW) compared with a read-only CD drive (CD-ROM), it is $\left(e^{0.08916} - 1\right)100 = 9.326$ percent. There is some bias in these estimated coefficients on dummy variables for the (semi-) logarithmic equation; one-half of the variance of the regression equation should be added to the coefficient before using it (Teekens and Koerts, 1972). For CD-ROM, the t-statistic is 2.88; this is equal to the coefficient divided by its standard error. The standard error is 0.08916/2.88 = 0.03096, and the variance is $0.03096^2 = 0.000958$. To adjust to variance of the regression equation, add 0.000958/2 to 0.08916 = 0.089639 or 8.9639 percent.

99. The approach is particularly useful when the market does not reveal the price of the quality characteristics required for the adjustment. Markets reveal prices of products, not quality characteristics, so it is useful to consider products as tied bundles of characteristics. A sufficiently large data set of products with their characteristics and sufficient variability in the mix of characteristics between the products allows the hedonic regression to provide estimates of the implicit prices of the characteristics. The formal theory is provided in Chapter 21. There are a number of ways of implementing the method, which are outlined below. Before doing so, it is useful to note how these coefficients should be interpreted in light of theoretical needs.

E.4.2 On theory

100. Some mention should be made of the interpretation of the coefficients from hedonic regressions. The matter will be discussed in further detail in Chapter 21, Section B.5. This section summarizes the conclusion. There used to be an erroneous perception that the coefficients from hedonic methods represented estimates of user value as opposed to resource cost. Rosen (1974) showed that hedonic coefficients may be reflective of both user value and resource cost, both supply and demand influences. There is, in econometrics terms, an identification problem, in which the observed data do not permit the estimation of the underlying demand-and-supply parameters. However, suppose the *production technology of sellers is the same* but buyers differ. Then the hedonic function describes the prices of characteristics the firm will supply with the given ruling technology to the current mixture of tastes. There are different tastes on the user side, so what appears in the market is the result of firms trying to satisfy purchaser's preferences all for a constant technology and profit level; the structure of supply is revealed by the hedonic price function. Now suppose sellers differ but *buyers' tastes are the same*. Here the hedonic function $p(z)$ identifies the structure of demand. Of these possibilities, uniformity of tastes is unlikely while uniformity of technologies is more likely, especially when

access to technology is unrestricted in the long run. Griliches (1988, p. 120) has argued in the context of a CPI:

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

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

101. The proper use of hedonic regression requires an examination of the coefficients of the estimated equations to see if they make sense. It might be argued that the very multitude of distributions of tastes and technologies and interplay of supply and demand make it unlikely that *reasonable* estimates will arise from such regressions. A firm may apply and cut a profit margin and prices for reasons related to long-run strategic plans, for example, yielding coefficients that prima facie do not look reasonable. This does not negate the usefulness of examining hedonic coefficients as part of a strategy for evaluating estimated hedonic equations. First, there has been extensive empirical work in this field, and the results for individual coefficients are, for the most part, quite reasonable. Even over time, individual coefficients can show quite sensible patterns of decline (van Mulligen, 2003). Second, as shall be seen, it might be argued that the prediction and its error should be our concern and not the values of individual coefficients (Pakes, 2001).

E.4.3 Implementation

102. The implementation of hedonic methods to estimate quality adjustments to noncomparable replacements can take a number of forms. The first form is when the repricing is for a product with different characteristics. What is required is to adjust either the price of the old or replacement (new) product for some valuation of the difference in quality between the two products. This patching of missing prices is quite different from the use of hedonic price indices to be discussed in Section 7.G.2 and in Chapter 21. These use hedonic regressions to provide hedonic price indices of overall quality-adjusted prices. The former is a partial application, used on noncomparable replacements when products are no longer available. The latter, as will be seen in Section 7.G.2, is a general application to a sample from the whole data set. The partial patching is considered here.

103. Hedonic imputation: *predicted vs. actual*—In this approach, a hedonic regression of the (natural logarithm of the) price of model *i* in period *t* on its characteristics set z_{ki}^t is estimated *for each month*, as given by

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

Let us say the price of a product *m* available in January (period *t*) is unavailable in March (period *t* + 2). The price of product *m* can be predicted for March by inserting the characteristics of the old unavailable product *m* into the estimated regression equation for March; this process is repeated for successive months. The predicted price for the old product in March and the price comparison with January (period *t*) are given, respectively, by

$$(7.25a)\ \hat{p}_m^{t+2} \equiv \exp[\beta_0^{t+2} + \sum_{k=1}^{K} \beta_k^{t+2} z_{mk}^t]$$

and $\hat{p}_m^{t+2} / p_m^t$, that is, the *old* model's price is adjusted. In the example in Table 7.2(a), $\hat{p}_2^3$, $\hat{p}_2^4$, etcetera and $\hat{p}_6^3$, $\hat{p}_6^4$, etc. would be estimated and compared with p_2^1 and p_6^1, respectively. The blanks for products 2 and 6 in Table 7.2(a) would be effectively filled in by the estimated price from the regression equation.

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

$$(7.25b)\ \hat{p}_m^t \equiv \exp[\beta_0^t + \sum_{k=1}^{K} \beta_k^t z_{mk}^{t+2}],$$

and $p_n^{t+2} / \hat{p}_n^t$, that is, the *new* model's price is adjusted. In this case, the characteristics of product *n* are inserted into the right-hand side of an estimated regression for period *t*. The price comparisons of equation (7.25a) may be weighted by w_m^t, as would those of its replaced price comparison in equation (7.25b).

105. A final alternative is to take the geometric mean of the formulations in equations (7.25a) and (7.25b) on grounds analogous to those discussed in Chapter 15 and by Diewert (1997) for similar index number issues.

106. Hedonic imputation: *predicted vs. predicted*—A further approach is the use of predicted values for the product in *both* periods, for example, $\hat{p}_n^{t+2} / \hat{p}_n^t$, where *n* represents the product. Consider a misspecification problem in the hedonic equation. For example, there may be an interaction effect between a brand dummy and a characteristic, say, between Dell and speed in the example in Table 7.4. Having both characteristics may be worth more on price (from a semi-logarithmic form) than their separate individual components (for evidence of interaction effects see, Curry, Morgan, and Silver, 2000). The use of $p_n^{t+2} / \hat{p}_n^t$ would be misleading since the actual price in the numerator would incorporate the 5 percent premium while the one predicted from a straightforward semi-logarithmic form would not. It is stressed that in adopting this

approach, a recorded, actual price is being replaced by an imputation. Neither this nor the form of bias discussed above are desirable. Diewert (2002e) considers a similar problem and suggests an adjustment to bring the actual price back in line with the hedonic one.

107. The comparisons using predicted values in both periods are given as

$\hat{p}_n^{t+2} / \hat{p}_n^t$ for the *new* product,

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

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

as a (geometric) mean of the two.

108. Hedonic adjustments using *coefficients*—In this approach, a replacement product is used and any differences between the characteristics of the replacement *n* in period $t + 2$ and *m* in period *t* are ascertained. A predicted price for *n* in period *t*, that is, $\hat{p}_n^t$, is compared with the actual price p_n^{t+2}. However, unlike the formulation in equation (7.25b) for example, $\hat{p}_n^t$ may be estimated by applying the subset of the *k* characteristics that distinguished *m* from *n,* to their respective implicit prices in period *t* estimated from the hedonic regression, and adjusting the price of p_m^t. For example, if the nearest replacement for product 2 was product 3, then the characteristics that differentiated product 3 from product 2 are identified and the price in the base period p_3^1 is estimated by adjusting p_2^1 using the appropriate coefficients from the hedonic regression in that month. For example, for washing machines, if product 2 had an 800 revolutions per minute (rpm) spin speed and product 3 had an 1,100 rpm spin speed, other things being equal, the shadow price of the 300 rpm differential would be estimated from the hedonic regression, and p_2^1 would be adjusted for comparison with p_3^3. Note that if the *z* variables in the characteristic set are perfectly independent of each other, the results from this approach will be similar to those from equation (7.25b). This is because interdependence among the variables on the right-hand side of the hedonic equation—multicollinearity—leads to imprecise estimates of the coefficients (see Chapter 21, Appendix 21.1).

109. Hedonic *indirect adjustment*—An indirect current period hedonic adjustment may be used, which only requires the hedonic regression to be estimated in the base period *t*.

(7.27) $\dfrac{p_n^{t+2}}{p_m^t} \div \dfrac{\hat{p}_n^t}{\hat{p}_m^t}$.

The first term is the change in price between the old and replacement items in periods *t* and $t + 2$, respectively. But the quality of the product has changed, so this price change needs to be divided by a measure of the change in quality. The second term uses the hedonic regression in period *t* in both the numerator and denominator. So the coefficients—the shadow prices of each characteristic—remain the same. It is not prices that change. The predicted prices differ because different *quantities* of the characteristics are being inserted into the numerator and denominator; the replacement *n* characteristics in the former and old product *m* characteristics in the latter. The measure is the change in price after removing (by division) the change in the quantity of characteristics each valued at a constant period t price. Conceptually, the constant valuation by a period $t + 2$ regression would be equally valid and a geometric mean of the two ideal. However, if hedonic regressions cannot be run in real time, equation (7.27) is a compromise. As the spread between the current and base-period results increases, its validity decreases. As such, the regression estimates should be updated regularly using old- and current-period estimates, and results compared retrospectively as a check on the validity of the results.

E.4.4 Need for caution

110. The limitations of the hedonic approach should be kept in mind. Some points are summarized below though readers are referred to the Bibliography and to Chapter 21, Appendix 21.1. First, the approach requires statistical expertise for the estimation of the equations. The prevalence of user-friendly software with regression capabilities makes this less problematic. 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; F-test and t-test statistics to enable tests to be conducted as to determine whether the differences between the coefficients on the explanatory variables are jointly and individually different from zero at specified levels of statistical significance. Most of these statistics make use of the errors from the estimated equation. The regression equation can be used to predict prices for each product by inserting the values of the characteristics of the products into the explanatory variables. The differences between the actual prices and these predicted results are the residual errors. Biased or imprecise results may arise from a range of factors, including heteroskedasticity (nonconstant variances in the residuals suggesting nonlinearities or omission of relevant explanatory variables), a nonnormal distribution for the errors, and multicollinearity, where two or more explanatory variables are related. The latter, in particular, has been described as the "bane of hedonic regressions..." (Triplett, 1990). Such econometric issues are well discussed in the context of hedonic regressions (Berndt, 1991; Berndt, Griliches, and Rappaport, 1995; Triplett, 1990; Gordon, 1990; Silver, 1999; and Chapter 21, Appendix 21.1) and more generally in introductory econometric texts such as Kennedy (2003) and Maddala (1988). The use of predicted values when multicollinearity is suspected is advised, rather than individual coefficients, for reasons discussed above.

111. Second, the estimated coefficients should be updated regularly. However, if the adjustment is to the old model, then the price comparison is between the price of the new model and the quality-adjusted price of the old model. The quality difference between the old and new model is derived using coefficients from a hedonic regression from a previous period as estimates of the value of such differences. There is, at first glance, no need to update the hedonic regression each month. The valuation of a characteristic in the price reference period may, however, be quite out of line with its valuation in the new period. For example, a feature may be worth an additional 5 percent in the reference period instead of 10 percent in the current period because it might have been introduced at a discount at that point in its life cycle to encourage usage. Continuing to use the coefficients from some far-off period to make price adjustments in the current period is similar to using out-of-date base-period weights. The comparison may be well defined but have little meaning. If price adjustments for quality differences are being made to the old item in the price reference period using hedonic estimates from that period, then there is a need to update the estimates if they are considered out of date, for example, due to changing tastes or technology, and splice the new estimated comparisons onto the old. Therefore, regular

updating of hedonic estimates when using the adjustments to the old price is recommended, especially when there is evidence of parameter instability over time.

112. Third, the sample of prices and characteristics used for the hedonic adjustments should be suitable for the purpose. If they are taken from a particular industry, trade source, or web page and then used to adjust noncomparable prices for products sold in quite different markets, then there must at least be an intuition that the marginal returns for characteristics are similar among the industries. A similar principle applies for the brands of products used in the sample for the hedonic regression. It should be kept in mind that high $\bar{R}^2$ statistics do not alone ensure reliable results. Such high values arise from regressions in periods before their application and inform us of the proportion of variation in prices across many products and brands. They are not in themselves a measure of the prediction error for a particular product, sold by a specific outlet of a given brand in a subsequent period, although they can be an important constituent of this.

113. Fourth, there is the issue of functional form and the choice of variables to include in the model. Simple functional forms generally work well. These include linear, semi-logarithmic (logarithm of the left-hand side), and double-log (logarithms of both sides) forms. Such issues are discussed in Chapter 21, Appendix 21.1. The specification of a model should include all price-determining characteristics. Some authors advise quite simple forms with only the minimum number of variables, as long as the predictive capacity is high (Koskimäki and Vartia, 2001). For the CPI, Shepler (2000) included 33 variables in her hedonic regressions of refrigerators, a fairly homogeneous product. These included 9 dummy variables for brand, 4 dummy variables for color, 5 types of outlets, 3 regions as control variables, and 11 characteristics. These characteristics included capacity, type of ice-maker, energy-saving control, number of extra drawers, sound insulation, humidifier, and filtration device. Typically, a study would start with a larger number of explanatory variables and a general econometric model of the relationship; the final model is a more specific, parsimonious one since it has dropped a number of variables. The dropping of variables occurs after experimenting with different formulations and seeing their effects on diagnostic test statistics, including the overall fit of the model and the accordance of signs and magnitudes of coefficients with prior expectations. Reese (2000), for example, started with a hedonic regression for U.S. college textbooks. It included about 50 explanatory variables; subsequently, those variables were reduced to 14 with little loss of explanatory power.

114. Finally, Bascher and Lacroix (1999) list several requirements for successful design and use of hedonic quality adjustment in the CPI, noting that these requirements require heavy investments over a long period. They involve (i) intellectual competencies and sufficient time to develop and reestimate the model and employ it when products are replaced; (ii) access to detailed, reliable information on product characteristics; and (iii) a suitable organization of the infrastructure for collecting, checking, and processing information.

115. It should be noted that hedonic methods may also improve quality adjustment in CPIs by indicating which product attributes do *not* appear to have material impacts on price. That is, if a replacement product differs from the old product only in characteristics that have been rejected as price-determining variables in a hedonic study, this would support a decision to treat the products as comparable or equivalent and include the entire price difference (if any) as pure price change. Care has to be exercised in such analysis because a feature of multicollinearity in

regression estimates is that the imprecision of the parameter estimates may give rise to statistical tests that do not reject null hypotheses that are false, that is, they do not find significant parameter estimates. However, the results from such regressions can nonetheless provide valuable information on the extent to which different characteristics influence price variation. This in turn can help in the selection of replacement products. The enhanced confidence in product substitution and the quality adjustment of prices from the hedonic approach with its parallel reduction in reliance on linking have been cited as significant benefits in the reliability of the measurement of price changes for apparel in the United States' CPI (Reinsdorf, Liegey, and Stewart, 1996). The results from hedonic regressions have a role to play in identifying price-determining characteristics and may be useful in the design of quality checklists in price collection (Chapter 6).

F. Choosing a Quality Adjustment Method

116. Choosing a method for quality-adjusting prices is not straightforward. The analyst must consider the technology and market for each product group and devise appropriate methods. This is not to say the methods selected for one product group will be independent of those selected for other industries. Expertise built up using one method may encourage its use elsewhere, and intensive use of resources for one product group may lead to less resource-intensive methods in others. The methods adopted for individual industries may vary among countries as access to data, relationships with the price collectors, resources, expertise and features of the production, and market for the product vary. Guidelines on choosing a method arise directly from the features of the methods outlined above. A good understanding of the methods and their implicit and explicit assumptions is essential when choosing a method.

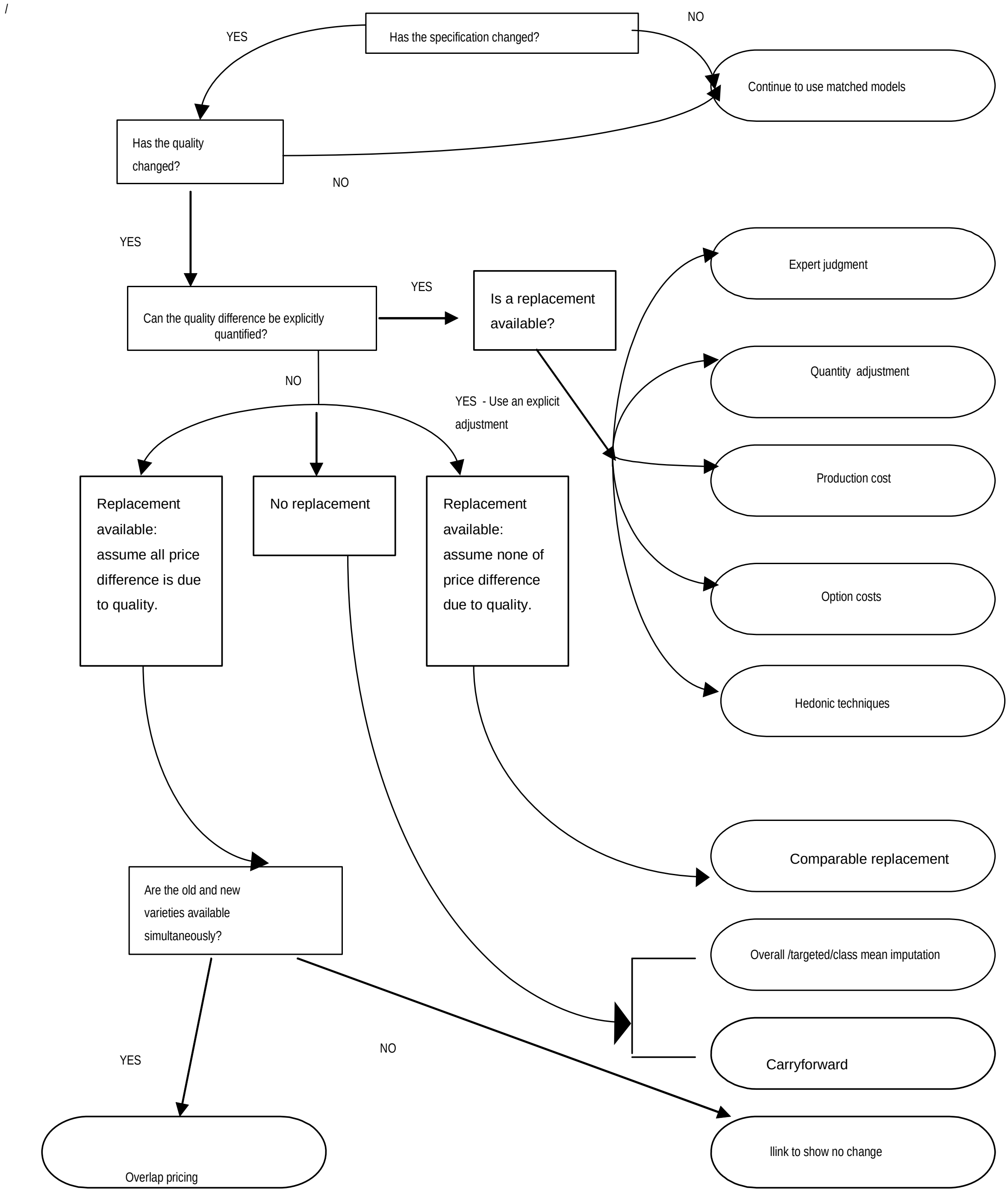
117. Consider Figure 7.3, which provides a useful guide to the decision-making process. Assume the matched-models method is being used. If the product is matched for repricing—without a change in the specification—no quality adjustment is required. This is the simplest of procedures. However, a caveat applies. If the product belongs to a high-technology product market where model replacement is rapid, the matched sample may become unrepresentative of the universe of transactions. Alternatively, matching may be under a chained framework, where prices of products in a period are matched to those in the preceding period to form a link. A series of successive links of matched comparisons combined by successive multiplication makes up the chained matched index. Alternatively, hedonic indices may be used, which require no matching. The use of such methods is discussed in Section G. At the very least, attention should be directed to more regular product resampling. Continued long-run matching would deplete the sample, and an alternative framework to long-run matching would be required.

118. Consider a change in the quality of a product, and assume a replacement product is available. The selection of a comparable product to the same specification and the use of its price as a *comparable replacement* require that none of the price difference is due to quality. They also require confidence that all price-determining factors are included on the specification. The replacement product should also be representative and account for a reasonable proportion of sales. Caution is required when nearly obsolete products at the end of their life cycles are replaced with unusual pricing by similar products that account for relatively low sales, or with products that have substantial sales but are at different points in their cycle. Strategies for

ameliorating such effects are discussed below and in Chapter 8, including early substitutions before pricing strategies become dissimilar.

119. Figure 7.3 shows where quality differences can be quantified. *Explicit estimates* are generally considered to be more reliable, but they are also more resource intensive (at least initially). Once an appropriate methodology has been developed, explicit estimates can often be easily replicated. General guidelines are more difficult here since the choice depends on the host of factors discussed above, which are likely to make the estimates more reliable in each situation. Central to all of this is the quality of the data on which the estimates are based. If reliable data are unavailable, subjective judgments may be used. Product differences are often quite technical and very difficult to specify and quantify. The reliability of the method depends on the knowledge of the experts and the variance in opinions. Estimates based on objective data are, as a result, preferred. Good *production cost* estimates, along with good data on markups

Figure 7.3. Flowchart for Making Decisions on Quality Change



Source: Chart based on work of Fenella Maitland-Smith and Rachel Bevan, OECD; see also a version in Triplett (2002).

and indirect taxes, where applicable, in industries with stable technologies where differences between the old and replacement products are well specified and exhaustive, are reliable by definition. The *option cost* approach is generally preferable when old and new products differ by easily identifiable characteristics that have once been separately priced as options, but the price of an option will overstate its value when it becomes standard so care must be taken when using this method. The use of *hedonic regressions* for partial patching is most appropriate where data on price and characteristics are available for a range of models and where the characteristics are found to predict and explain price variability well in terms of a priori reasoning and econometrics. Use of hedonic regressions is appropriate where the price of an option or change in characteristics cannot be separately identified and has to be gleaned from the prices of products 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.

120. The estimates are particularly useful for valuing changes in the quality of a product when only a given set of characteristics change, 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 product selection. The synergy between the selection of prices according to characteristics defined as price determining by the hedonic regression and the subsequent use of hedonics for quality adjustment should reap rewards. The method should be applied where there are high ratios of noncomparable replacements and where the differences between the old and new products can be well defined by a large number of characteristics.

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

122. Bias from using imputation is directly related to the proportion of missing products and the difference between quality-adjusted prices of available matched products and the quality-adjusted prices of unavailable ones (see Table 7.3). The nature and extent of the bias depends on whether short-run or long-run imputations are being used (the former being preferred) and on market conditions (see Section H). Imputation in practical terms produces the same result as deletion of the product, and the inclusion of imputed prices may give the illusion of larger sample sizes. Imputation is less likely to give bias for products where the proportion of missing prices is low. Table 7.2 can be used to estimate likely error margins arising from its use, and a judgment can be made as to whether they are acceptable. Its use across many industries need not compound the errors since, as noted in the discussion of this method, the direction of bias need not be systematic. It is cost-effective for industries with large numbers of missing products because of its ease of use. But the underlying assumptions required must be carefully considered if widely used. Imputation should by no means be the overall, catchall strategy, and statistical agencies are advised against its use as a default device without due consideration to the nature of

the markets, possibility of targeting the imputation, and the viability of estimates from the sample sizes involved if such targeting is employed.

123. If the old and replacement products are available simultaneously and the quality difference cannot be quantified, an implicit approach can be used whereby the price difference between the old and replacement product in a period in which they both exist is assumed to be due to quality. This *overlap* method, by replacing the old product with 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 products are taken. The assumption of relative prices equating to quality differences at the time of the splice is unlikely to hold true if the old and replacement products 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 product to clear inventories and price skimming of market segments that will purchase new models at relatively high prices. As with comparable replacements, early substitutions are advised so that the overlap is at a time when products are at similar stages in their life cycles.

124. The use of the *linked to show no change* method and the *carryfoward* method is not generally advised for making quality adjustment imputations for the reasons discussed unless there is deemed to be some validity to the implicit assumptions.

G. High-Technology and Other Sectors with Rapid Turnover of Models

125. The measurement of price changes of products unaffected by quality changes is primarily achieved by matching models, the aforementioned techniques being applicable when the matching breaks down. But what about 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 depleted sample. There is both a dynamic universe of all products sold and a static universe of the products selected for repricing (Dalén, 1998). For example, if the sample is initiated in December, by the subsequent May the static universe will be matching prices of those products available in the static universe in both December and May but will omit the unmatched new products introduced in January, February, March, April, and May, and the unmatched old ones available in December but unavailable in May. There are two empirical questions to answer for any significant bias to be detected. First, whether the sample depletion is substantial; such depletion is a necessary condition for bias. Second, whether the unmatched new and unmatched old products are likely to have different quality-adjusted prices versus the matched ones in the current and base period.

126. Thus, the matching of prices of identical models over time may lead to the monitoring of a sample of models increasingly unrepresentative of the population of transactions. There are old models that existed when the sample was drawn but are not available in the current period, and there are new ones coming into the current period that are not available in the base period. It may be that the departures have relatively low prices and the entrants relatively high ones and that by ignoring these prices a bias is being introduced. Using old low-priced products and ignoring new high-priced ones has the effect of biasing the index downward. In some industries, the new product may be introduced at a relatively low price and the old one may become obsolete at a relatively high one, serving a minority segment of the market (Berndt, Ling, and

Kyle, 2003). In this case, the bias would take the opposite direction; the nature of the bias depends on the pricing strategies of firms for new and old products.

127. This sampling bias exists for most products. However, our concern is with product markets where the statistical agencies are finding the frequency of new product introductions and old product obsolescence sufficiently high that they may have little confidence in their results. First, some examples of such product markets will be given. Then, two procedures will be considered: the use of hedonic price indices instead of partial hedonic patching and chaining.

G.1 Some examples

128. Koskimäki and Vartia (2001) attempted to match prices of personal computers over three two-month periods (spring, summer, and fall) using a sample of prices collected as part of the standard price collection for the Finnish CPI, which has some similarities to trade price indexes. Of the 83 spring prices, only 55 matched comparisons could be made with the summer prices, and of those, only 16 continued through to the fall. They noted that the sample of matched pairs became increasingly biased: of the 79 models in the fall, the 16 matched ones had a mean processor speed of 518 MHz compared with 628 MHz for the remaining 63 unmatched ones; the respective hard disk sizes were 10.2 gigabytes (GB) and 15.0 GB; and the percentages of high-end processors (Pentium III and AMD Athlon) were 25 percent and 49.2 percent, respectively. Hardly any change in *matched* prices was found over this six-month period, while a hedonic regression analysis using all of the data found quality-adjusted price falls of around 10 percent. Instructions to price collectors to hold on to models until forced replacements are required may lead to a sample increasingly unrepresentative of the population and be biased toward technically inferior variants. In this instance, the hedonic price changes fell faster since the newer models became cheaper for the services supplied.

129. Kokoski, Moulton, and Zieschang (1999) used hedonic regressions in an empirical study of interarea price comparisons of food products across U.S. urban areas using U.S. CPI data. They found a negative sign on the coefficients of dummy variables for whether the sample products were from newly rotated samples, where the dummy variable = 1, or samples before rotation, where the dummy variable = 0. This indicated that quality-adjusted prices were lower for the newly included products compared with the quality-adjusted prices of the old products.

130. Silver and Heravi (2002) found evidence of sample degradation when matching prices of U.K. washing machines over a year. By December, only 53 percent of the January basket of model varieties was used for the December/January index, although this accounted for 81.6 percent of January expenditure. Models of washing machines with lower sales values dropped out quicker. However, the remaining models in December accounted for only 48.2 percent of the value of transactions *in December*. The active sample relating to the universe of transactions in December had substantially deteriorated. The prices of unmatched and matched models differed, as did their vintage and quality. Even when prices were adjusted for quality using hedonic regressions, prices of unmatched old models were found to be lower than matched ones; there was also evidence of higher prices for unmatched new models. Quality-adjusted prices fell faster for the matched sample than the full sample: about 10 percent for the former compared with about 7 percent for the latter. Residuals from a common hedonic surface and their leverage were also examined. The residuals from unmatched new models were higher than matched ones, while residuals from unmatched old models were much lower. Unmatched observations had nearly

twice the (unweighted) leverage than matched ones; their influence in the estimation of the parameters of the regression equation was much greater and their exclusion more serious.

131. These studies demonstrate how serious sample degradation can occur and how unmatched excluded products may be quite different from included ones. Two procedures for dealing with such situations will be considered: the use of hedonic price indices instead of the partial hedonic patching discussed above and chaining. Both rely on a data set of a representative sample of products and their characteristics *in each period*. A checklist of structured product characteristics to be completed each reporting period is one way changes in quality characteristics can be prompted and monitored: this is especially useful in high-technology industries (Merkel, 2000). If a new product is introduced and has or is likely to have substantial sales, then it is included as a replacement or even an addition. Its characteristics are marked off against a checklist of salient characteristics. The list will be developed when the sample is initiated and updated as required. Alternatively, web pages and trade associations may be able to provide lists of models and their prices; however, the need for transaction prices as opposed to list prices is stressed.

G.2 Hedonic price indices

132. It is important to distinguish between the use of hedonic regressions to make adjustments for quality differences when a noncomparable substitute is used, as in Section E, and their use in their own right as *hedonic price indices,* which are measures of quality-adjusted price changes. Hedonic price indices are suitable when the pace and scale of replacements of products are substantial. There are two reasons for this. First, an extensive use of quality adjustments may lead to errors. Second, the sampling will be from a matched or replacement universe likely to be biased. With new models being continually introduced and old ones dying, the coverage of a matched sample may deteriorate and bias may be introduced as the price changes of the new or old models differ from those of the matched ones. A sample must be drawn in each month, and price indices must be constructed, but, instead of being controlled for quality differences by matching, they will be controlled for, or partialed out, in the hedonic regression. Note that all the indices described below use a fresh sample of the data available in each period. If there is a new product in a period, it is included in the data set and its quality differences controlled for by the regression. Similarly, if old products drop out, they are still included in the data for the indices in the periods in which they exist. In Section E.4.4 of this chapter, the need for caution was stressed in the use of hedonic regressions for quality adjustments due to theoretical and econometric issues, some of which will be considered in the appendix to Chapter 21. This need for caution extends to the use of the results from hedonic indices and is not repeated here for the sake of brevity.

133. In Chapter 17, theoretical price indices will be defined and practical index number formulas considered as bounds or estimates of these indices. Theoretical index numbers will also be defined in Chapter 21 to include goods made up of tied characteristics, so that something can be said about how such theoretical indices relate to different forms of hedonic indices. A number of forms will be considered in Chapter 21, and the account is outlined here.

G.2.1 Hedonic functions with dummy variables on time

134. The sample covers the two time periods being compared—for example, t and $t + 2$—and does not have to be matched. The hedonic formulation regresses the price of product i, p_i, on the $k = 2 \ldots K$ characteristics of the products z_{ki}. A single regression is estimated on the data in the two time periods compared, the equation also including a dummy variable D^{t+2} being 1 in period $t + 2$, zero otherwise:

$$(7.28)\ \ln p_i = \beta_0 + \beta_1 D^{t+2} + \sum_{k=2}^{K} \beta_k z_{ik} + \varepsilon_i$$

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

135. The approach just described uses the dummy variables on time to compare prices in period t with prices in each subsequent period. In doing so, the β parameters are constrained to be constant over the period being compared. A fixed-base, bilateral comparison using equation (7.28) makes use of the constrained parameter estimates over the two periods compared and, given an equal number of observations in each period, is a form of a symmetric average. A *chained* formulation would estimate an index between periods 1 and 4—represented here as $I^{1,4}$—as

$$I^{1,4} = I^{1,2} \times I^{2,3} \times I^{3,4}.$$

136. There is no explicit weighting in these formulations; this is a serious disadvantage. In practice. cutoff sampling might be employed to include only the most important products. If sales data are available, a weighted least squares estimator (WLS) should be used, as opposed to an OLS estimator. It is axiomatic in normal index number construction that the same weight should not be given to each price comparison since some products may account for much larger sales revenues than others. The same consideration applies to these hedonic indices. Diewert (2002e) has argued that sales *values* should form the basis of the weights over quantities. Two products may have sales equal to the same quantity, but, if one is priced higher than another, its price changes should be weighted higher accordingly for the result to be meaningful in an economic sense. In addition, Diewert (2002e) has shown that it is value *shares* that should form the weights, since values will increase—over period $t + 2$, for example—with prices, the

residuals, and their variance thus being higher in period $t + 2$ than in t. This heteroskedasticity is an undesirable feature of a regression model resulting in increased standard errors. Silver (2002) has further shown that a WLS estimator does not purely weight the observations by their designated weights. The actual influence given is also due to a combination of the residuals and the leverage effect. The latter is higher since the characteristics of the observations diverge from the average characteristics of the data. He suggests that observations with relatively high leverage and low weights be deleted and the regression repeated.

G.2.2 Period-on-period hedonic indices

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

$$(7.29)\ P_{JHB} = \frac{\left[\prod_{i=1}^{N^t} \hat{p}_i^{t+2}(z_i^t)\right]^{1/N^t}}{\left[\prod_{i=1}^{N^t} p_i^t(z_i^t)\right]^{1/N^t}}$$

$$\approx \frac{\left[\prod_{i=1}^{N^t} \hat{p}_i^{t+2}(z_i^t)\right]^{1/N^t}}{\left[\prod_{i=1}^{N^t} \hat{p}_i^t\right]^{1/N^t}} \approx \frac{\left[\prod_{i=1}^{N^t} \hat{p}_i^{t+2}(z_i^t)\right]^{1/N^t}}{\left[\prod_{i=1}^{N^t} p_i^t\right]^{1/N^t}}.$$

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

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

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

139. For a fixed-base, bilateral comparison using either equation (7.29) or (7.30), the hedonic equation is estimated only for one period, the current period *t* + 2 in equation (7.29) and the base period *t* in equation (7.30). For reasons analogous to those explained in Chapters 15, 16, and 17, a symmetric average of these indices would have some theoretical support. It would be useful as a retrospective study to compare the results from both approaches (7.29) and (7.30). If the discrepancy is large, the results from either should be treated with caution, similar to the way a large Laspeyres and Paasche spread would cast doubt on the use of either of these indices individually. It would be evidence for the need to update the regressions more often.

140. Note that a geometric mean of equations (7.29) and (7.30) uses all of the data available in each period, as does the hedonic index using a time dummy variable in (7.28). If in (7.28) there is a new product in period *t* + 2, it is included in the data set and its quality differences controlled for by the regression. Similarly, if old products drop out, they are still included in the indices in the periods in which they exist. This is part of the natural estimation procedure, unlike using matched data and hedonic adjustments on noncomparable replacements when products are no longer available.

141. With the dummy variable approach, there is no explicit weighting in its formulation in equations (7.29) and (7.30), and this is a serious disadvantage. In practice, cutoff sampling might be employed to include only the most important products or if value of output data are available, a WLS—as opposed to OLS—estimator used with value of output shares as weights, as discussed in Chapter 21, Appendix 21.1.

142. The indices ask counterfactual questions. Asking what the price of a model with characteristics *z* would have been if it had been on the market in a period ignores the likelihood that the appearance of that model would in turn alter the demand for other computers, thus altering the coefficients of the hedonic regression as well. The matter is particularly problematic when *backcasting*, that is, using a current period's specification in some previous period's regression as in equations (7.29) and (7.30). If the specifications increase rapidly, it may not be sensible to ask the value of some high-tech model when such technology was in an earlier stage of development. It should be kept in mind that hedonic coefficients may as much reflect production technology as demand (see Chapter 21), and old technologies simply may not have been able to produce goods to the standards of later ones. The question reversed—what would be the value of a previous period's specification in a subsequent period's regression—while subject to similar problems, may be more meaningful. In general, the solution lies in estimating regressions as often as possible, especially in markets subject to rapidly changing technologies.

G.2.3 Superlative and exact hedonic indices (SEHI)

143. In Chapter 17, Laspeyres and Paasche bounds will be defined on a theoretical basis, as will superlative indices, which treat both periods' data symmetrically. These superlative formulas, in particular the Fisher index, are also seen in Chapter 14 to have desirable axiomatic properties. The Fisher index is supported from economic theory as a symmetric average of the Laspeyres and Paasche bounds and was found to be the most suitable such average of the two on axiomatic grounds. The Törnqvist index is shown to be best from the stochastic viewpoint and also does not require strong assumptions for its derivation from the economic approach as a superlative index. The Laspeyres and Paasche indices are found to correspond to (be *exact* for) underlying Leontief aggregator functions with no substitution possibilities, while superlative indices are exact for flexible functional forms including the quadratic and translog forms for the Fisher and Törnqvist indices, respectively.

144. If data on prices, characteristics, *and quantities* are available, analogous approaches and findings arise for hedonic indices (Fixler and Zieschang, 1992a, and Feenstra, 1995). Exact theoretical bounds on a hedonic index have been defined by Feenstra (1995). Consider a theoretical index now defined only over products defined in terms of their characteristics. The prices are still of products, but they are wholly defined through their characteristics *p(z)*. An arithmetic aggregation for a linear hedonic equation finds a Laspeyres upper bound (as quantities demanded *decrease* with increasing relative prices) given by:

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

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

$$(7.31b)\ \hat{p}_i^{t+2} \equiv p_i^{t+2} - \sum_{k=1}^{K} \beta_k^{t+2} (z_{ik}^{t+2} - z_{ik}^t)$$

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

A Paasche lower bound is estimated as:

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

where $s_i^{t+2} = \boldsymbol{q}_i^{t+2}\boldsymbol{p}_i^{t+2} \Big/ \sum_{i=1}^{N} \boldsymbol{q}_i^{t+2}\boldsymbol{p}_i^{t+2}$ and

(7.32b) $\hat{p}_i^t \equiv p_i^t + \sum_{k=1}^{K} \beta_k^t (z_{ik}^{t+2} - z_{ik}^t)$

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

In Chapter 17 it is shown that Laspeyres P_L and Paasche P_P price indices form bounds on their respective 'true' economic theoretic indexes. Using similar reasoning to that in Chapter 17 applied to equations (7.31a) and (7.32a) it can be shown that under homothetic preferences these true economic indices collapse into a single theoretical index $c(p^{t+2})/c(p^t)$, and:

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

145. The approach is similar to that used for adjustments to noncomparable replacement items in equation (7.27). First, the SEHI approach uses all of the data in each period, not just the matched sample and selected replacements. Second, it uses coefficients from hedonic regressions on changes in the characteristics to adjust observed prices for quality changes. Third, it incorporates a weighting system using data on the value of output of each model and their characteristics, rather than treating each model as equally important. Finally, it has a direct correspondence to formulation defined from economic theory.

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

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

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

$$\hat{\boldsymbol{p}}_i^{t+2} \equiv \boldsymbol{p}_i^{t+2} \exp[-\sum_{i=1}^{N} \boldsymbol{\beta}_k^{t+2} (\boldsymbol{z}_{ik}^{t+2} - \boldsymbol{z}_{ik}^t)]$$

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

147. In equation (7.34a), the two bounds on their respective theoretical indices have been shown to be brought together under an assumption of homothetic preference (see Chapter 17). The calculation of such indices is no small task. For examples of its application see Silver and Heravi (2001a and 2003) for comparisons over time and Kokoski, Moulton, and Zieschang (1999) for price comparisons across areas of a country.

148. Note that unlike the hedonic indices in Sections G.2.1 and G.2.2, the indices in equations (7.30b), (7.31b), and (7.33b) need not be based on matched data. Kokoski, Moulton, and Zieschang (1999) used a sample from a replacement universe of otherwise matched data from the U.S. Bureau of Labor Statistics CPI, although the sample benefited from rotation. Silver and Heravi (2001a and 2003) used scanner data for the universe of transactions via a two-stage procedure. First, cells were defined according to major price-determining features much like strata; such features included all combinations of brand, outlet type, and screen size (for television sets). There may be a gain in the efficiency of the final estimate, since the adjustment is for within-strata variation, much in the way that stratified random sampling improves on simple random sampling. The average price in each matched cell could then be used for the price comparisons using equations (7.30a), (7.31a), or (7.33a), except that to ensure that the quality differences in each cell coming from characteristics other than these major ones did not influence the price comparison, adjustments were made for quality changes using equations (7.30b), (7.31b), or (7.33b). This allowed all matched, old unmatched, and new unmatched data to be included. If the average price in a cell of equation (7.30a) was increased because of the inclusion of a new improved product, equation (7.30b) would be used to remove such improvements, on average. For example, consider a brand *X*, 14-inch television set without stereo sound in a given elementary aggregate product group. In the next period, there may be matched cells: 14-inch television set for brand *X*, which also includes stereo. The new model may have to be grouped in the same cell with the brand *X*, 14-inch television sets with and without stereo and the average price of the cells compared in equations (7.30a), (7.31a), or (7.33a), with a quality adjustment for the stereo of the form undertaken by equations (7.30b), (7.31b), or (7.33b). There may be a gain in the efficiency of the final estimate, since the adjustment is for within-strata variation, much in the way that stratified random sampling improves on simple random sampling. The estimated coefficient for stereo would be derived from a hedonic equation estimated from data of other television sets, some of which possess stereo.

149. The description above illustrates how weighted index number formulas such as Laspeyres, Paasche, Fisher, and Törnqvist might be constructed using data on prices, quantities, and characteristics for a product. Silver and Heravi (2003) show that as the number of characteristics over which the summation takes place in equations (7.30a), (7.31a), or (7.33a) increases, the more redundant becomes the adjustment in equations (7.30b), (7.31b,) or (7.33b). When all characteristic combinations are used (equations [7.30a], [7.31a], or [7.33a]) as strata, the calculation extends to a matched-models problem, in which each cell uniquely identifies a product. For matched data, equations (7.30b), (7.31b, or (7.33b) serve no purpose and the aggregation in equations (7.30a), (7.31a), or (7.33a) would be over all products and reduce to the usual index number problem. Diewert (2003), commenting on the method, explains that when matching is relatively large, the results given are similar to those from superlative hedonic index numbers. Note that the theoretical indices in Chapter 21 are concerned with both goods that are hedonic tied bundles of characteristics *and* goods that are nonhedonic commodities. The framework of equations (7.30), (7.31), or (7.33) allows both types of goods to be included, and there are no adjustments necessary in equations (7.30b), (7.31b), or (7.33b) for the latter non-hedonic ones.

150. The above has illustrated how weighted index number formulas might be constructed using data on prices, quantities, and characteristics for a product when the data are not matched. This is because continuing with matched data may lead to errors from (i) multiple quality

adjustments from products no longer available and their noncomparable replacements and (ii) sample selectivity bias from sampling from a replacement universe as opposed to a double universe.

G.2.4 Difference between hedonic indices and matched indices

151. In previous sections, the advantages of hedonic indices over matched comparisons were referred to in terms of the inclusion by the former of unmatched data. This relationship is developed more formally here. Triplett (2002) argued and Diewert (2003) showed that an unweighted geometric mean (Jevons) index for matched data gives the same result as a logarithmic hedonic index run on the same data. Consider the matched sample *m* and z^{t+2} and z^t as overall quality adjustments to the dummy variables for time in equation (7.28), that is, $\sum_{k=2}^{K}\beta_k z_{ki}$. The very first line in equation (7.35) is shown by Aizcorbe, Corrado, and Doms (2001) to equal the difference between two geometric means of quality-adjusted prices. The sample space $m = M^t = M^{t+2}$ is the same model in each period. Consider the introduction of a new model *n* introduced in period $t + 2$ with no counterpart in *t* and the demise of an old model *o* so it has no counterpart in $t + 2$. So M^{t+2} is composed of *m* and *n*, and M^t is composed of *m* and *o*, and *M* consists only of the matched models *m*. Silver and Heravi (2002) have shown the dummy variable hedonic comparison to now be

$$
\begin{aligned}
(7.35)\ ln\ p^{t+2}/p^t
&= [m/(m+n)\sum_m (ln\ p_m^{t+2} - Z_m)/m \\
&+ n/(m+n)\sum_n (ln\ p_n^{t+2} - Z_n)/n] \\
&- [m/(m+o)\sum_m (ln\ p_m^{t} - Z_m)/m \\
&+ o/(m+o)\sum_o (ln\ p_o^{t} - Z_o)/o] \\
&= [m/(m+n)\sum_m (ln\ p_m^{t+2} - Z_m)/m \\
&- m/(m+o)\sum_m (ln\ p_m^{t} - Z_m)/m] \\
&+ [n/(m+n\sum_n (ln\ p_n^{t+2} - Z_n)\ /n \\
&- o/(m+o)\sum_o (lnp_o^{t} - Z_o)/o].
\end{aligned}
$$

152. Consider the *second* expression in equation (7.35). First, there is the change for the *m* matched observations, the quality adjustment being redundant. This is the change in mean prices of matched models *m* in period *t + 2* and *t* adjusted for quality. Note that the weight in period *t + 2* for this matched component is the proportion of matched to all observations in period $t + 2$. Similarly, for period *t* the matched weight depends on how many unmatched old observations are in the sample in this period. In the last line of equation (7.35), the change is between the unmatched new and the unmatched old mean (quality-adjusted) prices in periods $t + 2$ and *t*.

Thus, matched methods can be seen to ignore the last line in equation (7.35) and will differ from the hedonic dummy variable approach in at least this respect. The hedonic dummy variable approach, in its inclusion of unmatched old and new observations, can be seen from equation (7.35) to possibly differ from a geometric mean of matched price change. The extent of any difference depends, in this unweighted formulation, on the proportions of old and new products leaving and entering the sample and on the price changes of old and new ones relative to those of matched ones. If the market for products is one in which old quality-adjusted prices are unusually low while new quality-adjusted prices are unusually high, then the matched index will understate price changes (see Silver and Heravi, 2002, and Berndt, Ling, and Kyle, 2003, for examples). Different market behavior and changes in technology will lead to different forms of bias.

153. If sales weights replace the number of observations in equation (7.35), then different forms of weighted hedonic indices can be derived as explained in Chapter 21, Section A.5. Silver (2002) has also shown that the hedonic approach will differ from a corresponding weighted or unweighted hedonic regression in respect to the leverage and influence the hedonic regression gives to observations.

G.3 Chaining

154. An alternative approach for dealing with products with a high turnover is to use a chained index instead of the long-term fixed-base comparison. A chained index compares prices of items in period *t* with period *t* + 1 (Index$_{t,t+1}$) and then as a new exercise, studies the universe of products in period *t* + 1 and matches them with items in period *t* + 2. These links, Index$_{t,t+1}$ and Index$_{t+1,\, t+2,}$ are combined by successive multiplication continuing to, say, Index$_{t+5,t+6}$ to form Index$_{t,t+6.}$ Only items available in both period t and period *t* + 6 would be used in a fixed-base trade price index. Consider the five products 1, 2, 5, 6, and 8 over the four months January to April as shown in Table 7.2. The price index for January compared with February (J:F) involves price comparisons for all five products. For (F:M), it involves products 1, 4, 5, and 8; for (M:A), it involves products 1, 3, 4, 5, 7, and 8. The sample composition changes for each comparison as products die and are born. Price indices can be calculated for each of these successive price comparisons using any of the unweighted formulas described in Chapter 21. The sample will grow when new products appear and shrink when old products disappear, changing in composition through time (Turvey, 1999).

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

156. Chaining, unlike hedonic indices, does not use all the price information in the comparison for each link. Products 2 and 6, for example, may be missing in March. The index makes use of the price information on products 2 and 6, when they exist, for the January–February comparison but does not allow their absence to disrupt the index for the February–March comparison. It may be that product 4 is a replacement for product 2. Note how easily it is

included as soon as two price quotes become available. There is no need to wait for rebasing or sample rotation. It may be that product 7 is a replacement for product 6. A quality adjustment to prices may be required for the February–March comparison between products 6 and 7, but this is a short-run, on-off adjustment, The compilation of the index continues for March–April using product 7 instead of product 6. *SNA* (1993, Chapter 16, paragraph 16.54) picks up on the point in its sections on price and volume measurement:

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

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

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

159. There is a possibility that the introduction of new models and exits of old ones instantaneously affects the prices of all existing models. In such a case, the price changes of existing models will suffice. They will reflect the price changes of new entrants and old departures not part of the sample. This argument is used for the case that direct matched-models comparisons, chained matched-models comparisons, and hedonic indices should give the same results. It is an empirical matter, and its plausibility will vary among industries. It is more likely to apply to fast-moving goods with little to no development costs or barriers to entry.

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

H. Long-Run and Short-Run Comparisons

161. This section outlines a formula to help quality adjustment. The procedure can be used with all of the methods outlined in Sections D and E. Its innovation arises from a possible concern with the long-run nature of the quality-adjusted price comparisons being undertaken. In the example in Table 7.2, prices in March were compared with those in January. Assumptions of similar price changes are required by the imputation method to hold over this period for long-run imputations. This gives rise to increasing concern when

Table 7.5. Example of Long-Run and Short-Run Comparisons

Item	January	February	March	April	May	June
Comparable replacement						
A	2	2	2	2	2	2
B	3	3	4	n/a	n/a	n/a
C	n/a	n/a	n/a	6	7	8
Total	5	5	6	8	9	10
Explicit adjustment						
A	2	2	2	2	2	2
B	3	3	4	**5/6** x **6=5**	**5/6** x **7=5.8**	**5/6**x**8= 6.67**
C	**6/5** x **3=3.60**	n/a	n/a	6	7	8
Total	5	5	6	8	9	10
Overlap						
A	2	2	2	2	2	2
B	3	3	4	**6** x **4/5=4.8**	n/a	n/a
C	n/a	n/a	5	6	7	8
Total	5	5	6	6.8	9	10
Imputation						
A	2	2	2.5	3.5	4	5
B	3	3	4	**3.5/2.5** x **4= 5.6**	**4/3.5** x **5.6=6.4**	**5/4** x **6.4=8**
Total	5	5	6.5	9.1	8.4	13

Figures in bold are estimated quality-adjusted prices described in the text.
Note: n/a = not available.

price comparisons continue over longer periods, such as between January and October, January and November, and January and December, and even subsequently. In this section, a *short-run* formulation outlined in Sections C.3.3 and D.2 is more formally considered to help alleviate such concerns. Consider Table 7.5, which, for simplicity, has a single product *A* that exists throughout the period, a product *B* that is permanently missing in April, and a possible replacement *C* in April.

H.1 Short-run comparisons: illustration of some quality adjustment methods

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

$$P_D \equiv \left[\frac{\sum_{i=1}^{N} p_i^{\mathrm{Apr}}/N}{\sum_{i=1}^{N} p_i^{\mathrm{Jan}}/N} \right],$$

which is 8/5 = 1.60, a 60 percent increase. The *short-run* equivalent is the product of a long-run index up to the immediately preceding period and an index for the preceding to the current period, that is, for period $t + 4$ compared with period $t+3$:

$$(7.36)\quad P_D \equiv \left[\frac{\sum_{i=1}^{N} p_i^{\mathrm{t+3}}/N}{\sum_{i=1}^{N} p_i^{\mathrm{t}}/N} \right] \times \left[\frac{\sum_{i=1}^{N} p_i^{\mathrm{t+4}}/N}{\sum_{i=1}^{N} p_i^{\mathrm{t+3}}/N} \right],$$

or, for example, using a comparison of January with April:

$$P_D \equiv \left[\frac{\sum_{i=1}^{N} p_i^{Mar}/N}{\sum_{i=1}^{N} p_i^{Jan}/N} \right] \times \left[\frac{\sum_{i=1}^{N} p_i^{Apr}/N}{\sum_{i=1}^{N} p_i^{Mar}/N} \right],$$

which is of course $\frac{6}{5} \times \frac{8}{5} = 1.6$ as before.

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

164. However, for the overall Dutot index, the results will differ because the Dutot index weights price changes by their price in the initial period as a proportion of total price (Chapter 20, equation [20.1]). The two quality-adjustment methods will have the same price changes but different implicit weights. The Dutot index in May is 9/5.6 = 1.607 using an adjustment to the initial period, January's price, and 7.8/5 = 1.56 using an adjustment to the current period, May's price. The short-run indices give the same results for each adjustment:

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

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

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

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

H.2 Implicit short-run comparisons using imputations

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

168. Consider why the short-run result of 82 percent is larger than the long-run result of 75 percent. The price change for *A* between March and April of 40 percent, on which the short-run imputation is based, is larger than the average *monthly* change of *A*, which is just over 20 percent. The extent of any bias from this approach was found in the previous section to depend on the ratio of missing values and the difference between the average price changes of the matched sample and the quality-adjusted price change of the product that was missing, had it continued to exist. The short-run comparison is to be favored if, as is likely, the assumption of similar price changes is considered more likely to hold than the long-run one.

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

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

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

and in the case of June

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

against long-run comparisons for May:

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

and long-run comparisons for June:

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

171. A note of caution is required here. The comparisons use an imputed value for product *B* in April and also an imputed one for May. The price comparison for the second term in equation (7.36), for the current versus immediately preceding period, uses imputed values for

product *B*. Similarly, for the January to June results, the May to June comparison uses imputed values for product *B* for both May and June. The pragmatic needs of quality adjustment may demand this. If comparable replacements, overlap links, and resources for explicit quality adjustment are unavailable, an imputation must be considered. However, using imputed values as lagged values in short-run comparisons introduces a level of error into the index that will be compounded with their continued use. Long-run imputations are likely to be preferable to short-run changes based on lagged imputed values unless there is something in the nature of the product that cautions against such long-run imputations. There are circumstances when the price collector may believe the missing product is missing temporarily, and the imputation is conducted under the expectation that the product will subsequently reappear. A wait-and-see policy is adopted under some rules—three months, for example—after which it is deemed to be permanently missing. These are the pragmatic situations that require imputations to extend over consecutive periods. These circumstance promote lagged imputed values to compare against current imputed values. This is cautioned against, especially over a period of several months. There is an intuition that the period in question should not be extensive. First, the effective sample size is being eaten up as the use of imputation increases. Second, the implicit assumptions of similar price movements inherent in imputations are less likely to hold over the longer run. Finally, there is some empirical evidence, albeit from a different context, against using imputed values as lagged actual values. (See Feenstra and Diewert's 2001 study using data from the U.S. Bureau of Labor Statistics for their International Price Program.)

172. The short-run approach described above will be developed in the next section, where weighted indices are considered. The practice of estimating quality-adjusted prices is usually at the elementary product level. At this lower level, the prices of products may subsequently be missing and replacements with or without adjustments and imputations are used to allow the series to continue. New products and varieties are also being introduced; the switching of sales between sections of the index becomes prevalent. The turmoil of changing quality is not just about the maintaining of similar price comparisons but also about the accurate reweighting of the mix of what is sold. Under a Laspeyres framework, the bundle is held constant in the base period, so any change in the relative importance of products sold is held to be of no concern until the next rebasing of the index. Yet capturing some of the very real changes in the mix of what is sold requires procedures for updating the weights. This was considered in Chapter 5. The concern here is with a higher-level procedure equivalent to the short-run adjustments discussed above. It is one particularly suited to countries where resource constraints prohibit the regular updating of weights through regular household surveys.

H.3 Single-stage and two-stage indices

173. Consider aggregation at the elementary level (Chapter 6). This is the level at which prices are collected from a representative selection of outlets across regions in a period and compared with the matched prices of the same products in a subsequent period to form an index for a good. Lamb is an example of a good in an index. Each price comparison is equally weighted unless the sample design gave proportionately more chance of selection to products with more sales. The elementary price index for lamb is then weighted and combined with the weighted elementary indices for other products to form the CPI. A Jevons elementary aggregate index, for example, for period $t + 6$ compared with period t is given as

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

Compare this with a two-stage procedure:

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

174. If a product is missing in period t + 6, an imputation may be undertaken. If equation (7.37) is used, the requisite assumption is that the price change of the missing product, had it continued, is equal to that of the average of the remaining products *over the period* t *to* t *+ 6*. In equation (7.38), the missing product in period t + 6 may be included in the first stage of the calculation, between periods t and t + 5, but excluded in the second stage, between periods t + 5 and t + 6. The requisite assumption is that price changes *between* t *– 1 and* t are similar. Assumptions of short-run price changes are generally considered to be more valid than their long-run counterparts. The two-stage framework also has the advantage of including in the worksheet prices for the current period and the immediately preceding one, which, as will be shown in Chapter 9, promotes good data validity checks.

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

Appendix 7.1. Data for Hedonic Regression Illustration

Pric (£)	Speed (MHz)	RAM	HD	Dell	Presaric	Prosigni	Celeron	Pentium III	CD-R	DVI	Dell × Speed
2,123	1,000	128	40	0	1	0	0	0	0	0	0
1,642	700	128	40	0	1	0	0	0	0	0	0
2,473	1,000	384	40	0	1	0	0	0	0	0	0
2,170	1,000	128	60	0	1	0	0	0	0	0	0
2,182	1,000	128	40	0	1	0	0	0	0	1	0
2,232	1,000	128	40	0	1	0	0	0	1	0	0
2,232	1,000	128	40	0	1	0	0	0	0	0	0
1,192	700	384	40	0	1	0	0	0	0	0	0
1,689	700	384	60	0	1	0	0	0	0	0	0
1,701	700	384	40	0	1	0	0	0	0	1	0
1,751	700	384	40	0	1	0	0	0	1	0	0
1,851	700	384	40	0	1	0	0	0	0	0	0
2,319	933	128	15	0	0	0	0	1	0	0	0
2,512	933	256	15	0	0	0	0	1	0	0	0
2,451	933	128	30	0	0	0	0	1	0	0	0
2,270	933	128	10	0	0	0	0	1	0	0	0
2,463	933	256	10	0	0	0	0	1	0	0	0
2,183	933	64	10	0	0	0	0	1	0	0	0
1,039	533	64	8	0	0	1	1	0	0	0	0
1,139	533	128	8	0	0	1	1	0	0	0	0
1,109	533	64	17	0	0	1	1	0	0	0	0
1,180	533	64	8	0	0	1	1	0	1	0	0
1,350	533	128	17	0	0	1	1	0	1	0	0
1,089	600	64	8	0	0	1	0	1	0	0	0
1,189	600	128	8	0	0	1	0	1	0	0	0
1,159	600	64	17	0	0	1	0	1	0	0	0
1,230	600	64	8	0	0	1	0	1	1	0	0
1,259	600	128	17	0	0	1	0	1	0	0	0
1,400	600	128	17	0	0	1	0	1	1	0	0
2,389	933	256	40	0	1	0	0	1	0	0	0

Pric (£)	Speed (MHz)	RA M	HD	Dell	Presaric	Prosigni	Celeron	Pentiu m III	CD-R'	DVI	Dell × Spee d
1,833	733	256	40	0	1	0	0	1	0	0	0
2,189	933	128	40	0	1	0	0	1	0	0	0
2,436	933	256	60	0	1	0	0	1	0	0	0
2,397	933	256	40	0	1	0	0	1	0	1	0
2,447	933	256	40	0	1	0	0	1	1	0	0
2,547	933	256	40	0	1	0	0	1	0	0	0
2,845	933	384	60	0	1	0	0	1	0	0	0
2,636	933	384	60	0	1	0	0	1	0	0	0
1,507	733	64	30	0	1	0	0	1	0	0	0
1,279	667	64	10	1	0	0	0	1	0	0	667
1,379	667	128	10	1	0	0	0	1	0	0	667
1,399	667	64	30	1	0	0	0	1	0	0	667
1,499	667	128	30	1	0	0	0	1	0	0	667
1,598	667	128	30	1	0	0	0	1	1	0	667
1,609	667	128	30	1	0	0	0	1	0	1	667
1,389	667	64	10	1	0	0	0	1	0	1	667
999	667	64	10	1	0	0	1	0	0	0	667
1,119	566	64	30	1	0	0	1	0	0	0	566
1,099	566	128	10	1	0	0	1	0	0	0	566
1,097	566	64	10	1	0	0	1	0	1	0	566
1,108	566	64	10	1	0	0	1	0	0	1	566
1,219	566	128	30	1	0	0	1	0	0	0	566
1,318	566	128	30	1	0	0	1	0	1	0	566
1,328	566	128	30	1	0	0	1	0	0	1	566
1,409	566	128	10	1	0	0	0	1	0	0	733
1,809	733	384	10	1	0	0	0	1	0	0	733
1,529	733	128	30	1	0	0	0	1	0	0	733
1,519	733	128	10	1	0	0	0	1	0	1	733
1,929	733	384	30	1	0	0	0	1	0	0	733
2,039	733	384	30	1	0	0	0	1	0	1	933
2,679	933	128	30	1	0	0	0	1	0	0	933
3,079	933	384	10	1	0	0	0	1	0	0	933
2,789	933	128	10	1	0	0	0	1	0	1	933
3,189	933	384	10	1	0	0	0	1	0	1	933