20 ELEMENTARY INDICES

Introduction

20.1 In all countries, the calculation of a consumer price index (CPI) proceeds in two (or more) stages. In the first stage of calculation, elementary price indices are estimated for the elementary expenditure aggregates of a CPI. In the second and higher stages of aggregation, these elementary price indices are combined to obtain higher-level indices using information on the expenditures on each of the elementary aggregates as weights. An elementary aggregate consists of the expenditures on a small and relatively homogeneous set of products defined within the consumption classification used in the CPI. Samples of prices are collected within each elementary aggregate, so that elementary aggregates serve as strata for sampling purposes.

20.2 Data on the expenditures, or quantities, of the different goods and services are typically not available within an elementary aggregate. As there are no quantity or expenditure weights, most of the index number theory outlined in Chapters 15 to 19 is not directly applicable. As was noted in Chapter 1, an elementary price index is a more primitive concept that relies on price data only.

20.3 The question of what is the most appropriate formula to use to estimate an elementary price index is considered in this chapter. The quality of a CPI depends heavily on the quality of the elementary indices, which are the basic building blocks from which CPIs are constructed.

20.4 As is explained in Chapter 6, compilers have to select representative products within an elementary aggregate and then collect a sample of prices for each of the representative products, usually from a sample of different outlets. The individual products for which prices are actually collected are described as the *sampled products*. Their prices are collected over a succession of time periods. An elementary price index is therefore typically calculated from two sets of matched price observations. In most of this chapter[1], it is assumed that there are no missing observations and no changes in the quality of the products sampled so that the two

[1] The problem of sample attrition and the lack of matching over time is discussed briefly in the context of classification issues in paragraphs 20.23 to 20.37.

sets of prices are perfectly matched. The treatment of new and disappearing products, and of quality change, is a separate and complex issue that is discussed in detail in Chapters 7, 8 and 21 of this manual.

20.5 Even though quantity or expenditure weights are usually not available to weight the individual elementary price quotes, it is useful to consider an *ideal framework* where expenditure information is available. This is done in the next section. The problems involved in aggregating narrowly defined price quotes over time are also discussed in that section. Thus the discussion provides a theoretical target for "practical" elementary price indices that are constructed using only information on prices.

20.6 Paragraphs 20.23 to 20.37 provide some discussion about the difficulties involved in picking a suitable level of disaggregation for the elementary aggregates. Should the elementary aggregates have a regional dimension in addition to a product dimension? Should prices be collected from retail outlets or from households? These are the types of question discussed in this section.

20.7 Paragraphs 20.38 to 20.45 introduce the main elementary index formulae that are used in practice, and paragraphs 20.46 to 20.57 develop some numerical relationships between the various indices.

20.8 Chapters 15 to 17 develop the various approaches to index number theory when information on both prices and quantities is available. It is also possible to develop axiomatic, economic or sampling (stochastic) approaches to elementary indices, and these three approaches are discussed below in paragraphs 20.58 to 20.70, 20.71 to 20.86, and 20.87, respectively.

20.9 Paragraphs 20.88 to 20.99 look at some of the recent scanner data literature that computes elementary aggregates using both price and quantity information.

20.10 Paragraphs 20.100 to 20.111 develop a simple statistical approach to elementary indices that resembles a highly simplified hedonic regression model. The concluding section presents an overview of the various results[2].

Ideal elementary indices

20.11 The aggregates covered by a CPI or a producer price index (PPI) are usually arranged in the form of a tree-like hierarchy, such as the Classification of Individual Consumption according to Purpose (COICOP)[3] or the Nomenclature générale des Activités économiques dans les Communautés Européennes [General Industrial Classification of Economic Activities within the European Communities] (NACE). Any aggregate is a set of economic transactions pertaining to a set of commodities over a specified time period. Every economic transaction relates to the change of ownership of a specific, well-defined commodity (good or service) at a particular place and date, and comes with a quantity and a price. The price index for an aggregate is calculated as a weighted average of the price indices for the sub-aggregates, the (expenditure or sales) weights and type of average being determined by the index formula. One can descend in such a hierarchy as far as available information allows the weights to be decomposed. The lowest-level aggregates are called *elementary aggregates*. They are basically of two types:

- those for which all detailed price and quantity information is available;
- those for which the statistician, considering the operational cost or the response burden of getting detailed price and quantity information about all the transactions, decides to make use of a representative sample of commodities or respondents.

20.12 The practical relevance of studying this topic is large. Since the elementary aggregates form the building blocks of a CPI or a PPI, the choice of an inappropriate formula at this level can have a tremendous impact on the overall index.

20.13 In this section, it will be assumed that detailed price and quantity information for all transactions pertaining to the elementary aggregate for the two time periods under consideration is available. This assumption allows us to define an ideal elementary aggregate.

---

[2] This chapter draws heavily on the recent contributions of Dalén (1992), Balk (1994; 1998b; 2002) and Diewert (1995a; 2002c).

[3] Triplett (2003, p. 160) is quite critical of the COICOP classification scheme and argues that economic theory and empirical analysis should be used to derive a more appropriate CPI classification scheme. It is nevertheless very difficult to coordinate a classification scheme that can be used by all countries.

Subsequent sections will relax this strong assumption about the availability of detailed price and quantity data on transactions, but it is necessary to have a theoretically ideal target for the "practical" elementary index.

20.14 The detailed price and quantity data, although perhaps not available to the statistician, are in principle available in the outside world. It is frequently the case that at the respondent level (i.e., at the outlet or firm level) some aggregation of the individual transactions information has been executed, usually in a form that suits the respondent's financial or management information system. This level of information that is determined by the respondent could be called the *basic information level*. It is, however, not necessarily the finest level of information that could be made available to the price statistician. One could always ask the respondent to provide more disaggregated information. For instance, instead of monthly data one could ask for weekly data; or, whenever appropriate, one could ask for regional instead of global data; or one could ask for data according to a finer commodity classification. The only natural barrier to further disaggregation is the individual transaction level.[4]

20.15 It is now necessary to discuss a problem that arises when detailed data on individual transactions are available, either at the level of the individual household or at the level of an individual outlet. Recall that Chapter 15 introduces the price and quantity indices, $P(p^0,p^1,q^0,q^1)$ and $Q(p^0,p^1,q^0,q^1)$. These (bilateral) price and quantity indices decompose the value ratio V^1/V^0 into a price change part $P(p^0,p^1,q^0,q^1)$ and a quantity change part $Q(p^0,p^1,q^0,q^1)$. In this framework, it is taken for granted that the period t price and quantity for commodity i, p_i^t and q_i^t respectively, are well defined. These definitions are not, however, straightforward since individual consumers may purchase the *same* item during period t *at different prices*. Similarly, if one considers the sales of a particular shop or outlet that sells to consumers, the same item may sell at very different prices during the course of the period. Hence before a traditional bilateral price index of the form $P(p^0,p^1,q^0,q^1)$ considered in previous chapters of this manual can be applied, a non-trivial time aggregation problem must be resolved in order to obtain the basic prices p_i^t and quantities q_i^t that are the components of the price vectors p^0 and p^1 and the quantity vectors q^0 and q^1.

[4] See Balk (1994).

20.16 Walsh[5] and Davies (1924; 1932) suggested a solution to this time aggregation problem: in their view, the appropriate quantity at this very first stage of aggregation is the *total quantity purchased* of the narrowly defined item and the corresponding price is the value of purchases of this item divided by the total amount purchased, which is a *narrowly defined unit value*.

In more recent times, other researchers have adopted the Walsh and Davies solution to the time aggregation problem.[6] Note that this solution has the following advantages:

- The quantity aggregate is intuitively plausible, being the total quantity of the narrowly defined item purchased by the household (or sold by the outlet) during the time period under consideration.
- The product of the price times quantity equals the total value purchased by the household (or sold by the outlet) during the time period under consideration.

20.17 The above solution to the time aggregation problem will be adopted as the concept for the price and quantity at this very first stage of aggregation. This leaves open the question of how long the time period should be over which the unit value is calculated. This question will be considered in the following section.

20.18 Having decided on an appropriate theoretical definition of price and quantity for an item at the very lowest level of aggregation (i.e., a narrowly defined unit value and the total quantity sold of that item at the individual outlet, or the total quantity purchased by a single household or a group of households), it is necessary to consider how to aggregate these narrowly defined elementary prices and quantities into an overall elementary aggregate.

---

[5] Walsh explained his reasoning as follows:

> Of all the prices reported of the same kind of article, the average to be drawn is the arithmetic; and the prices should be weighted according to the relative mass quantities that were sold at them (Walsh (1901, p. 96)).
>
> Some nice questions arise as to whether only what is consumed in the country, or only what is produced in it, or both together are to be counted; and also there are difficulties as to the single price quotation that is to be given at each period to each commodity, since this, too, must be an average. Throughout the country during the period a commodity is not sold at one price, nor even at one wholesale price in its principal market. Various quantities of it are sold at different prices, and the full value is obtained by adding all the sums spent (at the same stage in its advance towards the consumer), and the average price is found by dividing the total sum (or the full value) by the total quantities (Walsh (1921a, p. 88)).

[6] See for example Szulc (1987, p. 13), Dalén (1992, p. 135), Reinsdorf (1994), Diewert (1995a, pp. 20-21), Reinsdorf and Moulton (1997), Balk (2002) and Richardson (2003).

Suppose that there are M lowest-level items or specific commodities in this chosen elementary category. Denote the period t quantity of item m by q_m^t and the corresponding time-aggregated unit value by p_m^t for $t = 0,1$ and for items $m = 1,2,...,M$. Define the period t quantity and price vectors as $q^t \equiv [q_1^t,q_2^t,...,q_M^t]$ and $p^t \equiv [p_1^t,p_2^t,...,p_M^t]$ for $t = 0,1$. It is now necessary to choose a theoretically ideal index number formula $P(p^0,p^1,q^0,q^1)$ that will aggregate the individual item prices into an overall aggregate price relative for the M items in the chosen elementary aggregate. This problem of choosing a functional form for $P(p^0,p^1,q^0,q^1)$ is identical to the overall index number problem that is addressed in Chapters 15–17. In these previous chapters, four different approaches to index number theory are studied, and specific index number formulae are seen as being "best" from each perspective. From the viewpoint of *fixed basket approaches*, the Fisher (1922) and Walsh (1901) price indices, P_F and P_W, appear to be "best". From the viewpoint of the *test approach*, the Fisher index appears to be "best". From the viewpoint of the *stochastic approach* to index number theory, the Törnqvist-Theil (1967) index number formula P_T emerges as being "best". Finally, from the viewpoint of the *economic approach* to index number theory, the Walsh price index P_W, the Fisher ideal index P_F and the Törnqvist-Theil index number formula P_T are all regarded as being equally desirable. It is also shown that these three index number formulae numerically approximate each other very closely, and so it does not matter very much which of these alternative indices is chosen.[7] Hence, the *theoretically ideal elementary index number formula* is taken to be one of the three formulae $P_F(p^0,p^1,q^0,q^1)$, $P_W(p^0,p^1,q^0,q^1)$ or $P_T(p^0,p^1,q^0,q^1)$ where the period t quantity of item m, q_m^t, is the total quantity of that narrowly defined item purchased by the household during period t (or sold by the outlet during period t) and the corresponding price for item m is p_m^t, the time-aggregated unit value, for $t = 0,1$ and for items $m = 1,2,...,M$.[8]

20.19 Various "practical" elementary price indices are defined in paragraphs 20.38 to 20.45.

[7] Theorem 5 in Diewert (1978, p. 888) shows that P_F, P_T and P_W approximate each other to the second order around an equal price and quantity point; see Diewert (1978, p. 894), Hill (2000) and Chapter 19 for some empirical results.

[8] Of course, all these ideal elementary index number formulae require current period quantity (or expenditure) weights and thus are not usually "practical" formulae that can be used to produce the usual type of month-to-month CPI. Nevertheless, as statistical agencies introduce superlative indices on a retrospective basis, it may be possible to obtain more current information on weights, at least at higher levels of aggregation; see Greenlees (2003). Gudnason (2003, p. 16) also gives some examples where the Icelandic CPI obtains enough information to be able to calculate some elementary indices using a superlative formula. In any case, a target index is required at the elementary level just as one is required at higher levels of aggregation.

These indices do not have quantity weights and thus are functions only of the price vectors p^0 and p^1, which contain time-aggregated unit values for the M items in the elementary aggregate for periods 0 and 1. Thus when a practical elementary index number formula, say $P_E(p^0,p^1)$, is compared to an ideal elementary price index, say the Fisher price index $P_F(p^0,p^1,q^0,q^1)$, then obviously P_E will differ from P_F because the prices are not weighted according to their economic importance in the practical elementary formula.[9] Call this difference between the two index number formulae *formula approximation error*.

20.20 Practical elementary indices are subject to other types of error as well:

- The statistical agency may not be able to collect information on all M prices in the elementary aggregate; i.e., only a sample of the M prices may be collected. Call the resulting divergence between the incomplete elementary aggregate and the theoretically ideal elementary index the *sampling error*.
- Even if a price for a narrowly defined item is collected by the statistical agency, it may not be equal to the theoretically appropriate time-aggregated unit value price. This use of an inappropriate price at the very lowest level of aggregation gives rise to *time aggregation error*.[10]
- The statistical agency may classify certain distinct products as being essentially equivalent and this may result in *item aggregation error*. For example, when the same product is sold in different package sizes, only the per unit price may be collected over the different package sizes. As another example, small quality differences between products may be ignored.
- The unit value for a particular item may be constructed by aggregating over all households in a region or a certain demographic class or by aggregating over all outlets or shops that sell the item in a particular region. This may give rise to an

[9] Hausman (2002, p. 14) also noted the importance of collecting quantity data along with price data at the elementary level so that more accurate quality change adjustments can be made by statistical agencies.

[10] Many statistical agencies send price collectors to various outlets on certain days of the month to collect list prices of individual items. Usually, price collectors do not work on weekends, when many sales take place. Thus the collected prices may not be fully representative of all transactions that occur. These collected prices can be regarded as approximations to the time-aggregated unit values for those items, but they are only approximations.

aggregation over agents or entities error.

20.21 The problems of aggregation and classification are discussed in more detail in paragraphs 20.23 to 20.37.

20.22 The five main elementary index number formulae are defined in paragraphs 20.30 to 20.45, and in paragraphs 20.46 to 20.57 various numerical relationships between these five indices are developed. Paragraphs 20.58 to 20.86 develop the axiomatic and economic approaches to elementary indices, and the five main elementary formulae used in practice will be evaluated in the light of these approaches.

Aggregation and classification problems for elementary aggregates

20.23 Hawkes and Piotrowski (2003) note that the definition of an elementary aggregate involves aggregation over *four* possible dimensions:[11]

- a *time* dimension; i.e., the item unit value could be calculated for all item transactions for a year, a month, a week, or a day;
- a *spatial* dimension; i.e., the item unit value could be calculated for all item transactions in the country, province or state, city, neighbourhood, or individual location;
- a *product* dimension; i.e., the item unit value could be calculated for all item transactions in a broad general category (e.g., food), in a more specific category (e.g., margarine), for a particular brand (ignoring package size) or for a particular narrowly defined item (e.g., a particular AC Nielsen universal product code);
- a *sectoral* (or *entity* or *economic agent*) dimension; i.e., the item unit value could be calculated for a particular class of households or a particular class of outlets.

20.24 Each of the above dimensions for choosing the domain of definition for an elementary aggregate will be discussed in turn.

[11] Hawkes and Piotrowski (2003, p. 31) combine the spatial and sectoral dimensions into the spatial dimension. They also acknowledge the pioneering work of Theil (1954), who identified three dimensions of aggregation: aggregation over individuals, aggregation over commodities, and aggregation over time.

20.25 As the time period is compressed, several problems emerge:

- Purchases (by households) and sales (by outlets) become erratic and sporadic. Thus the frequency of unmatched purchases or sales from one period to the next increases and in the limit (choose the time period to be one minute), nothing will be matched and bilateral index number theory fails.[12]
- As the time period becomes shorter, chained indices exhibit more "drift"; i.e., if the value at the end of a chain of periods reverts to the value in the initial period, the chained index does not revert to unity. As is discussed in paragraphs 15.76 to 15.97 of Chapter 15, it is only appropriate to use chained indices when the underlying price and quantity data exhibit relatively smooth trends. When the time period is short, seasonal fluctuations[13] and periodic sales and advertising campaigns[14] can cause prices and quantities to oscillate (or "bounce", to use Szulc's (1983, p. 548) term), and hence it is not appropriate to use chained indices under these circumstances. If fixed base indices are used in this short time period situation, then the results will usually depend very strongly on the choice of the base period. In the seasonal context, not all commodities may even be in the marketplace during the chosen base period.[15] All these problems can be mitigated by choosing a longer time period so that trends in the data will tend to dominate the short-term fluctuations.
- As the time period contracts, virtually all goods become *durable* in the sense that they yield services not only for the period of purchase but for subsequent periods. Thus the period of purchase or acquisition becomes different from the periods of use, leading to many complications.[16]

---

[12] This point is noted in paragraphs 15.65 to 15.71 of Chapter 15 in relation to the Divisia index. David Richardson (2003, p. 51) also made the point: "Defining items with a finer granularity, as is the case if quotes in different weeks are treated as separate items, results in more missing data and more imputations."

[13] See Chapter 22 for a monthly seasonal example where chained month-to-month indices are useless.

[14] See Feenstra and Shapiro (2003) for an example of a weekly superlative index that exhibits massive chain drift. Richardson (2003, pp. 50-51) discusses the issues involved in choosing weekly unit values versus monthly unit values.

[15] See Chapter 22 for suggested solutions to these seasonality problems.

[16] See Chapter 23 for more material on the possible CPI treatment of durable goods.

- As the time period contracts, users will not be particularly interested in the short-term fluctuations of the resulting index and there will be demands for smoothing the necessarily erratic results. Put another way, users will want the many, say, weekly or daily movements in the index to be summarized as monthly or quarterly movements in prices. Hence from the viewpoint of meeting the needs of users, there will be relatively little demand for high-frequency indices.

In view of the above considerations, it is recommended that the index number time period be at least four weeks or a month.[17]

20.26 It is also necessary to choose the spatial dimension of the elementary aggregate. Should item prices in each city or region be considered as separate aggregates or should a national item aggregate be constructed? Obviously, if it is desired to have regional CPIs which aggregate up to a national CPI, then it will be necessary to collect item prices by region. It is not clear, however, how fine the "regions" should be. They could be as fine as a grouping of households in a postal code area or as individual outlets across the country.[18] There does not seem to be a clear consensus on what the optimal degree of spatial disaggregation should be.[19] Each statistical agency will have to make its own judgements on

[17] If there is very high inflation in the economy (or even hyperinflation), then it may be necessary to move to weekly or even daily indices. Also, it should be noted that some index number theorists feel that new theories of consumer behaviour should be developed that could use weekly or daily data: "Some studies have endorsed unit values to reduce high frequency price variation, but this implicitly assumes that the high frequency variation represents simply noise in the data and is not meaningful in the context of a COLI. That is debatable. We need to develop a theory that confronts the data, not truncate the data to fit the theory" (Triplett (2003, p. 153)). Until such new theories are adequately developed, however, a pragmatic approach is to define the item unit values over months or quarters rather than days or weeks.

[18] Iceland no longer uses regional weights but uses individual outlets as the primary geographical unit; see Gudnason (2003, p. 18).

[19] William J. Hawkes and Frank W. Piotrowski note that it is quite acceptable to use national elementary aggregates when making international comparisons between countries:

> When we try to compare egg prices across geography, however, we find that lacing across outlets won't work, because the eyelets on one side of the shoe (or outlets on one side of the river) don't match up with those on the other side. Thus, in making interspatial comparisons, we have no choice but to aggregate outlets all the way up to the regional (or, in the case of purchasing power parities, national) level. We have no hesitation about doing this for interspatial comparisons, but we are reluctant to do so for intertemporal ones. Why is this? (Hawkes and Piotrowski (2003, pp. 31-32)).

An answer to their question is that it is preferable to match like with like as closely as possible. This leads statisticians to prefer the finest possible level of aggregation, which, in the case of intertemporal comparisons, would be the individual household or the individual outlet. In making cross-region comparisons, however, matching is not

the matter of the optimal degree of spatial disaggregation, taking into account the costs of data collection and the demands of users for a spatial dimension for the CPI.

20.27 How detailed should the product dimension be? The possibilities range from regarding all commodities in a general category as being equivalent to regarding only commodities in a particular package size made by a particular manufacturer or service provider as being equivalent. All things being equal, Triplett (2002) stresses the advantages of matching products at the most detailed level possible, since this will prevent quality differences from clouding the period-to-period price comparisons. This is sensible advice, but then what are the drawbacks to working with the finest possible commodity classification? The major drawback is that the finer the classification, the more difficult it will be to match the item purchased or sold in the base period to the same item in the current period. Hence, the finer the product classification, the smaller will be the number of matched price comparisons that are possible. This would not be a problem if the unmatched prices followed the same trend as the matched ones in a particular elementary aggregate; but in at least some circumstances, this will not be the case.[20] The finer the classification system, the more work (in principle) there will be for the statistical agency to adjust for quality or impute the prices that do not match. Choosing a relatively coarse classification system leads to a very cost-efficient system of quality adjustment (i.e., essentially no explicit quality adjustment or imputation is done for the prices that do not exactly match), but it may not be very accurate. Thus all things considered, it seems preferable to choose the finest possible classification system.

20.28 The final issue in choosing a classification scheme is the issue of choosing a sectoral dimension; i.e., should the unit value for a particular item be calculated for a particular outlet or a particular household, or for a class of outlets or households?

20.29 Before the above question can be answered, it is necessary to ask whether the

---

possible unless regional item aggregates are formed, as Hawkes and Piotrowski point out above.

[20] Silver and Heravi (2001a; 2001b; 2002; 2003, p. 286) and Koskimäki and Vartia (2001) stressed this point and presented empirical evidence to back up their point. Feenstra (1994) and Balk (2000b) developed some economic theory based methods to deal with the introduction of new items.

individual outlet or the individual household is the appropriate finest level of entity classification. If the economic approach to the CPI is taken, then the individual household is the appropriate finest level of entity classification.[21] Obviously, a single household will not work very well as the basic unit of entity observation because of the sporadic nature of many purchases by an individual household; i.e., there will be tremendous difficulties in matching prices across periods for individual households. For a grouping of households that is sufficiently large, however, it does become feasible in theory to use the household as the entity classification, rather than the outlet as is usually done. It is not usual to use households because of the costs and difficulties involved in collecting individual household data on prices and expenditures.[22] Price information is usually collected from retail establishments or outlets that sell mainly to households. Matching problems are mitigated (but not eliminated) using this strategy because the retail outlet generally sells the same items on a continuing basis.

20.30 If expenditures by all households in a region are aggregated together, will they equal sales by the retail outlets in the region? Under certain conditions, the answer to this question is "yes". The conditions are that the outlets do not sell any items to purchasers who are not local households (no regional exports or sales to local businesses or governments) and that the regional households do not make any purchases of consumption items other than from the local outlets (no household imports or transfers of commodities to local households by governments). Obviously, these restrictive conditions will not be met in practice, but they may hold as a first approximation.

20.31 The effects of *regional aggregation* and *product aggregation* can be examined,

---

[21] This point has been made emphatically by two authors in the recent book on scanner data and price indices:

> In any case, unit values across stores are not the prices actually faced by households and do not represent the per period price in the COLI, even if the unit values are grouped by type of retail outlet (Triplett (2003, pp. 153-154)).
>
> Furthermore, note that the relationship being estimated is not a proper consumer demand function but rather an 'establishment sales function'. Only after making further assumptions – for example, fixing the distribution of consumers across establishments – is it permissible to jump to demand functions (Ley (2003, p. 380)).

[22] However, it is not impossible to collect accurate household data in certain circumstances; see Gudnason (2003), who pioneered a receipts methodology for collecting household price and expenditure data in Iceland.

thanks to a recent study by Koskimäki and Ylä-Jarkko (2003). This study used scanner data for the last week in September 1998 and September 2000 on butter, margarine and other vegetable fats, vegetable oils, soft drinks, fruit juices and detergents that were provided by the AC Nielsen company for Finland. At the finest level of item classification (the AC Nielsen Universal Product Code), the number of individual items in the sample was 1,028. The total number of outlets in the sample was 338. Koskimäki and Ylä-Jarkko then considered four levels of spatial disaggregation:

- the entire country (1 level);
- provinces (4 levels);
- AC Nielsen regions (15 levels);
- individual outlets (338 levels).

They also considered four levels of product disaggregation:

- the COICOP 5-digit classification (6 levels);
- the COICOP 7-digit classification (26 levels);
- the AC Nielsen brand classification (266 levels);
- the AC Nielsen individual Universal Product Code (1,028 distinct products).

20.32 In order to illustrate the ability to match products over the two-year period as a function of the degree of fineness of the classification, Koskimäki and Ylä-Jarkko (2003, p. 10) presented a table showing that the proportion of transactions that could be matched across the two years fell steadily as the fineness of the classification scheme increased. At the highest level of aggregation (the national and COICOP 5-digit), all transactions could be matched over the two-year period, but at the finest level of aggregation (338 outlets times 1,028 individual products or 347,464 classification cells in all), only 61.7 per cent of the value of transactions in 2000 could be matched back to their 1998 counterparts. Koskimäki and Ylä-Jarkko's Table 7 is reproduced as Table 20.1.

Table 20.1 Proportion of transactions in 2000 that could be matched to 1998

	COICOP 5-digit	**COICOP 7-digit**	**AC Nielsen brand**	**AC Nielsen Universal Product Code**
Country	1.000	1.000	0.982	0.801

Province	1.000	1.000	0.975	0.774
AC Nielsen region	1.000	1.000	0.969	0.755
Individual outlet	0.904	0.904	0.846	0.617

20.33 For each of the above 16 levels of product and regional disaggregation, for the products that were available in September 1998 and September 2000, Koskimäki and Ylä-Jarkko (2003, p. 9) calculated Laspeyres and Fisher price indices. Their results are reproduced below as Tables 20.2 and 20.3.

Table 20.2 Laspeyres price indices by type of classification, September 1998–2000

	COICOP 5-digit	***COICOP* 7-digit**	**AC Nielsen brand**	**AC Nielsen Universal Product Code**
Country	1.079	1.031	1.046	1.023
Province	1.078	1.031	1.048	1.023
AC Nielsen region	1.078	1.031	1.048	1.025
Individual outlet	1.086	1.040	1.060	1.028

Table 20.3 Fisher price indices by type of classification, September 1998–2000

	COICOP 5-digit	**COICOP 7-digit**	**AC Nielsen brand**	**AC Nielsen Universal Product Code**
Country	1.080	1.032	1.048	1.015
Province	1.079	1.031	1.048	1.014
AC Nielsen region	1.079	1.030	1.047	1.014
Individual outlet	1.089	1.034	1.049	1.011

20.34 Some of the trends in Tables 20.2 and 20.3 can be explained. As the product classification is made more fine, the indices tend to fall.[23] This indicates that the new products entering the sample tend to be more expensive than the continuing products. The differences in the COICOP 5-digit results and the AC Nielsen Universal Product Code results are very big indeed and indicate that it is probably best to work at the finest level of product disaggregation, even if there is the possibility of bias because of neglecting new products. This possible bias would have to be very substantial to overturn a recommendation to work at the finest level of product disaggregation.

20.35 As the regional classification is made finer, there is a tendency for the Laspeyres indices to become larger. This can be explained by purchasers switching to the lowest-cost outlets so that the item unit values will be smaller the higher the degree of aggregation. Put another way, the Laspeyres indices calculated at the outlet level are subject to a certain amount of outlet substitution bias (if one is willing to regard this phenomenon as a bias).

20.36 What is striking in the Tables 20.1 to 20.3 are the differences between the Laspeyres and Fisher indices at the finer levels of aggregation. For the very finest level of aggregation, the Fisher at 1.011 is 1.7 percentage points below the corresponding Laspeyres at 1.028. Thus at the finest level of aggregation, the Laspeyres for this Finnish data set has a *representativity* or *elementary substitution bias* of about 0.85 percentage points per year.

20.37 Note that the above index number comparisons are free of chain drift problems since they make direct comparisons across the two years. They should also be free of seasonal problems, since the last week in September 1998 is compared with the last week of September 2000.

Elementary indices used in practice

20.38 Suppose that there are M lowest-level items or specific commodities in a chosen elementary category. Denote the period t price of item m by p_m^t for $t = 0,1$ and for items $m = 1,2,...,M$. Define the period t price vector as $p^t \equiv [p_1^t, p_2^t, ..., p_M^t]$ for $t = 0,1$.

[23] The results at the AC Nielsen brand level are a counter-example to this general assertion.

20.39 The first widely used elementary index number formula is attributable to the French economist Dutot (1738):

$$P_D(p^0, p^1) \equiv \frac{\sum_{m=1}^{M} \frac{1}{M} p_m^1}{\sum_{m=1}^{M} \frac{1}{M} p_m^0} = \frac{\sum_{m=1}^{M} p_m^1}{\sum_{m=1}^{M} p_m^0} \tag{20.1}$$

Thus the Dutot elementary price index is equal to the arithmetic average of the M period 1 prices divided by the arithmetic average of the M period 0 prices.

20.40 The second widely used elementary index number formula is attributable to the Italian economist Carli (1764):

$$P_C(p^0, p^1) \equiv \sum_{m=1}^{M} \frac{1}{M} \frac{p_m^1}{p_m^0}. \tag{20.2}$$

Thus the Carli elementary price index is equal to the arithmetic average of the M item price ratios or price relatives, $p_m{}^1/p_m{}^0$.

20.41 The third widely used elementary index number formula is attributable to the English economist Jevons (1863):

$$P_J(p^0, p^1) \equiv \prod_{m=1}^{M} \sqrt[M]{\frac{p_m^1}{p_m^0}} \tag{20.3}$$

Thus the Jevons elementary price index is equal to the geometric average of the M item price ratios or price relatives, $p_m{}^1/p_m{}^0$.

20.42 The fourth elementary index number formula P_H is the harmonic average of the M item price relatives. It was first suggested in passing as an index number formula by Jevons (1865, p. 121) and Coggeshall (1887):

$$P_H(p^0, p^1) \equiv \left[\sum_{m=1}^{M} \frac{1}{M} \left(\frac{p_m^1}{p_m^0} \right)^{-1} \right]^{-1} \tag{20.4}$$

20.43 Finally, the fifth elementary index number formula is the geometric average of the Carli and harmonic formulae; i.e., it is the geometric mean of the arithmetic and harmonic means of the *M* price relatives:

$$P_{CSWD}(p^0, p^1) \equiv \sqrt{P_C(p^0, p^1)P_H(p^0, p^1)} \qquad (20.5)$$

This index number formula was first suggested by Fisher (1922, p. 472) as his formula 101. Fisher also observed that, empirically for his data set, P_{CSWD} was very close to the Jevons index, P_J, and these two indices were his "best" unweighted index number formulae. In more recent times, Carruthers, Sellwood and Ward (1980, p. 25) and Dalén (1992, p. 140) also proposed P_{CSWD} as an elementary index number formula.

20.44 Having defined the most commonly used elementary formulae, the question now arises: which formula is "best"? Obviously, this question cannot be answered until desirable properties for elementary indices are developed. This will be done in a systematic manner in paragraphs 20.46 to 20.57, but one desirable property for an elementary index will be noted in the present section. This is the *time reversal test*, which was noted in Chapter 15. In the present context, this test for the elementary index $P(p^0,p^1)$ becomes:

$$P(p^0, p^1)\ P(p^1, p^0) = 1 \qquad (20.6)$$

This test says that if the prices in period 2 revert to the initial prices of period 0, then the product of the price change going from period 0 to 1, $P(p^0,p^1)$, times the price change going from period 1 to 2, $P(p^1,p^0)$, should equal unity; i.e., under the stated conditions, we should end up where we started. It can be verified that the Dutot, Jevons, and Carruthers-Sellwood-Ward-Dalén indices, P_D, P_J and P_{CSWD}, all satisfy the time reversal test, but that the Carli and harmonic indices, P_C and P_H, fail this test. In fact, these last two indices fail the test in the following *biased* manner:

$$P_C(p^0, p^1)\ P_C(p^1, p^0) \geq 1 \qquad (20.7)$$

$$P_H(p^0, p^1)\ P_H(p^1, p^0) \leq 1 \qquad (20.8)$$

with strict inequalities holding in (20.7) and (20.8) provided that the period 1 price vector p^1 is not proportional to the period 0 price vector p^0.[24] Thus the Carli index will generally have

[24] These inequalities follow from the fact that a harmonic mean of *M* positive numbers is always equal to or less than the corresponding arithmetic mean; see Walsh (1901, p. 517) or

an upward bias, while the harmonic index will generally have a downward bias. Fisher (1922, pp. 66 and 383) seems to have been the first to establish the upward bias of the Carli index,[25] and he made the following observations on its use by statistical agencies: "In fields other than index numbers it is often the best form of average to use. But we shall see that the simple arithmetic average produces one of the very worst of index numbers. And if this book has no other effect than to lead to the total abandonment of the simple arithmetic type of index number, it will have served a useful purpose" (Fisher (1922, pp. 29-30).

20.45 The following section establishes some numerical relationships between the five elementary indices defined in this section. Then in the subsequent section, a more comprehensive list of desirable properties for elementary indices is developed and the five elementary formulae are evaluated in the light of these properties or tests.

Numerical relationships between the frequently used elementary indices

20.46 It can be shown[26] that the Carli, Jevons and harmonic elementary price indices satisfy the following inequalities:

$$P_H(p^0, p^1) \leq P_J(p^0, p^1) \leq P_C(p^0, p^1) \qquad (20.9)$$

i.e., the harmonic index is always equal to or less than the Jevons index, which in turn is always equal to or less than the Carli index. In fact, the strict inequalities in (20.9) will hold provided that the period 0 vector of prices, p^0, is not proportional to the period 1 vector of prices, p^1.

20.47 The inequalities (20.9) do not tell us by how much the Carli index will exceed the Jevons index and by how much the Jevons index will exceed the harmonic index. Hence, in the remainder of this section, some approximate relationships between the five indices

---

Fisher (1922, pp. 383-384). This inequality is a special case of Schlömilch's inequality; see Hardy, Littlewood and Polya (1934, p. 26).

[25] See also Pigou (1920, pp. 59 and 70), Szulc (1987, p. 12) and Dalén (1992, p. 139). Dalén (1994, pp. 150-151) provides some nice intuitive explanations for the upward bias of the Carli index.

[26] Each of the three indices P_H, P_J and P_C is a mean of order r where r equals −1, 0 and 1, respectively, and so the inequalities follow from Schlömilch's inequality; see Hardy, Littlewood and Polya (1934, p. 26).

defined in the previous section are developed that provide some practical guidance on the relative magnitudes of each of the indices.

20.48 The first approximate relationship to be derived is between the Jevons index P_J and the Dutot index P_D.[27] For each period t, define the arithmetic mean of the M prices pertaining to that period as follows:

$$p^{t*} \equiv \sum_{m=1}^{M} \frac{1}{M} p_m^t \quad ; t = 0,1 \qquad (20.10)$$

Now define the multiplicative deviation of the mth price in period t relative to the mean price in that period, e_m^t, as follows:

$$p_m^t = p^{t*}\left(1 + e_m^t\right); \qquad m = 1,...,M \; ; \; t = 0,1. \qquad (20.11)$$

Note that equations (20.10) and (20.11) imply that the deviations e_m^t sum to zero in each period; i.e.

$$\sum_{m=1}^{M} e_m^t = 0; \qquad t = 0,1 \qquad (20.12)$$

20.49 Note that the Dutot index can be written as the ratio of the mean prices, p^{1*}/p^{0*}; i.e.

$$P_D(p^0, p^1) = \frac{p^{1*}}{p^{0*}} \qquad (20.13)$$

20.50 Now substitute equation (20.11) into the definition of the Jevons index (20.3):

---

[27] It should be noted that the Dutot index can also be written as a weighted average of the price relatives; i.e., $P_D(p^0,p^1) \equiv \sum_{i=1}^{n} p_i^1 \Big/ \sum_{j=1}^{n} p_j^0 = \sum_{i=1}^{n} (p_i^1/p_i^0) p_i^0 \Big/ \sum_{j=1}^{n} p_j^0 = \sum_{i=1}^{n} (p_i^1/p_i^0) w_i^0$ where the ith weight is defined as $w_i^0 \equiv p_i^0 \Big/ \sum_{j=1}^{n} p_j^0$. Thus if the commodities in the elementary aggregate are heterogeneous, the commodities that are more expensive in the chosen units of measurement will get a large weight, which may not be warranted from the viewpoint of expenditures on the commodity.

$$P_J(p^0, p^1) = \prod_{m=1}^{M} \sqrt[M]{\frac{p^{1*}(1+e_m^1)}{p^{0*}(1+e_m^0)}}.$$

$$= \frac{p^{1*}}{p^{0*}} \prod_{m=1}^{M} \sqrt[M]{\frac{(1+e_m^1)}{(1+e_m^0)}}$$

$$= P_D(p^0,p^1)\, f(e^0,e^1) \quad \text{using equation (20.13)} \qquad (20.14)$$

where $e^t \equiv [e_1^t,...,e_M^t]$ for $t = 0$ and 1, and where the function f is defined as follows:

$$f(e^0,e^1) \equiv \prod_{m=1}^{M} \sqrt[M]{\frac{(1+e_m^1)}{(1+e_m^0)}} \qquad (20.15)$$

20.51 Expand $f(e^0,e^1)$ by a second-order Taylor series approximation around $e^0 = 0_M$ and $e^1 = 0_M$. Using equation (20.12), it can be verified[28] that the following second-order approximate relationship between P_J and P_D is obtained:

$$P_J(p^0, p^1) \approx P_D(p^0, p^1)[1 + (1/2M)e^0e^0 - (1/2M)e^1e^1]$$

$$= P_D(p^0, p^1)[1 + (1/2)\,\mathrm{var}(e^0) - (1/2)\,\mathrm{var}(e^1)] \qquad (20.16)$$

where $\mathrm{var}(e^t)$ is the variance of the period t multiplicative deviations. Thus, for $t = 0,1$:

$$\mathrm{var}(e^t) \equiv \frac{1}{M}\sum_{m=1}^{M}(e_m^t - e^{t*})^2$$

$$= \frac{1}{M}\sum_{m=1}^{M}(e_m^t)^2$$

$$= \frac{1}{M} e^t e^t \qquad \text{since } e^{t*} = 0 \text{ using equation (20.12)} \qquad (20.17)$$

20.52 Under normal conditions,[29] the variance of the deviations of the prices from their

[28] This approximate relationship was first obtained by Carruthers, Sellwood and Ward (1980, p. 25).

[29] If there are significant changes in the overall inflation rate, some studies indicate that the variance of deviations of prices from their means can also change. Also if M is small, then there will be sampling fluctuations in the variances of the prices from period to period.

means in each period is likely to be approximately constant and so, under these conditions, the Jevons price index will approximate the Dutot price index to the second order.

20.53 Note that with the exception of the Dutot formula, the remaining four elementary indices defined in paragraphs 20.23 to 20.37 are functions of the relative prices of the M items being aggregated. This fact is used in order to derive some approximate relationships between these four elementary indices. Thus define the mth price relative as

$$r_m \equiv \frac{p_m^1}{p_m^0}; \quad m = 1,...,M \tag{20.18}$$

20.54 Define the arithmetic mean of the m price relatives as

$$r^* \equiv \frac{1}{M}\sum_{m=1}^{M} r_m = P_C(p^0,p^1) \tag{20.19}$$

where the last equality follows from the definition (20.2) of the Carli index. Finally, define the deviation e_m of the mth price relative r_m from the arithmetic average of the M price relatives r^* as follows:[30]

$$r_m = r^*(1+e_m); \quad m = 1,...,M \tag{20.20}$$

20.55 Note that equations (20.19) and (20.20) imply that the deviations e_m sum to zero:

$$\sum_{m=1}^{M} e_m = 0 \tag{20.21}$$

20.56 Now substitute equation (20.20) into the definitions (20.2)–(20.5) of P_C, P_J, P_H and P_{CSWD} in order to obtain the following representations for these indices in terms of the vector of deviations, $e \equiv [e_1,...,e_M]$:

$$P_C(p^0, p^1) = \sum_{m=1}^{M} \frac{1}{M} r_m = r^* 1 \equiv r^* f_C(e) \tag{20.22}$$

[30] Note that the ratio-type deviations e_m, defined by equation (20.20), are different from the level-type deviations e_m^t, defined by equation (20.11).

$$P_J(p^0,\ p^1) = \prod_{m=1}^{M} \sqrt[M]{r_m} = r^* \prod_{m=1}^{M} \sqrt[M]{1+e_m} \equiv r^* f_J(e) \tag{20.23}$$

$$P_H(p^0,\ p^1) = \left[\sum_{m=1}^{M} \frac{1}{M}(r_m)^{-1}\right]^{-1} = r^* \left[\sum_{m=1}^{M} \frac{1}{M}(1+e_m)^{-1}\right]^{-1} \equiv r^* f_H(e) \tag{20.24}$$

$$P_{CSWD}(p^0, p^1) = \sqrt{P_C(p^0, p^1) P_H(p^0, p^1)} = r^* \sqrt{f_C(e) f_H(e)} \equiv r^* f_{CSWD}(e) \tag{20.25}$$

where the last identity in each of equations (20.22)–(20.25) serves to define the deviation functions, $f_C(e)$, $f_J(e)$, $f_H(e)$ and $f_{CSWD}(e)$. The second-order Taylor series approximations to each of these functions[31] around the point $e = 0_M$ are:

$$f_C(e) = 1 \tag{20.26}$$

$$f_J(e) \approx 1 - (1/2M)ee = 1 - (1/2)\,\mathrm{var}(e) \tag{20.27}$$

$$f_H(e) \approx 1 - (1/M)ee = 1 - \mathrm{var}(e) \tag{20.28}$$

$$f_{CSWD}(e) \approx 1 - (1/2M)ee = 1 - (1/2)\,\mathrm{var}(e) \tag{20.29}$$

where repeated use of equation (20.21) is made in deriving the above approximations.[32] To the second order, the Carli index P_C will exceed the Jevons and Carruthers-Sellwood-Ward-Dalén indices, P_J and P_{CSWD}, by (1/2) r^* var(e), which is r^* times half the variance of the M price relatives p_m^1/p_m^0. Similarly, to the second order, the harmonic index P_H will lie below the Jevons and Carruthers-Sellwood-Ward-Dalén indices, P_J and P_{CSWD}, by r^* times half the variance of the M price relatives p_m^1/p_m^0.

20.57 Empirically, it is expected that the Jevons and Carruthers-Sellwood-Ward-Dalén indices will be very close to each other. Using the previous approximation result (20.16), it is expected that the Dutot index P_D will also be fairly close to P_J and P_{CSWD}, with some fluctuations over time as a result of changing variances of the period 0 and 1 deviation vectors, e^0 and e^1. Thus it is expected that these three elementary indices will give much the same numerical answers in empirical applications. In contrast, the Carli index can be

[31] From equation (20.22), it can be seen that $f_C(e)$ is identically equal to 1 so that the expression (20.26) will be an exact equality rather than an approximation.

[32] These second-order approximations are attributable to Dalén (1992, p. 143) for the case r^* =1 and to Diewert (1995a, p. 29) for the case of a general r^*.

expected to be substantially above these three indices, with the degree of divergence growing as the variance of the M price relatives grows. Similarly, the harmonic index can be expected to be substantially below the three middle indices, with the degree of divergence growing as the variance of the M price relatives grows.

The axiomatic approach to elementary indices

20.58 Recall the axiomatic approach to bilateral price indices $P(p^0,p^1,q^0,q^1)$ developed in Chapter 16. In the present chapter, the elementary price index $P(p^0,p^1)$ depends only on the period 0 and 1 price vectors, p^0 and p^1, respectively, so that the elementary price index does not depend on the period 0 and 1 quantity vectors, q^0 and q^1. One approach to obtaining new tests or axioms for an elementary index is to look at the 20 or so axioms listed in the Fisher axiomatic approach in Chapter 16 for bilateral price indices $P(p^0,p^1,q^0,q^1)$ and adapt those axioms to the present context; i.e., use the old bilateral tests for $P(p^0,p^1,q^0,q^1)$ that do not depend on the quantity vectors q^0 and q^1 as tests for an elementary index $P(p^0,p^1)$.[33] This is the approach taken in the present section.

20.59 The first eight tests or axioms are reasonably straightforward and uncontroversial.

T1: *Continuity*: $P(p^0,p^1)$ is a continuous function of the M positive period 0 prices $p^0 \equiv [p_1{}^0,...,p_M{}^0]$ and the M positive period 1 prices $p^1 \equiv [p_1{}^1,...,p_M{}^1]$.

T2: *Identity*: $P(p,p) = 1$; i.e., if the period 0 price vector equals the period 1 price vector, then the index is equal to unity.

T3: *Monotonicity in current period prices*: $P(p^0,p^1) < P(p^0,p)$ if $p^1 < p$; i.e., if any period 1 price increases, then the price index increases.

T4: *Monotonicity in base period prices*: $P(p^0,p^1) > P(p,p^1)$ if $p^0 <$ p; i.e., if any period 0 price increases, then the price index decreases.

T5: *Proportionality in current period prices*: $P(p^0,\lambda p^1) = \lambda P(p^0,p^1)$ if $\lambda > 0$; i.e., if all period 1 prices are multiplied by the positive number λ, then the initial price index is also multiplied by λ.

---

[33] The approach was used by Diewert (1995a, pp. 5-17), who drew on the earlier work of Eichhorn (1978,
pp. 152-160) and Dalén (1992).

T6: *Inverse proportionality in base period prices*: $P(\lambda p^0,p^1) = \lambda^{-1} P(p^0,p^1)$ if $\lambda > 0$; i.e., if all period 0 prices are multiplied by the positive number λ, then the initial price index is multiplied by $1/\lambda$.

T7: *Mean value test*: $\min_m \{p_m^1/p_m^0 : m = 1,...,M\} \leq P(p^0,p^1) \leq \max_m \{p_m^1/p_m^0 : m = 1,...,M\}$; i.e., the price index lies between the smallest and largest price relatives.

T8: *Symmetric treatment of outlets*: $P(p^0,p^1) = P(p^{0*},p^{1*})$, where p^{0*} and p^{1*} denote the *same* permutation of the components of p^0 and p^1; i.e., if we change the ordering of the outlets (or households) from which we obtain the price quotations for the two periods, then the elementary index remains unchanged.

Eichhorn (1978, p. 155) showed that tests T1, T2, T3 and T5 imply test T7, so that not all of the above tests are logically independent.

20.60 The following tests are more controversial and are not necessarily accepted by all price statisticians.

T9: *The price bouncing test*: $P(p^0,p^1) = P(p^{0*},p^{1**})$ where p^{0*} and p^{1**} denote possibly *different* permutations of the components of p^0 and p^1; i.e., if the ordering of the price quotes for both periods is changed in possibly different ways, then the elementary index remains unchanged.

20.61 Obviously, test T8 is a special case of test T9 where the two permutations of the initial ordering of the prices are restricted to be the same. Thus test T9 implies test T8. Test T9 is attributable to Dalén (1992, p. 138). He justified this test by suggesting that the price index should remain unchanged if outlet prices "bounce" in such a manner that the outlets are just exchanging prices with each other over the two periods. While this test has some intuitive appeal, it is not consistent with the idea that outlet prices should be matched to each other in a one-to-one manner across the two periods. This outlet price matching is preferable to not matching prices across outlets in case there are quality differences across outlets.

20.62 The following test was also proposed by Dalén (1992) in the elementary index context:

T10: *Time reversal*: $P(p^1,p^0) = 1/P(p^0,p^1)$; i.e., if the data for periods 0 and 1 are interchanged, then the resulting price index should equal the reciprocal of the original price index.

Since many price statisticians approve of the Laspeyres price index in the bilateral index context and this index does not satisfy the time reversal test, it is obvious that not all price statisticians would regard the time reversal test in the elementary index context as being a fundamental test that must be satisfied. Nevertheless, many other price statisticians do regard this test as a fundamental one since it is difficult to accept an index that gives a different answer if the ordering of time is reversed.

20.63 The following test is a strengthening of the time reversal test:

T11: *Circularity*: $P(p^0,p^1)P(p^1,p^2) = P(p^0,p^2)$; i.e., the price index going from period 0 to 1 times the price index going from period 1 to 2 equals the price index going from period 0 to 2 directly.

The circularity and identity tests imply the time reversal test (just set $p^2 = p^0$). Thus the circularity test is essentially a strengthening of the time reversal test, and so price statisticians who do not accept the time reversal test are unlikely to accept the circularity test. In general, however, the circularity test seems to be a very desirable property: it is a generalization of a property that holds for a single price relative.

20.64 The following test is a very important one:

T12: *Commensurability*: $P(\lambda_1 p_1^0,..., \lambda_M p_M^0; \lambda_1 p_1^1,..., \lambda_M p_M^1) = P(p_1^0,...,p_M^0; p_1^1,...,p_M^1) = P(p^0,p^1)$ for all $\lambda_1 > 0, ... , \lambda_M > 0$; i.e., if the units of measurement for each commodity are changed, then the elementary index remains unchanged.

In the bilateral index context, virtually every price statistician accepts the validity of this test. In the elementary context, however, this test is more controversial. If the M items in the elementary aggregate are all homogeneous, then it makes sense to measure all the items in the same units. Hence, if the unit of measurement of the homogeneous commodity is changed, then a modified version of test T12 should restrict all the λ_m to be the same number (say λ) and the modified test T12 becomes

$$P(\lambda p^0,\lambda p^1) = P(p^0,p^1) \; ; \; \lambda > 0. \qquad (20.30)$$

Note that this modified test T12 will be satisfied if tests T5 and T6 are satisfied. Thus if the items in the elementary aggregate are homogeneous, then there is no need for the original (unmodified) test T12.

20.65 In actual practice, there will usually be thousands of individual items in each

elementary aggregate and the hypothesis of item homogeneity is not warranted. Under these circumstances, it is important that the elementary index satisfy the commensurability test, since the units of measurement of the heterogeneous items in the elementary aggregate are arbitrary, and hence the price statistician can change the index simply by changing the units of measurement for some of the items.

20.66 This completes the listing of the tests for an elementary index. There remains the task of evaluating how many tests are passed by each of the five elementary indices defined in paragraphs 20.38 to 20.45.

20.67 Straightforward computations show that the Jevons elementary index P_J satisfies all the tests, and hence emerges as being "best" from the viewpoint of this particular axiomatic approach to elementary indices.

20.68 The Dutot index P_D satisfies all the tests with the important exception of the commensurability test T12, which it fails. If there are heterogeneous items in the elementary aggregate, this is a rather serious failure and hence price statisticians should be careful in using this index under these conditions.

20.69 The geometric mean of the Carli and harmonic elementary indices, P_{CSWD}, fails only the price bouncing test T9 and the circularity test T11. The failure of these two tests is probably not a disqualifying condition, and so this index could be used by price statisticians if, for some reason, it was decided not to use the Jevons formula. As was observed in paragraphs 20.38 to 20.45, numerically, P_{CSWD} will be very close to P_J.

20.70 The Carli and harmonic elementary indices, P_C and P_H, fail the price bouncing test T9, the time reversal test T10 and the circularity test T11, and pass the other tests. The failure of tests T9 and T11 is again not a disqualifying condition, but the failure of the time reversal test T10 is a rather serious matter and so price statisticians should be cautious in using these indices.

The economic approach to elementary indices

20.71 Recall the notation and discussion in paragraphs 20.38 to 20.45. Suppose that each

purchaser of the items in the elementary aggregate has preferences over a vector of purchases $q \equiv [q_1,...,q_M]$ that can be represented by the linearly homogeneous aggregator (or utility) function $f(q)$. Further assume that each purchaser engages in cost-minimizing behaviour in each period. Then, as seen in Chapter 17, it can be shown that certain specific functional forms for the aggregator or utility function $f(q)$ or its dual unit cost function $c(p)$[34] lead to specific functional forms for the price index $P(p^0,p^1,q^0,q^1)$ with

$$P(p^0,p^1,q^0,q^1) \equiv \frac{c(p^1)}{c(p^0)} \tag{20.31}$$

20.72 Suppose that the purchasers have aggregator functions f defined as follows:[35]

$$f(q_1,...,q_M) \equiv \min_m \{q_m/\alpha_m : m = 1,...,M\} \tag{20.32}$$

where the α_m are positive constants. Then under these assumptions, it can be shown that equation (20.31) becomes:[36]

$$\frac{c(p^1)}{c(p^0)} = \frac{p^1q^0}{p^0q^0} = \frac{p^1q^1}{p^0q^1} \tag{20.33}$$

and the quantity vectors of purchases during the two periods must be proportional; i.e.,

$$q^1 = \lambda q^0 \text{ for some } \lambda > 0 \tag{20.34}$$

20.73 From the first equation in (20.33), it can be seen that the true cost of living index, $c(p^1)/c(p^0)$, under assumptions (20.32) about the aggregator function f, is equal to the Laspeyres price index, $P_L(p^0,p^1,q^0,q^1) \equiv p^1q^0/p^0q^0$. It is shown below how various elementary formulae can estimate this Laspeyres formula under alternative assumptions about the sampling of prices.

---

[34] The unit cost function is defined as $c(p) \equiv \min_q \left\{ \sum_{m=1}^{M} p_m q_m : f(q) = 1 \right\}$.

[35] The preferences which correspond to this f are known as Leontief (1936) or no substitution preferences.

[36] See Pollak (1983). Notation: p^1q^0 is defined as $\sum_{i=1}^{n} p_i^1 q_i^0$, etc.

20.74 In order to provide a justification for the use of the Dutot elementary formula, write the Laspeyres index number formula as follows:

$$P_L(p^0,p^1,q^0,q^1) \equiv \frac{\sum_{m=1}^{M} p_m^1 q_m^0}{\sum_{m=1}^{M} p_m^0 q_m^0} = \frac{\sum_{m=1}^{M} \rho_m^0 p_m^1}{\sum_{m=1}^{M} \rho_m^0 p_m^0} \qquad (20.35)$$

where the *base period item probabilities* ρ_m^0 are defined as follows:

$$\rho_m^0 \equiv \frac{q_m^0}{\sum_{m=1}^{M} q_m^0}; \quad m = 1,\ldots M \qquad (20.36)$$

Thus the base period probability for item *m*, ρ_m^0, is equal to the purchases of item *m* in the base period relative to total purchases of all items in the commodity class in the base period. Note that these definitions require that all items in the commodity class have the same units.[37]

20.75 Now it is easy to see how formula (20.35) could be turned into a rigorous sampling framework for sampling prices in the particular commodity class under consideration.[38] If item prices in the commodity class were sampled proportionally to their base period probabilities ρ_m^0, then the Laspeyres index defined by the first equality in (20.35) could be estimated by a probability-weighted Dutot index defined by the second equality in (20.35). In general, with an appropriate sampling scheme, the use of the Dutot formula at the elementary level of aggregation *for homogeneous items* can be perfectly consistent with a Laspeyres index concept.

20.76 The Dutot formula can also be consistent with a Paasche index concept. If the Paasche formula is used at the elementary level of aggregation, then the following formula is obtained:

---

[37] The probabilities defined by equation (20.36) are meaningless unless the items are homogeneous.

[38] For the details, see Balk (2002, pp. 8-10).

$$P_P(p^0,p^1,q^0,q^1) \equiv \frac{\sum_{m=1}^{M} p_m^1 q_m^1}{\sum_{m=1}^{M} p_m^0 q_m^1} = \frac{\sum_{m=1}^{M} \rho_m^1 p_m^1}{\sum_{m=1}^{M} \rho_m^1 p_m^0} \qquad (20.37)$$

where the period 1 item probabilities ρ_m^1 are defined as follows:

$$\rho_m^1 \equiv \frac{q_m^1}{\sum_{m=1}^{M} q_m^1}; \qquad m = 1,...,M \qquad (20.38)$$

Thus the period 1 probability for item m, ρ_m^1, is equal to the quantity purchased of item m in period 1 relative to total purchases of all items in the commodity class in that period.

20.77 Again, it is easy to see how formula (20.37) could be turned into a rigorous sampling framework for sampling prices in the particular commodity class under consideration. If item prices in the commodity class were sampled proportionally to their period 1 probabilities ρ_m^1, then the Paasche index defined by the first equality in (20.37) could be estimated by the probability-weighted Dutot index defined by the second equality in (20.37). In general, with an appropriate sampling scheme, the use of the Dutot formula at the elementary level of aggregation (for a homogeneous elementary aggregate) can be perfectly consistent with a Paasche index concept.

20.78 Rather than use the fixed basket representations for the Laspeyres and Paasche indices, it is possible to use the expenditure share representations for the Laspeyres and Paasche indices, and to use the expenditure shares s_m^0 or s_m^1 as probability weights for price relatives. Thus if the relative prices of items in the commodity class under consideration are sampled using weights that are proportional to their base period expenditure shares in the commodity class, then the following probability-weighted Carli index

$$P_C\left(p^0,p^1,s^0\right) \equiv \sum_{m=1}^{M} s_m^0 \frac{p_m^1}{p_m^0} \qquad (20.39)$$

will be equal to the Laspeyres index.[39] Of course, formula (20.39) does not require the

[39] For a rigorous derivation of a sampling framework, see Balk (2002, p. 13-14).

assumption of homogeneous items as did equations (20.35) and (20.37).

20.79 If the relative prices of items in the commodity class under consideration are sampled using weights that are proportional to their period 1 expenditure shares in the commodity class, then the following probability-weighted harmonic index

$$P_H\left(p^0, p^1, s^1\right) \equiv \left(\sum_{m=1}^{M} s_m^1 \left(\frac{p_m^1}{p_m^0} \right)^{-1} \right)^{-1} \tag{20.40}$$

will be equal to the Paasche index.

20.80 The above results show that the Dutot elementary index can be justified as an approximation to an underlying Laspeyres or Paasche price index for a homogeneous elementary aggregate under appropriate price sampling schemes. The above results also show that the Carli and harmonic elementary indices can be justified as approximations to an underlying Laspeyres or Paasche price index for a heterogeneous elementary aggregate under appropriate price sampling schemes.

20.81 Recall that assumption (20.32) on *f* justified the Laspeyres and Paasche indices as being the "true" elementary aggregate from the viewpoint of the economic approach to elementary indices. Suppose now that assumption (20.32) is replaced by the following assumption of Cobb-Douglas (1928) preferences:[40]

$$f(q_1,...,q_M) \equiv \prod_{m=1}^{M} q^{\beta_m}; \quad \beta_m > 0 \quad \text{for } m = 1,...,M \quad \text{and} \quad \sum_{m=1}^{M} \beta_m = 1 \tag{20.41}$$

20.82 Under assumption (20.41), the true economic elementary price index is:[41]

$$\frac{c(p^1)}{c(p^0)} = \prod_{m=1}^{M} \left(\frac{p_m^1}{p_m^0} \right)^{\beta_m} \tag{20.42}$$

[40] These preferences were introduced slightly earlier by Konüs and Byushgens (1926).

[41] See Pollak (1983).

20.83 It turns out that if purchasers have the above Cobb-Douglas preferences, then item expenditures will be proportional over the two periods so that:

$$p_m^1 q_m^1 = \lambda p_m^0 q_m^0 \text{ for } m = 1,...,M \text{ and for some } \lambda > 0. \quad (20.43)$$

Under these conditions, the base period expenditure shares s_m^0 will equal the corresponding period 1 expenditure shares s_m^1, as well as the corresponding β_m; i.e., assumption (20.41) implies:

$$s_m^0 = s_m^1 \equiv \beta_m ; \quad m = 1,...,M. \quad (20.44)$$

Thus if the relative prices of items in the commodity class under consideration are sampled using weights that are proportional to their base period expenditure shares in the commodity class, then the following probability-weighted Jevons index

$$\ln P_J \left(p^0, p^1, s^0 \right) \equiv \sum_{m=1}^{M} s_m^0 \ln \frac{p_m^1}{p_m^0} \quad (20.45)$$

will be equal to the logarithm of the true elementary price aggregate defined by equation (20.42).[42]

20.84 The above results show that the Jevons elementary index can be justified as an approximation to an underlying Cobb-Douglas price index for a heterogeneous elementary aggregate under an appropriate price sampling scheme.

20.85 The assumption of Leontief preferences implies that the quantity vectors pertaining to the two periods under consideration will be proportional; recall equation (20.34). In contrast, the assumption of Cobb-Douglas preferences implies that expenditures will be proportional over the two periods; recall equation (20.43). Index number theorists have been debating the relative merits of the proportional quantities versus proportional expenditures assumption for a long time. Authors who thought that the proportional expenditures assumption was empirically more likely include Jevons (1865, p. 295) and Ferger (1931, p. 39; 1936, p. 271). These early authors did not have the economic approach to index number theory at their disposal but they intuitively understood, along with Pierson (1895, p. 332), that substitution effects occurred and hence the proportional expenditures assumption was more plausible than the proportional quantities assumption.

[42] See Balk (2002, pp. 11-12) for a rigorous derivation.

20.86 The results in the previous section gave some support for the use of the unweighted Jevons elementary index over the use of the unweighted Dutot, Carli and harmonic indices, provided that the proportional expenditures assumption is more likely than the proportional quantities assumption. This support is very weak, however, since an appropriate item price sampling scheme is required in order to justify the results. Thus, using an unweighted Dutot, Carli or harmonic index (without the appropriate sampling scheme) cannot really be justified from the viewpoint of the economic approach. The results in this section nevertheless give considerable support to the use of an appropriately weighted Jevons index over the other weighted indices, since from the economic perspective, cross-item elasticities of substitution are much more likely to be close to unity (this corresponds to the case of Cobb-Douglas preferences) than to zero (this corresponds to the case of Leontief preferences). If the probability weights in the weighted Jevons index are taken to be the arithmetic average of the period 0 and 1 item expenditure shares and narrowly defined unit values are used as the price concept, then the weighted Jevons index becomes an ideal type of elementary index discussed in paragraphs 20.11 to 20.22.

The sampling approach to elementary indices

20.87 In the previous section, it is shown that appropriately weighted elementary indices are capable of approximating various economic population elementary indices, with the approximation becoming exact as the sampling approaches complete coverage. Conversely, it can be seen that, in general, it is impossible for an unweighted elementary price index of the type defined in paragraphs 20.38 to 20.45 to approach the theoretically ideal elementary price index defined in paragraphs 20.11 to 20.22, even if all item prices in the elementary aggregate are sampled.[43] Hence, rather than just sampling prices, it will be necessary for the price statistician to collect information on the transaction values (or quantities) associated with the sampled prices in order to form sample elementary aggregates that will approach the target ideal elementary aggregate as the sample size becomes large. Thus, instead of just collecting a sample of prices, it will be necessary to collect corresponding sample quantities (or values) so that a sample Fisher, Törnqvist or Walsh price index can be constructed. This sample-based superlative elementary price index will approach the population ideal

[43] The numerical example given in paragraphs 20.91 to 20.99 illustrates this point.

elementary index as the sample size becomes large. This approach to the construction of elementary indices in a sampling context was recommended by Pigou (1920, pp. 66-67), Fisher (1922, p. 380), Diewert (1995a, p. 25) and Balk (2002).[44] In particular, Pigou (1920, p. 67) suggested that the sample-based Fisher ideal price index be used to deflate the value ratio for the aggregate under consideration in order to obtain an estimate of the quantity ratio for the aggregate under consideration.

The use of scanner data in constructing elementary aggregates

20.88 Until fairly recently, it was not possible to determine how close an unweighted elementary index of the type defined in paragraphs 20.38 to 20.45 was to an ideal elementary aggregate. With the availability of scanner data (i.e., of detailed data on the prices and quantities of individual items that are sold in retail outlets), it has now become possible to compute ideal elementary aggregates for some item strata and compare the results with statistical agency estimates of price change for the same class of items. Of course, the statistical agency estimates of price change are usually based on the use of the Dutot, Jevons or Carli formulae. The following quotations reflect the results of many of the scanner data studies:

> A second major recent development is the willingness of statistical agencies to experiment with scanner data, which are the electronic data generated at the point of sale by the retail outlet and generally include transactions prices, quantities, location, date and time of purchase and the product described by brand, make or model. Such detailed data may prove especially useful for constructing better indices at the elementary level. Recent studies that use scanner data in this way include Silver (1995), Reinsdorf (1996), Bradley, Cook, Leaver and Moulton (1997), Dalén (1997), de Haan and Opperdoes (1997) and Hawkes (1997). Some estimates of elementary index bias (on an annual basis) that emerged from these studies were: 1.1 percentage points for television sets in the United Kingdom; 4.5 percentage points for coffee in the United States; 1.5 percentage points for ketchup, toilet tissue, milk and tuna in the United States; 1 percentage point for fats, detergents, breakfast cereals and frozen fish in Sweden; 1 percentage point for coffee in the Netherlands and 3 percentage points for coffee in the United States respectively. These bias estimates incorporate both elementary and outlet substitution biases and are significantly higher than our earlier ballpark estimates of .255 and .41 percentage points. On the other hand, it is unclear to what extent these large bias estimates can be generalized to other commodities (Diewert (1998a, pp. 54-55)).

[44] Balk (2002) provides the details for this sampling framework. Hausman (2002) is another recent author who stressed the importance of collecting quantity information along with price information at the elementary level.

> Before considering the results it is worth commenting on some general findings from scanner data. It is stressed that the results here are for an experiment in which the same data were used to compare different methods. The results for the U.K. Retail Prices Index can not be fairly compared since they are based on quite different practices and data, their data being collected by price collectors and having strengths as well as weaknesses (Fenwick, Ball, Silver and Morgan (2003)). Yet it is worth following up on Diewert's (2002c) comment on the U.K. Retail Prices Index electrical appliances section, which includes a wide variety of appliances, such as irons, toasters, refrigerators, etc. which went from 98.6 to 98.0, a drop of 0.6 percentage points from January 1998 to December 1998. He compares these results with those for washing machines and notes that "…it may be that the non washing machine components of the electrical appliances index increased in price enough over this period to cancel out the large apparent drop in the price of washing machines but I think that this is somewhat unlikely." A number of studies on similar such products have been conducted using scanner data for this period. Chained Fisher indices have been calculated from the scanner data, (the RPI (within year) indices are fixed base Laspeyres ones), and have been found to fall by about 12 per cent for televisions (Silver and Heravi, 2001a), 10 per cent for washing machines (Table 7 below), 7.5 per cent for dishwashers, 15 per cent for cameras and 5 per cent for vacuum cleaners (Silver and Heravi, 2001b). These results are quite different from those for the RPI section and suggest that the washing machine disparity, as Diewert notes, may not be an anomaly. Traditional methods and data sources seem to be giving much higher rates for the CPI than those from scanner data, though the reasons for these discrepancies were not the subject of this study (Silver and Heravi (2002, p. 25)).

20.89 The above quotations summarize the results of many elementary aggregate index number studies that are based on the use of scanner data. These studies indicate that when detailed price and quantity data are used in order to compute superlative indices or hedonic indices for an expenditure category, the resulting measures of price change are often below the corresponding official statistical agency estimates of price change for that category.[45] Sometimes the measures of price change based on the use of scanner data are considerably below the corresponding official measures.[46] These results indicate that there may be large gains in the precision of elementary indices if a weighted sampling framework is adopted.

---

[45] Recall also the results obtained by Koskimäki and Ylä-Jarkko (2003) that showed the Laspeyres index considerably above the corresponding Fisher index using Finnish scanner data.

[46] Scanner data studies do not, however, always show large potential biases in official CPIs. Masato Okamoto has informed us that a large-scale comparative study in Japan is under way. Using scanner data for about 250 categories of processed food and daily necessities collected over the period 1997 to 2000, it was found that the indices based on scanner data averaged only about 0.2 percentage points below the corresponding official indices per year. Japan uses the Dutot formula at the elementary level in its official CPI.

20.90 Is there a simple intuitive explanation for the above empirical results? A partial explanation may be possible by looking at the dynamics of item demand. In any market economy, there are firms and outlets that sell items that are either declining or increasing in price. Usually, the items that decline in price experience an increase in their volume of sales. Thus the expenditure shares that are associated with items that are declining in price usually increase, and conversely for the items that increase in price. Unfortunately, elementary indices cannot pick up the effects of this negative correlation between price changes and the induced changes in expenditure shares, because elementary indices depend only on prices and not on expenditure shares.

20.91 An example can illustrate the above point. Suppose that there are only three items in the elementary aggregate and that in period 0, the price of each item is $p_m^0 = 1$ and the expenditure share for each item is equal so that $s_m^0 = 1/3$ for $m = 1,2,3$. Suppose that in period 1, the price of item 1 increases to $p_1^1 = 1 + i$, the price of item 2 remains constant at $p_2^1 = 1$ and the price of item 3 decreases to $p_3^1 = (1 + i)^{-1}$ where the item 1 rate of increase in price is $i > 0$. Suppose further that the expenditure share of item 1 decreases to $s_1^1 = (1/3) - \sigma$ where σ is a small number between 0 and 1/3 and the expenditure share of item 3 increases to $s_3^1 = (1/3) + \sigma$.[47]

The expenditure share of item 2 remains constant at $s_2^1 = 1/3$. The five elementary indices defined in paragraphs 20.23 to 20.37 can all be written as functions of the item 1 inflation rate i (which is also the item 3 deflation rate) as follows:

$$P_J(p^0, p^1) = \left[(1-i)(1+i)^{-1}\right]^{1/3} = 1 \equiv f_J(i) \qquad (20.46)$$

The term (1-i) should be replaced by (1+i) in equation (20.46).

$$P_C(p^0, p^1) = (1/3)(1+i) + (1/3) + (1/3)(1+i)^{-1} \equiv f_C(i) \qquad (20.47)$$

[47] The parameter σ is a measure of the degree of substitutability between the various items in the elementary aggregate. It is not precisely equal to the elasticity of substitution parameter σ which appeared in the Lloyd Moulton formula explained in paragraphs 17.61 to 17.64 of Chapter 17. However, the bigger is the elasticity of substitution, the bigger will be the σ parameter which appears in this section. David E. Lebow and Jeremy B. Rudd note that the marketing literature finds that the elasticity of substitution between brands in a narrowly defined elementary aggregate is around 2.5 (which is much higher than the Cobb-Douglas case where the elasticity of substitution is 1): "And, Gerard Tellis (1988) analyzed the results from a large number of papers in the marketing literature that estimate cross brand elasticities and found a mean elasticity (after adjusting for certain biases in the results) of 2.5" (Lebow and Rudd (2003, p. 167-168).

$$P_H(p^0,p^1) = \left[(1/3)(1+i)^{-1} + (1/3) + (1/3)(1+i)\right]^{-1} \equiv f_H(i) \tag{20.48}$$

$$P_{CSWD}(p^0,p^1) = \left[P_C(p^0,p^1)P_H(p^0,p^1)\right]^{1/2} \equiv f_{CSWD}(i); \tag{20.49}$$

$$P_D(p^0,p^1) = (1/3)(1+i) + (1/3) + (1/3)(1+i)^{-1} \equiv f_D(i) \tag{20.50}$$

20.92 Note that in this particular example, the Dutot index $f_D(i)$ turns out to equal the Carli index $f_C(i)$. The second-order Taylor series approximations to the five elementary indices (20.46)–(20.50) are given by the approximations (20.51)–(20.55):

$$f_J(i) \approx 1 \tag{20.51}$$

$$f_C(i) \approx 1 + (1/3)i^2 \tag{20.52}$$

$$f_H(i) \approx 1 - (1/3)i^2 \tag{20.53}$$

$$f_{CSWD}(i) \approx 1 \tag{20.54}$$

$$f_D(i) \approx 1 + (1/3)i^2 \tag{20.55}$$

Thus for small i, the Carli and Dutot indices will be slightly greater than 1,[48] the Jevons and the Carruthers-Sellwood-Ward-Dalén indices will be approximately equal to 1 and the harmonic index will be slightly less than 1. Note that the first-order Taylor series approximation to all five indices is 1. Thus, to the accuracy of a first-order approximation, all five indices equal unity.

20.93 Now calculate the Laspeyres, Paasche and Fisher indices for the elementary aggregate:

$$P_L = (1/3)(1+i) + (1/3) + (1/3)(1+i)^{-1} \equiv f_L(i) \tag{20.56}$$

$$P_P = \left\{\left[(1/3) - \sigma\right](1+i) + (1/3) + \left[(1/3) + \sigma\right](1+i)^{-1}\right\}^{-1} \equiv f_P(i) \tag{20.57}$$

$$P_F = (P_L P_P)^{1/2} \equiv f_F(i) \tag{20.58}$$

[48] Recall the approximate relationship (20.16) in paragraph 20.51 between the Dutot and Jevons indices. In the present numerical example, var(e^0) = 0 whereas var(e^1) > 0. This explains why the Dutot index is not approximately equal to the Jevons index in this numerical example.

20.94 First-order Taylor series approximations to the above indices (20.56)–(20.58) around $i = 0$ are given by the approximations (20.59)–(20.61):

$$f_L(i) \approx 1 \qquad (20.59)$$

$$f_P(i) \approx 1 - 2\sigma i \qquad (20.60)$$

$$f_F(i) \approx 1 - \sigma i \qquad (20.61)$$

20.95 An ideal elementary index for the three items is the Fisher ideal index $f_F(i)$. The approximations (20.51)–(20.55) and (20.61) show that the Fisher index will lie below all five elementary indices by the amount σi, taking first-order approximations to all six indices. Thus all five elementary indices will have an approximate upward bias equal to σi compared to an ideal elementary aggregate.

20.96 Suppose that the annual item inflation rate for the item rising in price is equal to 10 per cent so that $i = 0.10$ (and hence the rate of price decrease for the item decreasing in price is approximately 10 per cent as well). If the expenditure share of the increasing price item declines by 5 percentage points, then $\sigma = 0.05$ and the annual approximate upward bias in all five elementary indices is $\sigma i = 0.05 \times 0.10 = 0.005$ or half of a percentage point. If i increases to 20 per cent and σ increases to 10 per cent, then the approximate bias increases to $\sigma i = 0.10 \times 0.20 = 0.02$ or 2 per cent. Note, however, if prices in period 2 revert to the prices prevailing in period 0, then the bias will reverse itself. Hence elementary bias of the type modelled above can only cumulate over successive periods if there are long-run trends in prices and market shares.[49]

20.97 The above example is highly simplified. More sophisticated models are capable of explaining at least some of the discrepancy between official elementary indices and superlative indices calculated by using scanner data for an expenditure class. Basically, elementary indices defined without using associated quantity or value weights are incapable

[49] White's (2000) research into Canadian outlet substitution bias indicated that not only did discount outlets have lower prices for the same items, but they also had lower inflation rates over time.

of picking up shifts in expenditure shares that are induced by fluctuations in item prices.[50] In order to eliminate the problem of an inability to pick up shifts in expenditure shares that are induced by fluctuations in item prices, it will be necessary to sample values along with prices in both the base and comparison periods.

20.98 A few words of caution are, however, in order at this point. The use of chained superlative indices can lead to very biased results if there are large period-to-period fluctuations in prices and quantities compared to longer-run trends in prices. In long runs, large fluctuations can be induced by seasonal factors[51] or by temporary sales.[52]

20.99 In the following section, a simple regression-based approach to the construction of elementary indices is outlined. The importance of weighting the price quotes will again emerge from the analysis.

A simple stochastic approach to elementary indices

20.100 Recall the notation used in paragraphs 20.38 to 20.45 above. Suppose the prices of the *M* items for period 0 and 1 are approximately equal to the right-hand sides of equations

---

[50] Put another way, elementary indices are subject to substitution or representativity bias. In the case of Cobb-Douglas preferences, however, the parameter σ in this section would be equal to zero and the Jevons elementary aggregate would be unbiased. But the results from the marketing literature (recall Tellis (1988)) indicate that σ will be greater than 0 and hence that the Jevons elementary index will have an upward bias. Thus Lebow and Rudd's (2003, p. 167) estimate that elementary substitution bias is only around 0.05 percentage points per year if the Jevons formula is used seems rather low.

[51] For an example where the use of chained superlative indices leads to a tremendous downward bias induced by seasonal fluctuations, see Chapter 22.

[52] For an example where the use of chained superlative indices leads to a tremendous upward bias induced by periodic sales, see Robert C. Feenstra and Matthew D. Shapiro (2003):

> The reason for this is that periods of *low* prices (i.e., sales) attract high purchases only when they are accompanied by advertising, and this tends to occur in the final weeks of a sale. Thus, the initial price decline, when the sale starts, does not receive as much weight in the cumulative index as the final price *increase* when the sale ends. The demand behavior that leads to this upward bias of the chained Törnqvist – with higher purchases at the end of a sale – means that consumers are very likely purchasing goods for inventory accumulation. The only theoretically correct index to use in this type of situation is a fixed base index, as demonstrated in section 5.3 (Feenstra and Shapiro (2003, p. 125)).

The use of a fixed base index in these circumstances may, however, lead to results that are highly dependent on the choice of the base period. Other solutions that could be tried in this type of circumstance are either lengthening the period of time (as discussed in paragraphs 20.23 to 20.37) or using the moving year idea explained in Chapter 22 below.

(20.62) and (20.63):

$$p_m^0 \approx \beta_m; \quad m = 1,\ldots M \tag{20.62}$$

$$p_m^1 \approx \alpha\beta_m; \quad m = 1,\ldots M \tag{20.63}$$

where α and the β_m are positive parameters. Note that there are $2M$ prices on the left-hand sides of equations (20.62) and (20.63), but only $M + 1$ parameters on the right-hand sides of these equations. The basic hypothesis in the model of price behaviour defined by equations (20.62) and (20.63) is that the two price vectors p^0 and p^1 are proportional (with $p^1 = \alpha p^0$ so that α is the factor of proportionality) except for random multiplicative errors. Hence α represents the value of the underlying elementary price aggregate. Taking logarithms of both sides of equations (20.62) and (20.63), and adding some random errors e_m^0 and e_m^1 to the right-hand sides of the resulting equations, the following linear regression model is obtained:

$$\ln p_m^0 = \delta_m + e_m^0; \quad m = 1,...,M; \tag{20.64}$$

$$\ln p_m^1 = \gamma + \delta_m + e_m^1; \quad m = 1,...,M \tag{20.65}$$

where

$$\gamma \equiv \ln\alpha \text{ and } \delta_m \equiv \ln\beta_m; \quad m = 1,...,M. \tag{20.66}$$

20.101 Note that equations (20.64) and (20.65) can be interpreted as a highly simplified hedonic regression model.[53] The only characteristic of each commodity is the commodity itself. This model is also a special case of the country product dummy method for making international comparisons between the prices of different countries.[54] A major advantage of this regression method for constructing an elementary price index is that standard errors for the index number α can be obtained. This advantage of the stochastic approach to index number theory was stressed by Selvanathan and Rao (1994).

20.102 It can be verified that the least squares estimator for γ is:

$$\gamma^* \equiv \sum_{m=1}^{M} \frac{1}{M} \ln \frac{p_m^1}{p_m^0} \tag{20.67}$$

[53] See Chapters 7, 8 and 21 for discussion of hedonic regression models.

[54] See Summers (1973). In our special case, there are only two "countries" which are the two observations on the prices of the elementary aggregate for two periods.

20.103 If γ^* is exponentiated, then the following estimator for the elementary aggregate α is obtained:

$$\alpha^* \equiv \prod_{m=1}^{M} \sqrt[M]{\frac{p_m^1}{p_m^0}} \equiv P_J(p^0, p^1) \tag{20.68}$$

where $P_J(p^0,p^1)$ is the Jevons elementary price index defined in paragraphs 20.38 to 20.45 above. Thus the simple regression model defined by equations (20.64) and (20.65) leads to a justification for the use of the Jevons elementary index.

20.104 Consider the following unweighted least squares model:

$$\min_{\gamma,\delta's} \sum_{m=1}^{M} \left(\ln p_m^0 - \delta_m\right)^2 + \sum_{m=1}^{M} \left(\ln p_m^1 - \gamma - \delta_m\right)^2 \tag{20.69}$$

It can be verified that the γ solution to the unconstrained minimization problem (20.69) is the γ^* defined by equation (20.67).

20.105 There is a problem with the unweighted least squares model defined by equation (20.69), namely, that the logarithm of each price quote is given exactly the same weight in the model no matter what the expenditure on that item was in each period. This is obviously unsatisfactory since a price that has very little economic importance (i.e., a low expenditure share in each period) is given the same weight in the regression model as a very important item. Thus it is useful to consider the following weighted least squares model:[55]

$$\min_{\gamma,\delta's} \sum_{m=1}^{M} s_m^0 \left(\ln p_m^0 - \delta_m\right)^2 + \sum_{m=1}^{M} s_m^1 \left(\ln p_m^1 - \gamma - \delta_m\right)^2 \tag{20.70}$$

where the period *t* expenditure share on commodity *m* is defined in the usual manner as:

$$s_m^t \equiv \frac{p_m^t q_m^t}{\sum_{m=1}^{M} p_m^t q_m^t}; \quad t = 0,1; \quad m = 1,...,M \tag{20.71}$$

In the model (20.70), the logarithm of each item price quotation in each period is weighted by its expenditure share in that period. Note that weighting prices by their economic importance

[55] Balk (1980c) considers a similar weighted least squares model for many periods but with different weights.

is consistent with Theil's (1967, pp. 136-138) stochastic approach to index number theory.[56]

20.106 The γsolution to the minimization problem (20.70) is

$$\gamma^{**} \equiv \frac{\sum_{m=1}^{M} h\left(s_m^0, s_m^1\right) \ln \frac{p_m^1}{p_m^0}}{\sum_{m=1}^{M} h\left(s_m^0, s_m^1\right)} \qquad (20.72)$$

where

$$h(a,b) \equiv [(1/2)a^{-1} + (1/2)b^{-1}]^{-1} = 2ab/(a + b) \qquad (20.73)$$

and $h(a,b)$ is the harmonic mean of the numbers a and b. Thus γ^{**} is a share-weighted average of the logarithms of the price ratios p_m^1/p_m^0. If γ^{**} is exponentiated, then an estimator α^{**} for the elementary aggregate α is obtained.

20.107 How does α^{**} compare to the three ideal elementary price indices defined in paragraphs 20.11 to 20.22? It can be shown[57] that α^{**} approximates those three indices to the second order around an equal price and quantity point; i.e., for most data sets, α^{**} will be very close to the Fisher, Törnqvist and Walsh elementary indices.

20.108 In fact, a slightly different weighted least squares problem that is similar to the minimization problem (20.70) will generate exactly the Törnqvist elementary index. Consider the following weighted least squares model:

$$\min_{\gamma,\delta's} \sum_{m=1}^{M} \frac{1}{2}\left(s_m^0 + s_m^1\right)\left(\ln p_m^0 - \delta_m\right)^2 + \sum_{m=1}^{M} \frac{1}{2}\left(s_m^0 + s_m^1\right)\left(\ln p_m^1 - \gamma - \delta_m\right)^2 \qquad (20.74)$$

Thus in the model (20.74), the logarithm of each item price quotation in each period is weighted by the arithmetic average of its expenditure shares in the two periods under consideration.

---

[56] Theil's approach is also pursued by Rao (2002), who considered a generalization of equation (20.70) to cover the case of many time periods.

[57] Using the techniques in Diewert (1978).

20.109 The γsolution to the minimization problem (20.74) is

$$\gamma^{***} = \sum_{m=1}^{M} \frac{1}{2}\left(s_m^0 + s_m^1\right)\ln\frac{p_m^1}{p_m^0} \qquad (20.75)$$

which is the logarithm of the Törnqvist elementary index. Thus the exponential of γ^{***} is precisely the Törnqvist price index.

20.110 The results in this section provide some weak support for the use of the Jevons elementary index, but they provide much stronger support for the use of weighted elementary indices of the type defined in paragraphs 20.11 to 20.22.

20.111 The results in this section also provide support for the use of value-based weights in hedonic regressions.

Conclusion

20.112 The main results in this chapter can be summarized as follows:

- In order to define a "best" elementary index number formula, it is necessary to have a target index number concept. In paragraphs 20.11 to 20.22, it is suggested that normal bilateral index number theory applies at the elementary level as well as at higher levels and hence the target concept should be one of the Fisher, Törnqvist or Walsh formulae.
- When aggregating the prices of the same narrowly defined item within a period, the narrowly defined unit value is a reasonable target price concept.
- The axiomatic approach to traditional elementary indices (i.e., no quantity or value weights are available) supports the use of the Jevons formula under all circumstances.[58] If the items in the elementary aggregate are homogeneous (i.e., they have the same unit of measurement), then the Dutot formula can be used. In the case of a heterogeneous elementary aggregate (the usual case), the Carruthers–Sellwood–Ward-Dalén formula can be used as an alternative to the

[58] One exception to this advice is when a price can be zero in one period and positive in another comparison period. In this situation, the Jevons index will fail and the corresponding item will have to be ignored in the elementary index, or the technique outlined in paragraphs 17.90 to 17.94 of Chapter 17 could be used.

Jevons formula, but both will give much the same numerical answers.

- The Carli index has an upward bias and the harmonic index has a downward bias.
- The economic approach to elementary indices weakly supports the use of the Jevons formula.
- None of the five unweighted elementary indices is really satisfactory. A much more satisfactory approach would be to collect quantity or value information along with price information, and form sample superlative indices as the preferred elementary indices. If a chained superlative index is calculated, however, it should be examined for chain drift; i.e., a chained index should only be used if the data are relatively smooth and subject to long-term trends rather than short-term fluctuations.
- A simple hedonic regression approach to elementary indices supports the use of the Jevons formula, but a weighted hedonic regression approach is more satisfactory. The resulting index will closely approximate the ideal indices defined in paragraphs 20.11 to 20.22.