8. Item Substitution, Sample Space, and New Goods

A. Introduction

8.1 In the introduction to Chapter 7, the use of the matched-models method was recognized as the accepted approach to ensure that the measurement of price changes was untainted by changes in their quality. However, it was noted that the approach might fail in three respects: missing items, sampling issues, and new goods and services (hereafter "goods" includes services). Missing items are the subject of Chapter 7, in which several implicit and explicit methods of quality adjustment to prices, and the choice between them, were discussed. In this chapter, attention is turned to two other reasons why the matched-models method may fail: sampling issues and new goods. The three sources of potential error are briefly outlined.

- **Missing items.** A problem arises when an item is no longer produced. An implicit quality adjustment may be made using the overlap or imputation method, or the respondent may choose a replacement item of a comparable quality, and its price may be directly compared with the missing item's price. If the replacement is of a noncomparable quality, an explicit price adjustment is required. This was the subject of Chapter 7, Sections C through F. In Section G of Chapter 7 a caveat was added. Items in industries where model replacements were rapid, continued long-run matching would deplete the sample, and quality adjustment becomes unfeasible on the scale required. Chained matching or hedonic indices were deemed preferable.

- **Sampling issues.** The matching of prices of identical items over time, by its nature, is likely to lead to monitoring of a sample of items increasingly unrepresentative of the population of transactions. Respondents may keep with their selected items until they are no longer produced—that is, continue to monitor items with unusual price changes and limited sales. Yet on item replacement, respondents may select unpopular comparable items to avoid explicit quality adjustments; obsolete items with unusual price changes are replaced by near-obsolete items that also have unusual price changes, compounding the problem of unrepresentative samples. The substitution of an item with relatively high sales for an obsolete one has its own problems, since the difference in quality is likely to be substantial and substantive, beyond what can be attributed to, say, the price difference in some overlap period. One would be in the last stage of its life cycle and the other in its first. The issue has implications for sample rotation and item substitution.

- **New products.** A third potential difficulty arises when something "new" is produced. There is a difficulty in distinguishing between new items and quality changes in old ones, which will be discussed below. When a new good is produced, there is a need for it to be included in the index as soon as possible, especially if the product is expected to be responsible for relatively high sales. New goods might have quite different price changes than existing ones, especially at the start of their life cycle. In the initial period of introduction, producers often gain more from their ability to receive higher prices from their recently introduced product than might be attainable once the market settles into a competitive equilibrium. But by definition, there is no price in the period preceding the introduction of the *new* product. So even if prices of new products were obtained and included in the index from the initial introduction date, there would still be something missing—the initial high price producers can reap by exploiting any monopoly power in the period of launch.

8.2 The problem of missing items was the subject of Chapter 7. In this chapter, sampling issues arising out of the matched-models approach and the problem of introducing new goods into the index are considered.

B. Sampling Issues and Matching

B.1 Introduction

8.3 The matching procedure has at its roots a conundrum. Matching is designed to avoid price changes being contaminated by quality changes. Yet its adoption constrains the sampling to a static universe of items that exist in *both* the reference and base periods. Outside of this there is, of course, something more: items that exist in the reference period but not in the current period, and are therefore not matched; and similarly those new items existing in the current period but not in the reference one—the dynamic universe (Dalén, 1998, and Sellwood, 2001). The conundrum is that the items not in the matched universe, the new items appearing after the reference period and the old items that disappeared from the current period, may be the ones whose price changes differ substantially from existing matched ones. They will embody different technologies and be subject to different (quality-adjusted) strategic price changes. The very device used to maintain a constant quality sample may itself give rise to a sample biased away from technological developments. Furthermore, when this sample is used to make imputations (Chapter 7, Sections D.1 and D.2) as to the price changes of replacement items, it reflects the technology of a sample not representative of current technological changes.

8.4 A formal consideration of matching and the dynamic universe is provided in Appendix 8.1. Three universes are considered:

- An *intersection* universe, which includes only matched items;
- A dynamic *double universe*, which includes all items in the base comparison period and all in the current period, although they may be of different qualities; and
- A *replacement* universe, which starts with the base-period universe but also includes one-to-one replacements when an item from the sample in the base period is missing in the current period.

8.5 It is, of course, difficult to ascertain the extent to which matching from the intersection universe constrains the penetration of the sample into the dynamic double universe, since statistical agencies generally do not collect data for the latter. Its extent will, in any event, vary between commodities. Sellwood (2001) advocated simulations using the universe of scanner data. Silver and Heravi (2002) undertook such an experiment using scanner data on the consumer prices of washing machines in the United Kingdom in 1998. A matched Laspeyres index—based on price comparisons with matched models existing in both January and December—covered only 48 percent of the December expenditure on washing machines, as a result of new models that were introduced after January not being included in the matched index. Furthermore, the January to December matched comparison covered only just over 80 percent of the January expenditure, because of the exclusion of models available in January but not in December. A biannual sample rotation (rebasing) increased the December expenditure coverage to just over 70 percent and a monthly (chained) rotation to about 98 percent (see also Chapter 7, Section G.1, for further examples). Two implications arise from this. First, selection of item substitutes (replacements) puts coverage of the sample to some extent under the control of the respondents. Guidelines on directed replacements in particular product areas have some merit. Second, chaining, hedonic indices (as considered in Chapter 7, Section G) and regular sample rotation also have merit in some commodity areas as devices to refresh the sample.

B.2 Sample space and item replacement or substitution

8.6 The respondents often are best placed to select replacement items for repricing. They are aware of not only the technological basis of the items being produced but also their terms of sale. The selection of the replacement for repricing might be quite obvious to the respondent. There may be only a slight, nominal improvement to the item. For example, the "improved" item is simply a replacement variety sold instead of the previous one. The replacement could have a different code or model number and will be known to the respondent as simply a different color or packaging. The specification list given to the respondent is a critical prompt as to when a repriced item is different, and it is important that this include all price-determining factors.

8.7 The respondent prompted by the specification list takes on the role of identifying whether an item is of comparable quality or otherwise. If it is judged to be comparable when it is not, the quality difference is taken to be a price difference, and a bias will result if the unrecognized quality changes are in a consistent direction. Informed comparable substitution requires general guidelines on what makes a good substitute as well as product-specific information on likely price-determining characteristics. It also requires timely substitution to maximize the probability of an appropriate substitute being available.

8.8 Liegey (1994), for CPI purposes, notes how useful the results from hedonic regressions are in the selection of items. The results provide an indication of the major quality factors that make up the product or service, in terms of explaining price variation. Not only would the selection of items be more representative, but the coefficients from hedonic regressions, for their subsequent use to estimate quality-adjusted prices, would be more tailored to the sample in hand.

8.9 On repricing, respondents traditionally are required to find substitute items that are as similar as possible to the items being replaced. This maximizes the likelihood that the old and replacement item will be judged equivalent and so minimizes the need to employ some method of quality adjustment. Yet, replacement items should be chosen so that they intrude into the sampled items in a substantial and representative manner so as to make the sampled items more representative of the dynamic universe. The inclusion of a popular replacement item to refresh the sample—one at the same point in its life cycle as the original popular one selected in the base period—allows for a useful and accurate price comparison and increases the chance of an appropriate quality adjustment being undertaken. It is of little merit to substitute a new item with limited sales for a missing item with limited sales, just because they have similar features of both being "old." The index would become more unrepresentative. Yet if replacements are made for items at the end of their life with popular replacements items at the start of their life, the quality adjustment will be substantial and substantive. More frequent sample rotation or directed replacements will be warranted for some commodity areas.

- Replacements offer an opportunity to cut back on and possibly remove sample bias in the period of replacement, though not prior to it;
- The more frequent the replacement, the less the bias;
- If there is more than one new (replacement) item in the market, there may still be bias since only the most popular one will be selected, and it may be at a different stage in its life cycle than others and priced differently;
- The analysis assumes that perfect quality adjustments are undertaken on replacements. The less frequent the replacement, the more difficult this might be, because the very latest replacement item on the market may have more substantial differences in quality than earlier ones;
- If the replacement item has relatively high sales, is of comparable quality, and is at the same stage in its life cycle as the existing one, then its selection will minimize bias;
- If there is more than one new (replacement) item and the most comparable one is selected at the old technology, it will have low market share and unusual price changes;
- Given advance market or production information, replacements undertaken before obsolescence are likely to increase the sample's share of the market, include items more representtative of the market, and facilitate quality adjustment.

8.10 The problem of item substitution is analogous to the problems that arise when an establishment closes. It may be possible to find a comparable establishment not already in the sample, or a noncomparable one for which, in principle, an adjustment can be made for the better quality of service of the new one. It is not unusual for an establishment to close following the introduction of a new factory. Thus, here is an obvious replacement factory. However, if the new establishment has comparable prices but a better range of items, delivery, and service quality, there is a gain to purchasers from substituting one factory's output for the other. Yet, since such facilities have no direct price, it is difficult to provide estimates of the value of such services in order for an adjustment to be made for the better quality of service of the new one. The index thus would have an upward bias, which would be lost on rebasing. In such cases, substituting the old establishment for a new one that provides a similar standard of service may be preferable.

B.3 Sample rotation, chaining, and hedonic indices

8.11 It is important also to recognize the interrelationships among the methods for handling item rotation, item replacement, and quality adjustment. When PPI item samples are rotated, this is a form of item substitution, except that it is not "forced" by a missing item but is undertaken for a general group of items to update the sample. Rotation has the effect of making future forced replacements less likely. Yet the assumptions implicit in its use are equivalent to those for the overlap adjustment technique: price differences are an adequate proxy for the change in price per unit of quality between items disappearing from the sample and replacement items. Consider the initiation of a new sample of items by probability or judgmental methods or a combination of the two. Prices for the old and new sample are returned in the same month and the new index is compiled on the basis of the new sample, with the results being linked to the old. This is an implicit use of the overlap method, in which all price differences between the new and old items are taken to be quality changes. Assume the initiation is in January. The prices of an old item in December and January are 10 and 11, respectively, a 10 percent increase, and those for the replacement item in January and February are 16 and 18, respectively, an increase of 12.5 percent. The new item in January is of a better quality than the old, and this difference in quality may be worth 16 – 11 = 5; that is, the price difference is assumed to be equal to the quality difference, which is the assumption implicit in the overlap method. Had the price of the old item in December been compared with the quality-adjusted price of the new item in January under this assumption, the price change would be the same: 10 percent (that is, (16 – 5)/10 = 1.10). If, however, the quality difference in January between the two prices was more than the revenue difference to the producer, the result would be wrong. In practice, the need to simultaneously replace and update a large number of items requires the assumptions of the overlap method. This process should not be regarded as error-free, and in cases where the assumptions are likely to be particularly untenable (discussed in Chapter 7, Section D.2), explicit adjustments of the form discussed in Chapter 7, Section E, should, resources permitting, be used.

8.12 It was noted above that when samples are updated, any difference in average quality between samples is dealt with in a way that is equivalent to the overlap adjustment technique. Sample rotations to freshen the sample between rebasing are an expensive exercise. However, if rebasing is infrequent and there is a substantial loss of items in particular industries, then this might be appropriate for those industries. In the next section the need for a metadata system to facilitate such decisions will be outlined. The use of a more frequent sample rotation aids the process of quality adjustment in two ways. First, the updated sample will include newer varieties, comparable replacements with substantial sales will be more likely to be available, and noncomparable ones will be of a more similar quality, which will aid good explicit adjustments. Second, because the sample has been rotated, there will be fewer missing items than otherwise and thus less need for quality adjustments.

8.13 A natural extension of more frequent sample rotation is to use a chained formulation in which the sample is reselected each period. In Chapter 7, Section G.3, the principles and methods are outlined in the context of sectors in which there is a rapid turnover of items, and such principles are echoed here. Similarly, the use of hedonic indices as outlined in Chapter 7, Section G.2, or the use of short-run comparisons discussed in Chapter 7, Section H, might be useful in this context.

8.14 Chaining, as discussed in Chapter 7, Section G.3, allows the price changes of a new commodity to be included as soon as the commodity can be priced for two successive periods. The new price's effect on the index in the initial period of introduction is ignored. Fisher and Shell (1972) suggest that the preceding price is imputed as the reservation price given the current-period technology, where the reservation price is defined as the maximum price at which zero production of the good is forthcoming, given current-period inputs and prices of other outputs in the preceding period (see Hicks, 1940, and Hausman, 1997, for equivalent considerations for the consumer price index). Similar concerns arise for disappearing commodities. A disappearing good's price has to be imputed in the current period, which is imputed as the reservation price given the preceding period technology, defined as the maximum price in the current period at which no production of the good is forthcoming, given inputs in the preceding period and prices of other outputs in the current period. The estimation of such reservation prices is not practical, though

Hausman (1997) provides an example in the context of the CPI. If the new commodity is not entirely new, in the sense that it is providing more services than those of the old product, a hedonic estimate of the reservation price can be used to estimate the cost of the base situation characteristics for the missing price of the disappearing good or the cost of the current situation characteristics for the missing reference price of the new variety (Zieschang, 1988). However, this applies only when the good is not entirely new, so that the price can be determined in terms of a different combination of the existing character set.

C. Information Requirements for a Strategy for Quality Adjustment

8.15 It should be apparent from the above that a strategy for quality adjustment must not only be linked to sample representativity, but it also requires building a statistical metadata system. The approach for the index as a whole cannot be described simply. It requires the continual development of market information and the recording and evaluation of methods on a commodity-by-commodity basis. The rationale for such a metadata system relates to the variety of procedures for quality adjustments to prices discussed in Chapter 7, Section C.3.4, and how their suitability might vary on a case-by-case basis, all of which require documentation.

C.1 Statistical metadata system

8.16 The methods used for estimating quality-adjusted prices should be well documented as part of a statistical metadata system. Metadata is systematic, descriptive information about data content and organization that helps those who operate the statistics production systems to remember what tasks they should perform and how they should perform them. A related purpose is to introduce new staff to and train them in the production routines (Sundgren, 1993). The metadata, as proposed in this context, are also to help identify where current methods of quality adjustment require reconsideration and will prompt the use of alternative methods.[1] The dramatic increase in the volume of statistical data in machine-readable form has some arguing for keeping metadata in such a form. This is to encourage transparency in the methods used and help ensure that they are understood and continued as staff members leave and others join. Changes in quality adjustment methodology can in themselves lead to changes in the index. Indices for products using new procedures should be spliced onto existing indices. The metadata system also should be used as a tool to help with quality adjustment. Because so much of the rationale for the employment of different methods is specific to the features of the industries concerned, data should be kept on such features. The metadata system should help in the following ways:

- Statistical agencies should monitor the incidence of missing items against each three-digit ISIC code, and if the incidence is high, then at the four- or more digit level or by elementary aggregate to the most detailed level of the system. Where the incidence is high, the ratios of temporary missing prices, comparable replacements, and noncomparable replacements to the overall number of prices, and the methods for dealing with each of these three circumstances, also should be monitored to provide the basis of a statistical metadata system. The advantage of a top-down approach is that resources are saved by monitoring at the detailed level only the product areas that are problematic.
- Product-specific information—such as the timing of the introduction of new models, pricing policies, especially in months when no changes were made, and popularity of models and brands according to different data sources—should be included as the system develops.
- An estimate, if available, of the weight of the product concerned should be given so that a disproportionate effort is not given to relatively low-weighted items. All of this will lead to increased transparency in the procedures used and allow effort to be directed where it is most needed.
- The statistical metadata system will benefit from contacts among market research organizations, retailers, manufacturers, and trade associations for items for which replacement levels are high. The development of such links may lead, for example, to option cost estimates, which can be easily introduced. Where possible, staff should be encouraged to learn more about specific industries whose weights are relatively high and where item replacement is

[1] Metadata may also serve user needs, the oldest and most extensive form being footnotes (Silver, 1993).

common. Such links to these organizations will allow staff to better judge the validity of the assumptions underlying implicit quality adjustments.

- Industries likely to be undergoing regular technological change should be identified. The system should attempt to ascertain the pace at which models change and, where possible, the timing.
- Price-determining characteristics for product areas using hedonic regressions, information from market research, store managers, trade and other such bodies, and the experience of price collectors should be identified. This should contribute to the statistical metadata system and be particularly useful in providing subsequent guidelines on item selection.
- The system should undertake an analysis of what have in the past been judged to be "comparable" replacements in terms of the factors that distinguish the replacement and old item. The analysis should identify whether different respondents are making similar judgments and whether such judgments are reasonable.
- When hedonic regressions are used either for partial patching of missing prices or as indices in their own right, information on the specification, estimated parameters, and diagnostic tests of the regression equations should be kept along with notes as to why the final formulation was chosen and used along with the data. This will allow the methodology for subsequent updated equations to be benchmarked and tested against the previous versions.
- Price statisticians may have more faith in some quality adjustment procedures than others. When such procedures are used extensively, it might be useful to note, as part of the metadata system, the degree of faith the statistician has in the procedures. Following Shapiro and Wilcox (1997b) this may be envisaged as a traditional confidence interval: the statistician believes at a 90 percent level of confidence that the quality-adjusted price change is 2 percent (0.02) with an overall width of 0.005, for example. There may be an indication as to whether the interval is symmetric or positively or negatively one-sided. Alternatively, statisticians may use a simple subjective coding on a scale of one to five.

D. Incorporating New Goods

D.1 What are new goods and how do they differ from quality changes?

8.17 A new model of a good may provide more of a currently available set of service flows. For example, a new model of an automobile is different from an existing one in that it may have a bigger engine. There is a continuation of a service and production flow, and this may be linked to the service flow and production technology of the existing model. The practical concern with the definition of a new good's quality changes against an updated existing model is that, first, the former cannot be easily linked to an existing item as a continuation of an existing resource base and service flow because of the very nature of its newness. Some forms of genetically modified seeds, frozen foods, microwave ovens, and mobile phones, while extensions of existing services, have a dimension of service that is quite new. Second, new goods can generate a welfare gain to consumers and surplus to producers by their very introduction, and the simple introduction of the new good into the index, once two successive price quotes are available, misses this gain.

8.18 Oi (1997) directs the problem of defining new goods to that of defining a monopoly. If there is no close substitute, the good is new. He argues that some individual new videos may have quite small cross-elasticities with other videos; their shared service is to provide movie entertainment and they are similar only in this respect. The same argument may apply to some new books and new breakfast cereals. However, Hausman (1997) found cross-elasticities for substitution to be quite substantial for new breakfast cereals. There are many new forms of existing commodities, such as fashionable toys, which are not easily substitutable for similar items, and thus manufacturers could generate a substantial surplus over and above what might be expected from their production costs. The ability of manufacturers to generate monopoly surpluses is one way of considering whether items are new.

8.19 However, Bresnahan (1997, p. 237) notes that for the United States, *Brandweek* counted more than 22,000 new-product introductions in 1994—the purpose of their introduction being, as differentiated products, to be distinct and not exact substitutes for existing ones. Their distinctiveness is in many cases the rationale behind their launch. However, the extent of differentiated markets makes impractical the definition and treatment of such things as new. Oi (1997, p. 110) provides the pragmatic case: "Our theory and statistics would be unduly cluttered if separate product codes had to be set aside for Clear Coke and Special K." Furthermore, the techniques for their inclusion are not readily applicable, and the sound practical advice given by Oi (1997) to keep matters uncluttered is therefore not unreasonable.

8.20 Merkel (2000, p. 6) is more practical in devising a classification scheme that will meet the needs of PPI compilation (see also Armknecht, Lane, and Stewart, 1997, for CPIs). Merkel considers *evolutionary* and *revolutionary* goods. The former are defined as

> …extensions of existing goods. From a production inputs standpoint, evolutionary goods are similar to pre-existing goods. They are typically produced on the same production line and/or use largely the same production inputs and processes as pre-existing goods. Consequently, in theory at least, it should be possible to quality adjust for any differences between a pre-existing good and an evolutionary good.
>
> In contrast, revolutionary goods are goods that are substantially different from pre-existing goods. They are generally produced on entirely new production lines and/or with substantially new production inputs and processes than those used to produce pre-existing goods. These differences make it virtually impossible, both from a theoretical and practical standpoint, to quality adjust between a revolutionary good and any pre-existing good.

8.21 Quality adjustments to prices are therefore suitable for evolutionary goods under the FIOPI framework but unsuitable for revolutionary goods. The definitions are designed to distinguish between the two types of goods not in terms of what is analytically appropriate, but by what is practically meaningful for the needs of PPI construction. It is quite possible for a new item made from the same inputs and processes as the old one to have a high cross-elasticity of substitution and, thus, command revenue for each item beyond what might be expected from a normal markup. Yet practical needs are important in this context, especially because the methods for estimating the producers' surplus are not practically possible given their substantial resource needs of data and econometric expertise.

D.2 The issues

8.22 There are two major concerns regarding the incorporation of new goods into the PPI. First is their identification and detection; second is the related decision on the need and timing for their inclusion. This refers to both the weight and price changes of the new goods. Consider some examples.

8.23 First, the production of cellular phones, for example, was in some countries at such a significant level that their early inclusion in the PPI became a matter of priority. They simply rose from nothing to quite a large proportion of output in their industry. Furthermore, their price changes were atypical of other goods in their industry. Being new, they may be produced using inputs and technologies quite different from those used for existing ones.

8.24 Many new goods can command substantial sales and be the subject of distinct pricing strategies at introduction because of substantial marketing campaigns. Dulberger (1993) provided some estimates for U.S. PPIs for dynamic random-access memory (DRAM) computer memory chips. She calculated price indices for the period from 1982 to 1988 with varying amounts of delay in introducing new chips into the index. The indices were chained so that new chips could be introduced, or not, as soon as they were available for two successive years. Using a Laspeyres chained index, the fall of 27 percent, if there is no delay in introducing new goods, was compared with falls of 26.2 percent, 24.7 percent, 19.9 percent, 7.1 percent, and 1.8 percent, if the introductions were delayed by one year, two years,…, five years, respectively. In all cases, the index is biased downward because of the delay. The longer the delay, the more the price changes of new products are estimated by products whose market shares may be quite small. Berndt and others (1997) provided a detailed study of the new anti-ulcer drug Tagamet and found the effects of preintroduction marketing on its price and market share at introduction to be quite substantial. Not unexpectedly, price falls were found for the generic form of a pharmaceutical on the expiration of the patent, but *increases* were found for the branded form because loyal customers were willing to pay a premium over the price prior to the patent expiration (Berndt, Ling, and Kyle, 2003).

8.25 Waiting for a new good to be established or waiting for the rebasing of an index before incorporating new products may lead to errors in the measurement of price changes if the unusual price movements cycles are ignored at critical stages in the product life. Strategies are required for the early identification of new products and mechanisms for their incorporation either at launch, if preceded by major marketing strategies, or soon after, if there is evidence of market acceptance. This should form part of the metadata system. Waiting for a new product to achieve market maturity may result in an implicit policy of ignoring the quite disparate price movements that accompany their introduction (Tellis, 1988, and Parker, 1992). This is not to say that new goods will always have different price changes. Merkel (2000) gives the example of "lite" varieties of foods and beverages, similar to the original ones but with less fat. They have prices very close to the original ones and serve to expand the market. While there is a need to capture such expansion when the weights are revised, the price changes for the existing items can be used to capture those of the lite ones.

D.3 Methods

8.26 The methods outlined here include those that fall under what should be normal PPI procedures and those that would require exceptional treatment. In the former case, consideration will be given in Section D.3.1 to the rebasing of the index, rotating of items, introduction of new goods as replacements for discontinued ones, and a strategy for dealing with new-item bias. In the latter, techniques that require different sets of data will be outlined. The use of chained matched models and hedonic indices were outlined and discussed in Chapter 7, Section G, "High-Technology and Other Sectors with Rapid Turnover of Models."

D.3.1 Sample rebasing, rotation, directed replacements, and sample augmentation

D.3.1.1 Sample rebasing and rotation

8.27 The concern here is mainly with *evolutionary goods*. A new good may be readily incorporated in the index at the time of rebasing the index or when the sample is rotated. If the new good has, or is likely to have, substantial sales and is not a replacement for a preexisting one, or is likely to command a much higher or lower market share than the preexisting one it is replacing, then new weights are necessary to reflect this. New weights are fully available only at rebasing, not on sample rotation. There will be a delay in the new item's full inclusion, and the extent of the delay will depend on how close its introduction is to the next rebasing and, more generally, the frequency with which the index is rebased. The term rebasing here is effectively concerned with the use of new weights for the index. Even if the index was rebased annually and chained, it would take until the annual rebasing before weights could be assigned, and even then there might be a further six-month delay in the sampling and collating of the survey results for the weights. More frequent rebasing allows for an earlier introduction of the new good and is advised when the weights are not keeping pace with product innovations.

8.28 At the elementary level of aggregation, equal (implicit) weight is given by the Jevons index—for example, to each price relative. The Dutot index gives each price change the weight of its price relative to the sum of the prices in the initial base period of the comparison (Chapter 20, Section B). If an industry is expected to be subject to dynamic innovations, then the sample may be increased without any changes to the weight for the group. There simply would be more items selected to form the arithmetic or geometric average price change. As new varieties become available, they can be substituted for some of the existing ones, with a wider range from which to draw a comparable one or with less effort involved in the quality adjustment procedure for a noncomparable one.

8.29 Some statistical agencies rotate (resample) items within industry groups. Opportunities exist to introduce new items within a weighted group under such circumstances. The resource practicalities of such schemes require items to be rotated on a staggered basis for different industries, with industries experiencing rapid change being rotated more frequently. For example, DVDs could replace VCR tapes using the overlap method, with the difference in prices in the overlap period assumed to be equal to their quality difference. The assumptions implicit in such procedures have been outlined above, and their likely veracity needs to be considered. Since evolutionary items are defined as continuations of the service flow of exiting ones, the hedonic framework may be more suitable; further methods and their choice were discussed in Chapter 7, Sections D through F. However, the principle remains for including new goods in an index within a weighting system as a substitute for old goods.

8.30 Yet in many countries rebasing is infrequent and sample rotation not undertaken. Furthermore, rotating samples on a frequent basis should not be considered as a panacea. Sample rotation is an arduous task, especially when performed over a range of industries experiencing rapid change. Even frequent rotation, say, every four years, may miss many new goods. Experience in the United States has found that frequent rotation (resampling) has had a negative impact on participation rates, since respondents shy away from incurring the indirect costs associated with being interviewed about their product range and technology (Merkel, 2000). Yet it is not necessary for statistical agencies to wait until an item is obsolete before the new one is introduced. It is quite feasible for statistical agencies to preempt the obsolescence of the old item and direct an early substitution of the new. In some industries, the arrival of a new good is well advertised in advance of the launch, while in others it is feasible for a statistical agency to have more general procedures for directed substitutions, as will be outlined below. Without such a strategy and infrequent rotation and rebasing, a country would be open to serious new product bias. In summary,

- The treatment of a new good as a replacement for an existing one can be undertaken if the old item's weights suitably reflect the new good's sales and if suitable quality adjustments can be made to its price to link it to the existing old price series.
- If the new good does not fit into the preexisting weighting structure, it can be included on rebasing, though this may be infrequent in some countries.
- Regular sample rotation provides a means by which the inclusion of such items can be formally reconsidered. Since this is undertaken on a staggered basis, only the weights within the industry are reallocated, not those between industries.
- Directed sample substitution, as opposed to waiting for sample rotation, may be used to preempt the arrival of new goods.
- Revolutionary items, tectonic shifts, and entirely new products will not fit into existing weighting structures, and alternative means are required.
- Directed replacements for evolutionary goods as replacement items and for revolutionary goods to augment the sample are considered below.
- The chained framework outlined in Chapter 15, Section F, may be more appropriate for product areas with high turnover of items.

D.3.1.2 Directed replacements and sample augmentation

8.31 For *evolutionary goods* in industries with a rapid replacement and introduction of such goods, a policy of directed substitution might be adopted. Judgment, experience, and a statistical metadata system should help identify such industries. The existing items should be coded into well-defined product lines. The respondents then are contacted on a regular (say, annual) basis to establish whether a new version has been introduced and, if so, what percentage of the product line's revenue is represented by the new version. Replacement could be decided by a number of criteria. If the new version is designed as a replacement for an existing one, then substitution might be automatic. Once a substitute has been made, the prices require adjustment for the quality differences using the overlap method, imputation, or an explicit estimate based on production or option costs or a hedonic regression. Examples of forms to help guide this process of directed substitution are given in Merkel (2000).

8.32 It is important to emphasize that, on the introduction of new versions of these evolutionary goods, a price may be charged over and above that which can be ascribed to the resource costs behind its difference from the old one. A new version of, for example, electrical cable may have stronger and more flexible plastic coating and the resource cost behind its production may be quite small. Yet it may be sold at a much higher price than the old version because it's seen to be superior to other products in the market. This price increase is a real one that should, after subtraction of the difference in resource costs, be captured by the PPI. After a while prices may be reduced as the novelty of the item wears off or as competitors bring out improved products. The directed substitution becomes important so that the unusual price increases at the introduction are captured by the PPI. It is also necessary so that the coverage of items becomes more representative. Directed substitution allows both.

8.33 However, for *revolutionary goods* substitution may not be appropriate. First, they may not be able to be defined within the existing classification systems. Second, they may be primarily produced by a new establishment, which will require extending the sample to include such establishments. Third, there will be no previous items to match it against and make a quality adjustment to prices since, by definition, they are substantially different from preexisting goods. Finally, there is no weight to attach to the new establishment or item.

8.34 The first need is to identify new goods and the proposal for contacts with market research companies, outlet managers, and manufacturers discussed in Section C.1 on producing a supporting metadata system. *Once identified, sample augmentation is appropriate for the introduction of revolutionary goods, as opposed to sample substitution for evolutionary goods*. It is necessary to bring the new revolutionary good into the sample in addition to what exists. This may involve extending the classification, the sample of outlets, and the item list within new or existing outlets. The means by which the new goods are introduced is more problematic.

8.35 Once two price quotes are available, it should be possible to splice the new good onto an existing or obsolete one. This of course misses the impact of the new item in its initial period, but as discussed below, including such effects is not a trivial exercise. Consider the linking of a good that is likely to be replaced in the market by the new good. For example, a quite new electrical kitchen appliance may use the price index for existing kitchen appliances up to the period of the link, and then the price changes for the new good in subsequent periods. This would create a separate and additional price series for the new good, which augments the sample, as illustrated in Table 8.1. Product *C* is new in period 2 and has no base-period weight. Its price change between periods 1 and 2, had it existed, is assumed to follow the overall index for products *A* and *B*. For period 3 onward a new, linked price series is formed for product *C*, which for period 3 is $101.40 \times 0.985 = 99.88$, and for period 4 is $101.40 \times 0.98 = 99.37$. New revised weights in period 2 show product *C*'s weight to be 20 percent of all products. The new index for period 3 is

$$101.40\,[(0.8\,(101.9/101.4) + 0.2\,(99.88/101.4))]$$

$$= 0.8\,(101.9) + 0.2\,(99.88)$$

$$= 101.50$$

and for period 4,

$$101.40\,[(0.8\,(102.7/101.4) + 0.2\,(99.37/101.4))]$$

$$= 0.8\,(102.7) + 0.2\,(99.37)$$

$$= 102.05.$$

8.36 If product *C* was an evolutionary good replacing product *B*, there would be no need to introduce new weights and no need to augment the sample, as undertaken above. However, since the revolutionary product *C* has no weight in the base period, the splicing requires a revision of the weights at the same time. The selection of the series onto which the new item is spliced, and, in turn, the product groups selected for the weight revision, requires some judgment. Items whose market share is likely to be affected by the introduction of the new good should be selected. If the new good is likely to be responsible for a significant share of revenue, such that it will affect the weights of a broad class of product groups, then there may be a case for a realignment of the overall weighting procedure. Such seismic shifts can of course occur, especially in the communications industry, and for a wide range of industries when regulations are removed or trade barriers are relaxed in less developed economies. In some countries, a new industry or plant can, in itself, amount to sizable proportions of a sector's weights. The change in weights also may be required for *disappearing* goods no longer produced in an economy. As noted in Chapter 15, Section F, chaining and hedonic indices may be appropriate when there is a rapid turnover in such new and obsolete goods. Chaining is an extension of the above procedure and can be used to introduce a new good as soon as it is available for two successive periods.

Table 8.1. Sample Augmentation Example

Products	Base Weight	Revised Weight	Period 1	Period 2	Period 3	Period 4
A	0.6	0.5	100.00	101.00	101.50	102.50
B	0.4	0.3	100.00	102.00	102.50	103.00
All items		*0.8*	*100.00*	*101.40*	*101.90*	*102.70*
C				100.00	98.50	98.00
Spliced *C*		0.2	100.00	101.40	99.88	99.37
Revised all items			*100.00*	*101.40*	*101.50*	*102.05*

8.37 Item augmentation also may be used for evolutionary goods that are likely to be responsible for a substantial share of the market, while not displacing the existing goods. For example, a country with a local brewery and a licensing agreement with a foreign brewery will have joint production of the two beers. The revenue share for beer from the brewery remains the same, but one segment of the market now drinks foreign as opposed to domestic beer. Respondents may be directed to a forced substitution of some of the sample of domestic beers for foreign ones, with the weight remaining the same. This would be similar to a quality adjustment using a noncomparable replacement as discussed in Chapter 7, Section F. Alternatively, the sample may be augmented since there is concern that a smaller sample of domestic beers may not be sufficiently representative. The augmentation process may be similar to that outlined in Table 8.1, with the new foreign beer *C* accounting for 20 percent of the market. Had the advent of foreign beers displaced some of the alcoholic spirits market, then the revision of weights would extend into this product group. As noted in Chapter 7, Section G, chaining and hedonic indices may be appropriate when there is a rapid turnover in new and obsolete goods. With chaining, the good needs to be available only for two successive periods to allow for its introduction.

E. Summary

8.38 The concern in this chapter with sample space and new goods arises out of a very real concern with the dynamic nature of modern markets. New goods and quality changes are far from new issues, and as Triplett (1999) has argued, it has not been demonstrated that the *rate* of new product developments and introductions is much higher now than in the past. However, it is certainly accepted that the *number* of new products and varieties is substantially greater than before. Computer technology provides cost-effective means for collecting and analyzing much larger sets of data. In Chapter 6, the use of handheld computers for data capture was considered, as was the availability of bar-code scanner data. Yet the proper handling of such data requires consideration of issues and methods that go beyond those normally considered for the static intersection universe, which underscores matched samples. In the appendix to this chapter a formal outline of such sampling issues is provided. In this section some of the more important issues are reiterated.

- Where nothing much in the quality and range of available goods changes, there is much that is advantageous to the use of the matched-models methods. It compares like with like from like establishments.
- Statistical metadata systems are needed for quality adjustment issues to help identify the industries in which matching provides few problems. This focuses attention on those that are problematic by collecting and providing information that will facilitate quality adjustment. It also allows for transparency in methods and facilitates retraining.
- Where there is a very rapid turnover in items, such that serious sample depletion takes place quickly, replacements cannot be relied on to replete the sample. Alternative mechanisms, which sample from or use the double universe

of items in each period, are required. These include chained formulations and hedonic indices as discussed in Chapter 7, Section G.

- Some new goods can be treated as evolutionary and incorporated using noncomparable replacements with an associated quality adjustment. The timing of the replacement is critical for both the efficacy of the quality adjustment and the representativity of the index.
- Instructions to respondents on the selection of replacement items are important because they also have a bearing on the representativity of the index. The replacement of obsolete items with newly introduced items leads to difficulties in undertaking quality adjustments, while their replacement with similar items leads to problems of representativity.
- Sample rotation is an extreme form of the use of replacements and is one mechanism for refreshing the sample and increasing its representativity. However, a disadvantage is the possible bias arising from the implicit assumptions underlying the quality adjustment overlap procedure not being met.
- Revolutionary goods may require the augmentation of the sample to make room for new price series and new weighting procedures. The classification of new goods into evolutionary goods and revolutionary goods has a bearing on the strategy for their introduction, directed replacement (substitution), and sample augmentation.

Appendix 8.1: Appearance and Disappearance of Products and Establishments

8.39 In earlier chapters, especially Chapter 5 on sampling, it was generally assumed that the target quantity for estimation could be defined on a fixed set of products. In this appendix the important complications arising from the products and establishments continually changing are considered. The rate of change is rapid in many industries. With this in mind, sampling for price change estimation is a dynamic rather than static problem. Somehow, the prices of new products and in new establishments have to be compared with old ones. *It is important to realize that whatever methods and procedures are used in a price index to handle these dynamic changes, the effects of these procedures will always amount to an explicit or implicit estimation approach for this* ***dynamic universe****.*

Representation of change in a price index[2]

8.40 From a sample selection perspective, there are three ways of handling dynamic changes in an elementary aggregate universe, where varieties and establishments move in and out: (i) by *resampling* the whole elementary aggregate at certain points in time, (ii) by a one-to-one *replacement* of one variety or establishment for another one, and (iii) by *adding and deleting* single observation points (items in establishments) within an index link.

Resampling

8.41 In *resampling*, the old sample is reconsidered as a whole so as to make it representative of the universe in a later period. This does not necessarily mean that all, or even most, sampling units have to be changed, only that a fresh look is taken at the representativity of the whole sample and changes undertaken as appropriate. The methods used for *resampling* could be any of those used for the initial sampling. In the case of probability sampling, it means that every unit belonging to the universe in the later period needs to have a nonzero probability equal to its relative market share of being included in the sample.

8.42 *Resampling* or *sample rotation* is traditionally combined with the overlap method outlined in Chapter 7, Section D. It is similar to the procedure used when combining two links in chained indices. The first period for which the new sample is used is also the last period for which the old sample is used. Thereby, price change estimation is always based on one sample only—the old sample up to the overlap period and the new sample from the overlap period onward, as discussed in further detail below. *Resampling* is the only method that is fully able to maintain the representativity of the sample and, resources permitting, should be undertaken frequently. The necessary frequency depends on the rate of change in a particular product group. It relies, however, on the assumption that the price differences between the old and new items are appropriate estimates of quality differences. At its ex-

[2]A fuller version of this appendix can be found in Dálen (1998).

treme, *resampling* amounts to drawing a new sample in each period and comparing the average price between the samples, instead of the usual procedure of averaging price changes for matched samples. Although being the logical end-point from a representativity point of view, resampling each period would aggravate the quality adjustment problem by its implicit quality adjustment procedure, and, thus, it is not recommended.

Replacement

8.43 A replacement can be defined as an individual successor to a sampled product that either disappeared completely from the market or lost market share in the market as a whole or a specific establishment. Criteria for selecting replacements may differ considerably. There is first the question of when to replace. Usual practices are to do it either when an item disappears completely or when its share of the sales is reduced significantly. Another possible, but less-used, rule would be to replace an item when another variety within the same group, or representative item definition, has become larger with regard to sales, even if the old variety still is sold in significant quantities.

8.44 Second is the question of how to select the replacement item. If the rule for initial selection was most sold or with probability proportionate to (sales) size, then the replacement rule could follow the same selection rule. Alternatively, the replacement could be that item that is most like the old one. The advantage of the former rule is better representativity. The advantage of the most-like rule is, at least superficially, that it might result in a smaller quality adjustment problem.

8.45 It is important to realize that, at least with today's practices, replacements cannot adequately represent new items coming into the market. This is because what triggers a replacement is not the appearance of something new but the disappearance or reduced importance of something old. If the range of varieties in a certain group is increasing, sampling can represent this increase only directly from the set of new varieties, such as in the case of *resampling*.

Adding and deleting

8.46 It is possible to add a new observation point into an elementary aggregate within an index link. If, for example, a new brand or model of a durable was introduced without replacing any particular old model, it would be desirable to add it to the sample starting from the time of its introduction. In order to accommodate this new observation into the index system, its reference price needs to be imputed. A practical way to do this is to divide its price in the month of introduction by the price index of all other items in the elementary aggregate from the reference period to the month of introduction. In this way, its effect on the index for months up to the introduction month will be neutral.

8.47 Similarly, an item that disappears could just be deleted from the sample without replacement. Price change can then be computed over the remaining items. If no further action is taken, this means that the price change for the deleted item that was measured up to the month prior to deletion will be disregarded from the month of deletion. This may or may not be desirable, depending on the circumstances in the particular product group.

Formulating an operational target in a dynamic universe

8.48 A rigorous approach to the problem of statistical estimation requires an *index estimation strategy* that includes both the *operational target of measurement* and the *sampling strategy* (design and estimator) needed for estimating this target. This strategy would have to consist of the following components:

(i) A definition of the universe of *transactions* or *observation points* (usually a product variety in an establishment) in each of the two time periods between which we want to estimate price change;
(ii) A list of all *variables* defined on these units. These variables should include prices and quantities (number of units sold at each price), but also all relevant price-determining characteristics of the products (and possibly also of the establishments)—the price basis;
(iii) The *target algorithm* (index formula) that combines the variable values defined in (ii) for the observation points in the universe defined in (i) into a single value;
(iv) Procedures used for *initial sampling* of items and establishments from the universe defined in (i);

(v) Procedures within the time span for *replacing, resampling and adding or deleting* observations; and

(vi) The *estimation algorithm* (index formula) applied to the sample with the purpose of minimizing the expected error of the sample estimate compared with the target algorithm under (iii). This algorithm, in principle, needs to consider all the procedures taken in replacement and *resampling* situations, including procedures for quality adjustment.

8.49 The kind of rigorous strategy outlined above is generally not used in practical index construction because of its complexity, though its required information system was discussed in Section C.1. A few comments on such possible strategies are made below.

A two-level aggregation system

8.50 A starting point for discussing this objective is a two-level structuring of the universe of commodities and establishments considered in the scope of a price index. These levels are

- The *aggregate* level. At this level there is a fixed structure of item groups $h = 1, \ldots, H$ (or perhaps a fixed cross-structure of item groups by regions or establishment types) within an index link. New goods and services for updating the universe of commodities would be defined in terms of new groups at this level and moved into the index only in connection with a new index link.
- The *elementary* level. Within this level the aim is to capture the properties of a changing universe in the index by comparing new and old items. The micro comparison from s to t must be defined so that new products and establishments enter into the market and old products and establishments disappear from the market.

The common starting point for three alternative approaches at the elementary level is a pure price formulation of price change from period s to period t at the aggregate level:

$$\text{(A8.1)}\; I^{st} = \frac{\sum_h Q_h P_h^t}{\sum_h Q_h P_h^s} = \sum_h W_h^s I_h^{st} \,,$$

where $W_h^s = \dfrac{Q_h P_h^s}{\sum_h Q_h P_h^s}$ and $I_h^{st} = \dfrac{P_h^t}{P_h^s}$.

The quantities, Q_h, are for $h = 1 \ldots H$ item groups from any period or functions of quantities from several periods, for example, a symmetric average of the base and current periods s and t. Special cases of such a pure price index are the Laspeyres ($Q_h = Q_h^s$), Paasche ($Q_h = Q_h^t$), Edgeworth ($Q_h = Q_h^s + Q_h^t$), and Walsh ($Q_h = [\,Q_h^t\, Q_h^t\,]^{1/2}$) price indices outlined in Chapters 15 through 17. Alternative formulations for an elementary-level estimation strategy now enter in the definition of I_h^{st}. As a further common starting point the set of items or establishments belonging to h in period u *(*= s or t*) are defined* as Ω_h^u. The concept of an *observation point* is introduced, usually a tightly specified item in a specific establishment, such that, say, $\Omega_h^u = \{1, \ldots, j, \ldots, N_h^u\}$. For each observation point $j \in \Omega_h^u$, there is a price p_j^u and a quantity sold q_j^u. There are now three possibilities for defining the operational target.

The intersection universe

8.51 The elementary index is defined over the intersection universe, that is, only over observation points existing in both s and t. This index may also be called the *identical units index*. It is equivalent to starting out with the observation points existing in s and then dropping (deleting) missing or disappearing points. An example of such an index is:

$$\text{(A8.2)}\quad I_h^{st} = \frac{\sum_{j \in \Omega_h^s \cap \Omega_h^t} q_j p_j^t}{\sum_{j \in \Omega_h^s \cap \Omega_h^t} q_j p_j^s} \,.$$

The intersection universe decreases successively over time as fewer matches are found for each long-run comparison between s and t, s and $t + 1$, s and $t + 2$, etc., until it eventually becomes empty. An attraction of the intersection universe is that there are, by definition, no replacements involved in this target and, thus, normally no quality adjustments. If the identical units index is combined with a short index link, followed by *resampling* from the universe in a later period, sampling from this universe is a perfectly reasonable strategy, as long as the assumptions implicit in the overlap procedure, that the price differences at that point in time reflect the quality differences, are valid.

The double universe

8.52 The polar opposite approach to the intersection universe is to consider P_h^s and P_h^t as average prices defined over two separately defined universes in the two periods. A double universe target could then be considered: one universe in period *s* and another in period *t*. This seems to be a natural way of defining the target, since both time periods should be of equal status, and all products existing in any of these should be taken into account. The difficulty with this approach is that the two universes are rarely comparable in terms of quality. Some kind of adjustment for average quality change would need to be brought into the index. The natural definition of the average prices involved in this approach is as unit values. This would lead to the following definition of a *quality-adjusted unit value index*:

$$\text{(A8.3)} \quad I_h^{st} = \bar{P}_h^t \Big/ \bar{P}_h^s g_h^{st} \quad ,$$

$$\text{where } \bar{P}_h^t = \frac{\sum_{j \in \Omega_h^t} q_j^t p_j^t}{\sum_{j \in \Omega_h^t} q_j^t} \text{ and } \bar{P}_h^s = \frac{\sum_{j \in \Omega_h^s} q_j^s p_j^s}{\sum_{j \in \Omega_h^s} q_j^s} .$$

In equation (A8.3), g_h^{st} is the average quality change in *h* (also interpretable as a *quality index*), which needs further definition. For example, g_h^{st} could be thought of as a hedonic adjustment procedure, where characteristics are held constant. Equation (A8.3) was outlined in Chapter 7, Section E, in forms that include explicit hedonic quality adjustments, g_h^{st}, but as part of Laspeyres, Paasche, Fisher, and Törnqvist indices. This operational target is attractive for commodities where the rate of turnover of varieties is very fast, but average quality changes slowly, or reliable estimates of quality changes can be made. The commonly used representative item method is not really compatible with a double-universe target. It implicitly focuses on preselected primary sampling units that are used for both periods *s* and *t*.

The replacement universe

8.53 Neither sampling from the intersection nor sampling from the double universe bears a close resemblance to usual practices for constructing price indices. In particular, the representative item method combined with one-to-one replacements, which is the most common sampling method used in practice, needs a rationalization in terms of operational targets, which differs from these alternatives. Such a rationalization of sampling from a *replacement universe* is considered below.

Definition 1a: For each $j \in \Omega_h^s$ and $j \notin \Omega_h^t$ we define replacement items $a_j \in \Omega_h^t$ whose price enters into *j*'s place in the formula. (For $j \in \Omega_h^s$ and $j \in \Omega_h^t$, $a_j = j$.) In addition to a replacement, a quality change from *j* to a_j is included, which gives rise to a quality adjustment factor g_j, interpreted as the factor with which p_j^s must be multiplied for the producer to be indifferent between producing items *j* and a_j at prices p_j^s and $p_{a_j}^t$.

$$\text{(A8.4)} \quad I_h^{st} = \frac{\sum_{j \in \Omega_h^s} q_j p_{a_j}^t}{\sum_{j \in \Omega_h^s} q_j p_j^s g_j} \quad .$$

However, this first step toward an operational use of the formula requires, first, a need to define g_j, possibly arising from a hedonic regression as described in Chapter 7, Section G.2. Second, there is a need to define a_j. A natural procedure is to use a *dissimilarity function* from *j* to a_j. The notation $d(j, a_j)$ is introduced for this function. The common procedure of choosing the most similar item in cases of replacement now corresponds to minimizing the dissimilarity function. However, some further specifications need to be made. *When* is the replacement defined to take place? In practice, this ought to be done when the first chosen variety is no longer representative. Mathematically, this could be defined as Definition 1b.

Definition 1b: Observation point *j* should be replaced in the first period in which $q_j^t < cq_j^s$, where *c* is a suitably chosen constant between 0 and 1 (a modification would be required for seasonal items).

The *choice* of replacement point would then be governed by a rule such as Definition 1c.

Definition 1c: a_j should be chosen so that $d(j, a_j)$ is minimized for j.

However, since some priority should be given to observation points that are important in terms of quantities or values, Definition 1c can then be modified to become Definition 1d.

Definition 1d: a_j should be chosen so that $d(j,a_j)/q^t_{a_j}$ is minimized for j. (Some other function of $d(.)$ and $q^t_{a_j}$ could be chosen in its place.)

8.54 The dissimilarity function needs to be specified; it may depend on the item group h. In general this must be some kind of metric defined on the set of characteristics of the product and establishment in question. For example, priority could be given to its dissimilarity either to "same establishment" or "same product," which could easily be worked into such a metric. A more troublesome concern is the inclusion of as many new points in Ω^t_h as possible in the index definition, to make the sample representative. As Definitions 1a–d now stand, the same new point could replace many predecessors, whereas there may be many new points that will not be sampled unless there is a need for a replacement. This shortcoming of the replacement universe is an inherent trait in the replacement method as such. The replacement method is designed only to maintain the representativity of the old sample, not that of the new sample.