AN INTRODUCTION TO CONSUMER PRICE INDEX METHODOLOGY 1

1.1 A price index is a measure of the proportionate, or percentage, changes in a set of prices over time. A consumer price index (CPI) measures changes in the prices of goods and services that households consume. Such changes affect the real purchasing power of consumers' incomes and their welfare. As the prices of different goods and services do not all change at the same rate, a price index can only reflect their average movement. A price index is typically assigned a value of unity, or 100, in some reference period and the values of the index for other periods of time are intended to indicate the average proportionate, or percentage, change in prices from this price reference period. Price indices can also be used to measure differences in price levels between different cities, regions or countries at the same point in time.

1.2 Much of this manual and the associated economic literature on price indices is concerned with two basic questions:

- Exactly what set of prices should be covered by the index?
- What is the most appropriate way in which to average their movements?

These two questions are addressed in the early sections of this introduction.

1.3 Consumer price indices (CPIs) are index numbers that measure changes in the prices of goods and services purchased or otherwise acquired by households, which households use directly, or indirectly, to satisfy their own needs and wants. Consumer price indices can be intended to measure either the rate of price inflation as perceived by households, or changes in their cost of living (that is, changes in the amounts that the households need to spend in order to maintain their standard of living). There need be no conflict between these two objectives. In practice, most CPIs are calculated as weighted averages of the percentage price changes for a specified set, or "basket", of consumer products, the weights reflecting their relative importance in household consumption in some period. Much depends on how appropriate and timely the weights are.

1.4 This chapter provides a general introduction to, and overview of, the methodology for compiling CPIs. It provides a summary of the relevant theory and practice of index number compilation that is intended to facilitate the reading and understanding of the detailed chapters that follow, some of which are inevitably quite technical. It describes all the various steps involved in CPI compilation starting with the basic concept, definition and purpose of a CPI, followed by the sampling procedures and survey methods used to collect and process the price data, and finishing with a summary of the actual calculation of the index and its dissemination.

1.5 An introductory presentation of CPI methodology has to start with the basic concept of a CPI and the underlying index number theory, including the properties and behaviour of the various kinds of index number that are, or might be, used for CPI purposes. In principle, it is necessary to settle what type of index to calculate before going on to consider the best way in which to estimate it in practice, taking account of the resources available.

1.6 The main topics covered in this chapter are as follows:

- the origins and uses of CPIs;
- basic index number theory, including the axiomatic and economic approaches to CPIs;
- elementary price indices and aggregate CPIs;
- the transactions, activities and households covered by CPIs;
- the collection and processing of the prices, including adjusting for quality change;
- the actual calculation of the CPI;
- potential errors and bias;
- organization, management and dissemination policy.

In contrast, in this manual, the chapters dealing with index theory come later on; thus the presentation in this chapter does not follow the same order as the corresponding chapters of the manual.

1.7 It is not the purpose of this introduction to provide a complete summary of the contents of the manual. The objective is rather to provide a short presentation of the core methodological issues with which readers need to be acquainted before tackling the detailed chapters that follow. Some special topics, such as the treatment of certain individual products whose prices cannot be directly observed, are not considered here as they are not central to CPI methodology.

The origins and uses of consumer price indices

1.8 CPIs must serve a purpose. The precise way in which they are defined and constructed depends very much on what they are meant to be used for, and by whom. As explained in Chapter 15, CPIs have a long history dating back to the eighteenth century. Laspeyres and Paasche indices, which are still widely used today, were first proposed in the 1870s. They are explained below. The concept of the cost of living index was introduced early in the twentieth century.

1.9 Traditionally, one of the main reasons for compiling a CPI was to compensate wage-earners for inflation by adjusting their wage rates in proportion to the percentage change in the CPI, a procedure known as indexation. For this reason, official CPIs tended to become the responsibility of ministries of labour, but most are now compiled by national statistical offices. A CPI that is specifically intended to be used to index wages is known as a compensation index.

1.10 CPIs have three important characteristics. They are published *frequently*, usually every month but sometimes every quarter. They are available *quickly*, usually about two weeks after the end of the month or quarter. They are also usually *not revised*. CPIs tend to be closely monitored and attract a lot of publicity.

1.11 As CPIs provide timely information about the rate of inflation, they have also come to be used for a wide variety of purposes in addition to indexing wages. For example:

- CPIs are widely used to index pensions and social security benefits.
- CPIs are also used to index other payments, such as interest payments or rents, or the prices of bonds.
- CPIs are also commonly used as a proxy for the general rate of inflation, even though they measure only consumer inflation. They are used by some governments or central banks to set inflation targets for purposes of monetary policy.
- The price data collected for CPI purposes can also be used to compile other indices, such as the price indices used to deflate household consumption expenditures in national accounts, or the purchasing power parities used to compare real levels of consumption in different countries.

1.12 These varied uses can create conflicts of interest. For example, using a CPI as an indicator of general inflation may create pressure to extend its coverage to include elements that are not goods and services consumed by households, thereby changing the nature and concept of the CPI. It should also be noted that because of the widespread use of CPIs to index a wide variety of payments – not just wages, but social security benefits, interest payments, private contracts, etc. – extremely large sums of money turn on their movements, enough to have a significant impact on the state of government finances. Thus, small differences in the movements of CPIs resulting from the use of slightly different formulae or methods can have considerable financial implications. CPI methodology is important in practice and not just in theory.

Choice of index number

1.13 The first question is to decide on the kind of index number to use. The extensive references dealing with index theory in the bibliography reflect the fact that there is a very large literature on this subject. Many different kinds of mathematical formulae have been proposed over the past two centuries. While there may be no single formula that would be preferred in all circumstances, most economists and compilers of CPIs seem to be agreed that, in principle, the index formula should belong to a small class of indices called *superlative* indices. A superlative index may be expected to provide an approximation to a cost of living index. A characteristic feature of a superlative index is that it treats the prices and quantities in both periods being compared symmetrically. Different superlative indices tend to have similar properties, yield similar results and behave in very similar ways. Because of their properties of symmetry, some kind of superlative index is also likely to be seen as desirable, even when the CPI is not meant to be a cost of living index.

1.14 When a monthly or quarterly CPI is first published, however, it is invariably the case that there is not sufficient information on the quantities and expenditures in the current period to make it possible to calculate a symmetric, or superlative, index. While it is necessary to resort to second-best alternatives in practice, being able to make a rational choice between the various possibilities means having a clear idea of the target index that would be preferred in principle. The target index can have a considerable influence on practical matters such as the frequency with which the weights used in the index should be updated.

1.15 A comprehensive, detailed, rigorous and up-to-date discussion of the relevant index number theory is provided in Chapters 15 to 23 of the manual. The following sections provide a summary of this material. Proofs of the various propositions or theorems stated in this chapter are to be found in the later chapters, to which the reader may refer for further explanation.

Price indices based on baskets of goods and services

1.16 The purpose of an index number may be explained as comparing the *values* of households' expenditures on consumer goods and services in two time periods. Knowing that expenditures have increased by 5 per cent is not very informative if we do not know how much of this change is attributable to changes in the *prices* of the goods and services, and how much to changes in the *quantities* purchased. The purpose of an index number is to decompose proportionate or percentage changes in value aggregates into their overall components of price and quantity change. A CPI is intended to measure the price component of the change in households' consumption expenditures. One way to do this is to measure the change in the value of an aggregate, holding the quantities constant.

Lowe indices

1.17 One very wide, and popular, class of price indices is obtained by defining the index as the percentage change, between the periods compared, in the total cost of purchasing a given set of quantities, generally described as a "basket". The meaning of such an index is easy to grasp and to explain to users. This class of index is called a Lowe index in this manual, after the index number pioneer who first proposed it in 1823 (see Chapter 15). Most statistical offices make use of some kind of Lowe index in practice.

1.18 Let there be n products in the basket with prices p_i and quantities q_i, and let the two periods compared be 0 and t. The Lowe index, P_{Lo}, is defined as follows:

$$P_{Lo} \equiv \frac{\sum_{i=1}^{n} p_i^t q_i}{\sum p_i^0 q_i}$$

1.19 In principle, any set of quantities could serve as the basket. The basket does not have to be restricted to the quantities purchased in one or other of the two periods compared, or indeed any actual period of time. The quantities could, for example, be arithmetic or geometric averages of the quantities in the two periods. For practical reasons, the basket of quantities used for CPI purposes usually has to be based on a survey of household consumption expenditures conducted in an earlier period than either of the two periods whose prices are compared. For example, a monthly CPI may run from January 2000 onwards, with January 2000 = 100, but the quantities may be derived from an annual expenditure survey made in 1997 or 1998, or even spanning both those years. As it takes a long time to collect and process expenditure data, there is usually a considerable time lag before such data can be introduced into the calculation of CPIs. The basket may also refer to a year, whereas the index may be compiled monthly or quarterly.

1.20 The period whose quantities are actually used in a CPI is described as the *weight reference period* and it will be denoted as period b. Period 0 is the *price reference period*. As just noted, b is likely to precede 0, at least when the index is first published, and this is assumed here, but b could be any period, including one between 0 and t, if the index is calculated some time after t. The Lowe index using the quantities of period b can be written as follows:

$$P_{Lo} \equiv \frac{\sum_{i=1}^{n} p_i^t q_i^b}{\sum_{i=1}^{n} p_i^0 q_i^b} \equiv \sum_{i=1}^{n} (p_i^t / p_i^0) s_i^{0b}$$

$$\text{where} \quad s_i^{0b} = \frac{p_i^0 q_i^b}{\sum_{i=1}^{n} p_i^0 q_i^b} \qquad (1.1)$$

The index can be written, and calculated, in two ways: either as the ratio of two value aggregates, or as an arithmetic weighted average of the price ratios, or *price relatives*, p_i^t/p_i^0, for the individual products using the hybrid expenditure shares s_i^{0b} as weights. The expenditures are described as *hybrid* because the prices and quantities belong to two different time periods, 0 and b respectively. The hybrid weights may be obtained by updating the actual expenditure shares in period b, namely $p_i^b q_i^b / \sum p_i^b q_i^b$, for the price changes occurring between periods b and 0 by multiplying them by the price relatives b and 0, namely p_i^0/p_i^b. Lowe indices are widely used for CPI purposes.

Laspeyres and Paasche indices

1.21 Any set of quantities could be used in a Lowe index, but there are two special cases which figure very prominently in the literature and are of considerable importance from a theoretical point of view. When the quantities are those of the price reference period, that is when $b = 0$, the *Laspeyres* index is obtained. When quantities are those of the other period, that is when $b = t$, the *Paasche* index is obtained. It is necessary to consider the properties of Laspeyres and Paasche indices, and also the relationships between them, in more detail.

1.22 The Laspeyres price index, P_L, is defined as:

$$P_L = \frac{\sum_{i=1}^{n} p_i^t q_i^0}{\sum_{i=1}^{n} p_i^0 q_i^0} \equiv \sum_{i=1}^{n} (p_i^t / p_i^0) s_i^0 \qquad (1.2)$$

where s_i^0 denotes the share of the *actual* expenditure on commodity i in period 0: that is, $p_i^0 q_i^0 / \sum p_i^0 q_i^0$.

1.23 The Paasche index, P_P, is defined as:

$$P_P = \frac{\sum_{i=1}^{n} p_i^t q_i^t}{\sum_{i=1}^{n} p_i^0 q_i^t} \equiv \left\{ \sum_{i=1}^{n} (p_i^t / p_i^0)^{-1} s_i^t \right\}^{-1} \qquad (1.3)$$

where s_i^t denotes the actual share of the expenditure on commodity i in period t; that is, $p_i^t q_i^t / \sum p_i^t q_i^t$. Notice that the Paasche index is a weighted *harmonic* average of the price relatives that uses the actual expenditure shares in the later period t as weights. It follows from equation (1.1) that the Paasche index can also be expressed as a weighted arithmetic average of the price relatives using hybrid expenditure weights, in which the quantities of t are valued at the prices of 0.

Decomposing current value changes using Laspeyres and Paasche indices

1.24 Laspeyres and Paasche quantity indices are defined in a similar way to the price indices, simply by interchanging the p and q values in formulae (1.2) and (1.3). They summarize changes over time in the flow of quantities of goods and services consumed. A Laspeyres quantity index values the quantities at the fixed prices of the earlier period, while the Paasche quantity index uses the prices of the later period. The ratio of the values of the expenditures in two periods (V) reflects the combined effects of both price and quantity changes. When Laspeyres and Paasche indices are used, the value change can be exactly decomposed into a price index times a quantity index only if the Laspeyres price (quantity) index is matched with the Paasche quantity (price) index. Let P_{La} and Q_{La} denote the Laspeyres price and quantity indices and let P_{Pa} and Q_{Pa} denote the Paasche price and quantity indices: then, $P_{La} Q_{Pa} \equiv V$ and $P_{Pa}\, Q_{La} \equiv V$.

1.25 Suppose, for example, a time series of household consumption expenditures at current prices in the national accounts is to be deflated by a price index to show changes in real consumption. To generate a series of consumption expenditures at constant base period prices (whose movements are identical with those of the Laspeyres volume index), the consumption expenditures at current prices must be deflated by a series of Paasche price indices.

Ratios of Lowe and Laspeyres indices

1.26 The Lowe index is transitive. The ratio of two Lowe indices using the same set of q^b values is also a Lowe index. For example, the ratio of the Lowe index for period $t+1$ with price reference period 0 divided by that for period t also with price reference period 0 is:

$$\frac{\sum_{i=1}^{n} p_i^{t+1} q_i^b \Big/ \sum_{i=1}^{n} p_i^0 q_i^b}{\sum_{i=1}^{n} p_i^t q_i^b \Big/ \sum_{i=1}^{n} p_i^0 q_i^b} = \frac{\sum_{i=1}^{n} p_i^{t+1} q_i^b}{\sum_{i=1}^{n} p_i^t q_i^b} = P_{Lo}^{t,t+1} \tag{1.4}$$

This is a Lowe index for period $t+1$ with period t as the price reference period. This kind of index is, in fact, widely used to measure short-term price movements, such as between t and $t+1$, even though the quantities may date back to some much earlier period b.

1.27 A Lowe index can also be expressed as the ratio of two Laspeyres indices. For example, the Lowe index for period t with price reference period 0 is equal to the Laspeyres index for period t with price reference period b divided by the Laspeyres index for period 0 also with price reference period b. Thus,

$$P_{Lo} = \frac{\sum_{i=1}^{n} p_i^t q_i^b}{\sum_{i=1}^{n} p_i^0 q_i^b} = \frac{\sum_{i=1}^{n} p_i^t q_i^b \Big/ \sum_{i=1}^{n} p_i^b q_i^b}{\sum_{i=1}^{n} p_i^0 q_i^b \Big/ \sum_{i=1}^{n} p_i^b q_i^b} = \frac{P_{La}^t}{P_{La}^0} \tag{1.5}$$

Updated Lowe indices

1.28 It is useful to have a formula that enables a Lowe index to be calculated directly as a chain index, in which the index for period $t+1$ is obtained by updating the index for period t. Because Lowe indices are transitive, the Lowe index for period $t+1$ with price reference period 0 can be written as the product of the Lowe index for period t with price reference period 0 multiplied by the Lowe index for period $t+1$ with price reference period t. Thus,

$$\frac{\sum_{i=1}^{n} p_i^{t+1} q_i^b}{\sum_{i=1}^{n} p_i^0 q_i^b} = \left[\frac{\sum_{i=1}^{n} p_i^t q_i^b}{\sum_{i=1}^{n} p_i^0 q_i^b}\right]\left[\frac{\sum_{i=1}^{n} p_i^{t+1} q_i^b}{\sum_{i=1}^{n} p_i^t q_i^b}\right] = \left[\frac{\sum_{i=1}^{n} p_i^t q_i^b}{\sum_{i=1}^{n} p_i^0 q_i^b}\right]\left[\sum_{i=1}^{n}\left(\frac{p_i^{t+1}}{p_i^t}\right) s_i^{tb}\right] \tag{1.6}$$

where the expenditure weights s_i^{tb} are hybrid weights defined as:

$$s_i^{tb} \equiv p_i^t q_i^b \Big/ \sum_{i=1}^{n} p_i^t q_i^b \tag{1.7}$$

1.29 Hybrid weights of the kind defined in equation (1.7) are often described as *price-updated* weights. They can be obtained by adjusting the original expenditure weights $p_i^b q_i^b / \sum p_i^b q_i^b$ by the price relatives p_i^t / p_i^b. By price-updating the expenditure weights from b to t in this way, the index between t and $t+1$ can be calculated directly as a weighted average of the price relatives p_i^{t+1}/p_i^t without referring back to the price reference period 0. The index can then be linked on to the value of the index in the preceding period t.

Interrelationships between fixed basket indices

1.30 Consider first the interrelationship between the Laspeyres and the Paasche indices. A well-known result in index number theory is that if the price and quantity changes (weighted by values) are *negatively* correlated, then the Laspeyres index exceeds the Paasche index. Conversely, if the weighted price and quantity changes are *positively* correlated, then the Paasche index exceeds the Laspeyres index. The proof is given in Appendix 15.1 of Chapter 15.

1.31 As consumers are usually price-takers, they typically react to price changes by substituting goods or services that have become *relatively* cheaper for those that have become *relatively* dearer. This is known as the *substitution effect*, a phenomenon that figures prominently in this manual and the wider literature on index numbers. Substitution tends to generate a negative correlation between the price and quantity relatives, in which case the Laspeyres index is greater than the Paasche index, the gap between them tending to widen over time.

1.32 In practice, however, statistical offices do not calculate Laspeyres or Paasche indices but instead usually calculate Lowe indices as defined in equation (1.1). The question then arises of how the Lowe index relates to the Laspeyres and Paasche indices. It is shown in the text of Chapter 15, and also in Appendix 15.2, that if there are persistent long-term trends in relative prices and if the substitution effect is operative, the Lowe index will tend to exceed the Laspeyres, and therefore also the Fisher and the Paasche indices. Assuming that period b precedes period 0, the ranking under these conditions will be:

$$\text{Lowe} \geq \text{Laspeyres} \geq \text{Fisher} \geq \text{Paasche}$$

Moreover, the amount by which the Lowe exceeds the other three indices will tend to increase, the further back in time period b is in relation to period 0.

1.33 The positioning of period b is crucial. Given the assumptions about long-term price trends and substitution, a Lowe index will tend to increase as period b is moved backwards in time, or to decrease as period b is moved forwards in time. While b may have to precede 0 when the index is first published, there is no such restriction on the positioning of b as price and quantity data become available for later periods with passage of time. Period b can then be moved forwards. If b is positioned midway between 0 and t, the quantities are likely to be equi-representative of both periods, assuming that there is a fairly smooth transition from the relative quantities of 0 to those of t. In these circumstances, the Lowe index is likely to be close to the Fisher and other superlative indices, and cannot be presumed to have either an upward or a downward bias. These points are elaborated further below, and also in Chapter 15.

1.34 It is important that statistical offices take these relationships into consideration in deciding upon their policies. There are obviously practical advantages and financial savings from continuing to make repeated use over many years of the same fixed set of quantities to calculate a CPI. However, the amount by which such a CPI exceeds some conceptually preferred target index, such as a cost of living index (COLI), is likely to get steadily larger the further back in time the period b to which the quantities refer. Most users are likely to interpret the difference as upward bias. A large bias may undermine the credibility and acceptability of the index.

Young index

1.35 Instead of holding constant the quantities of period b, a statistical office may calculate a CPI as a weighted arithmetic average of the individual price relatives, holding constant the revenue shares of period b. The resulting index is called a *Young* index in this manual, again after another index number pioneer. The Young index is defined as follows:

$$P_{Yo} \equiv \sum_{i=1}^{n} s_i^b \left(\frac{p_i^t}{p_i^0}\right) \quad \text{where} \quad s_i^b \equiv \frac{p_i^b q_i^b}{\sum_{i=1}^{n} p_i^b q_i^b} \tag{1.8}$$

In the corresponding Lowe index, equation (1.1), the weights are hybrid revenue shares that value the quantities of b at the prices of 0. As already explained, the price reference period 0 is usually later than the weight reference period b because of the time needed to collect and process and revenue data. In that case, a statistical office has the choice of assuming that either the quantities of period b remain constant or the expenditure shares in period b remain constant. Both cannot remain constant if prices change between b and 0. If the expenditure shares actually remained constant between periods b and 0, the quantities must have changed inversely in response to the price changes, which implies an elasticity of substitution of unity.

1.36 Whereas there is a presumption that the Lowe index will tend to exceed the Laspeyres index, it is more difficult to generalize about the relationship between the Young index and the Laspeyres index. The Young could be greater or less than the Laspeyres depending on how sensitive the quantities are to changes in relative prices. It is shown in Chapter 15 that with high elasticities of substitution (greater than unity) the Young will tend to exceed the Laspeyres, whereas with low elasticities the Young will tend to be less than the Laspeyres.

1.37 As explained later in this chapter, the Lowe index may be preferred to the Young index because the Young index has some undesirable properties that cause it to fail some critical index number tests (see also Chapter 16).

Geometric Young, Laspeyres and Paasche indices

1.38 In the geometric version of the Young index, a weighted geometric average is taken of the price relatives using the expenditure shares of period b as weights. It is defined as follows:

$$P_{GYo} \equiv \prod_{i=1}^{n} \left(\frac{p_i^t}{p_i^0}\right)^{s_i^b} \tag{1.9}$$

where s_i^b is defined as above. The geometric Laspeyres is the special case in which $b = 0$; that is, the expenditure shares are those of the price reference period 0. Similarly, the geometric Paasche uses the expenditure shares of period t. It should be noted that these geometric indices cannot be expressed as the ratios of value aggregates in which the quantities are fixed. They are not basket indices and there are no counterpart Lowe indices.

1.39 It is worth recalling that for any set of positive numbers the arithmetic average is greater than, or equal to, the geometric average, which in turn is greater than, or equal to, the harmonic average, the equalities holding only when the numbers are all equal. In the case of unitary cross-elasticities of demand and constant expenditure shares, the geometric Laspeyres and Paasche indices coincide. In this case, the ranking of the indices must be the ordinary Laspeyres $\geq$ the geometric Laspeyres and Paasche $\geq$ the ordinary Paasche, because the indices are, respectively, arithmetic, geometric and harmonic averages of the same price relatives which all use the same set of weights.

1.40 The geometric Young and Laspeyres indices have the same information requirements as their ordinary arithmetic counterparts. They can be produced on a timely basis. Thus, these geometric indices must be treated as serious practical possibilities for purposes of CPI calculations. As explained later, the geometric indices are likely to be less subject than their arithmetic counterparts to the kinds of index number biases discussed in later sections. Their main disadvantage may be that, because they are not fixed basket indices, they are not so easy to explain or justify to users.

Symmetric indices

1.41 A symmetric index is one that makes equal use of the prices and quantities in both the periods compared and treats them in a symmetric manner. There are three particular symmetric indices that are widely used in economic statistics. It is convenient to introduce them at this point. As already noted, these three indices are also superlative indices.

1.42 The first is the *Fisher price index*, P_F, defined as the *geometric* average of the Laspeyres and Paasche indices; that is,

$$P_F \equiv \sqrt{P_L P_P} \qquad (1.10)$$

1.43 The second is the *Walsh price index,* P_W. This is a basket index whose quantities consist of *geometric* averages of the quantities in the two periods; that is,

$$P_W \equiv \frac{\sum_{i=1}^{n} p_i^t \sqrt{q_i^t q_i^0}}{\sum_{i=1}^{n} p_i^0 \sqrt{q_i^t q_i^0}} \qquad (1.11)$$

By taking a *geometric* rather than an arithmetic average of the quantities, equal weight is given to the *relative* quantities in both periods. The quantities in the Walsh index can be regarded as being equi-representative of both periods.

1.44 The third index is the *Törnqvist price index,* P_T, defined as a *geometric* average of the price relatives weighted by the average expenditure shares in the two periods.

$$P_T = \prod_{i=1}^{n} (p_i^t / p_i^0)^{\sigma_i} \qquad (1.12)$$

where σ_i is the arithmetic average of the share of expenditure on product i in the two periods.

$$\sigma_i = \frac{S_i^t + S_i^0}{2} \qquad (1.13)$$

where the s_i values are defined as in equations (1.2) and (1.3) above.

1.45 The theoretical attractions of these indices become more apparent in the following sections on the axiomatic and economic approaches to index numbers.

Fixed base versus chain indices

1.46 This topic is examined in Chapter 15. When a time series of Lowe or Laspeyres indices is calculated using a fixed set of quantities, the quantities become progressively out of date and increasingly irrelevant to the later periods for which prices are being compared. The base period, in which quantities are set, has to be updated sooner or later and the new index series linked to the old. Linking is inevitable in the long run.

1.47 In a chain index, each link consists of an index in which each period is compared with the preceding one, the weight and price reference periods being moved forward each period. Any index number formula can be used for the individual links in a chain index. For example, it is possible to have a chain index in which the index for $t+1$ on t is a Lowe index defined as $\sum p^{t+1} q^{t-j} / \sum p^t q^{t-j}$. The quantities refer to some period that is j periods earlier than the price reference period t. The quantities move forward one period as the price reference period moves forward one period. If $j=0$, the chain Lowe becomes a chain Laspeyres, while if $j=-1$, it becomes a chain Paasche.

1.48 The CPIs in some countries are, in fact, annual chain Lowe indices of this general type, the quantities referring to some year or years that precede the price reference period 0 by a fixed period. For example, the 12 monthly indices from January 2000 to January 2001, with January 2000 as the price reference period, could be Lowe indices based on price-updated expenditures for 1998. The 12 indices from January 2001 to January 2002 are then based on price updated expenditures for 1999, and so on.

1.49 The expenditures lag behind the January price reference period by a fixed interval, moving forward a year each January as the price reference period moves forward one year. Although, for practical reasons, there has to be a time lag between the quantities and the prices when the index is first published, it is possible to recalculate the monthly indices for the current year later, using current expenditure data when they eventually become available. In this way, it is possible for the long-run index to be an annually chained monthly index, with contemporaneous annual weights. This method is explained in more detail in Chapter 9. It is used by one statistical office.

1.50 A chain index has to be "path dependent". It must depend on the prices and quantities in all the intervening periods between the first and last periods in the index series. Path dependency can be advantageous or disadvantageous. When there is a gradual economic transition from the first to the last period, with smooth trends in relative prices and quantities, chaining will tend to reduce the index number spread between the Lowe, Laspeyres and Paasche indices, thereby making the movements in the index less dependent on the choice of index number formula.

1.51 If there are fluctuations in the prices and quantities in the intervening periods, however, chaining may not only increase the index number spread but also distort the measure of the overall change between the first and last periods. For example, suppose all the prices in the last period return to their initial levels in period 0, which implies that they must have fluctuated in between; a chain Laspeyres index will not return to 100. It will tend to be greater than 100. If the cycle is repeated with all the prices periodically returning to their original levels, a chain Laspeyres index will tend to drift further and further above 100 even though there may be no long-term upward trend in the prices. Chaining is therefore not advised when prices fluctuate. When monthly prices are subject to regular and substantial seasonal fluctuations, for example, monthly chaining cannot be recommended. Seasonal fluctuations cause serious problems, which are analysed in Chapter 22. While a number of countries update their expenditure weights annually, the 12-monthly indices within each year are not chain indices but Lowe indices using fixed annual quantities.

1.52 *The Divisia index*. If the prices and quantities are continuous functions of time, it is possible to partition the change in their total value over time into price and quantity components following the method of Divisia. As shown in Chapter 15, the Divisia index may be derived mathematically by differentiating value (i.e. price multiplied by quantity) with respect to time to obtain two components: a relative-value-weighted price change and a relative-value-weighted quantity change.

These two components are defined to be price and quantity indices, respectively. The Divisia is essentially a theoretical index. In practice, prices can be recorded only at discrete intervals, even if they vary continuously with time. A chain index may, however, be regarded as a discrete approximation to a Divisia. The Divisia index itself offers limited practical guidance about the kind of index number formula to choose for the individual links in a chain index.

Axiomatic and stochastic approaches to index numbers

1.53 Various *axiomatic approaches* to index numbers are explained in Chapter 16. These approaches seek to determine the most appropriate functional form for an index by specifying a number of axioms, or tests, that the index ought to satisfy. They throw light on the properties possessed by different kinds of indices, some of which are not intuitively obvious. Indices that fail to satisfy certain basic or fundamental axioms, or tests, may be rejected completely because they are liable to behave in unacceptable ways. An axiomatic approach may also be used to rank indices on the basis of their desirable, and undesirable, properties.

First axiomatic approach

1.54 The first approach is the traditional test approach pioneered by Irving Fisher. The price and quantity indices are defined as functions of the two vectors of prices and two vectors of quantities relating to the two periods compared. The prices and quantities are treated as independent variables, whereas in the economic approach to index numbers considered later in this chapter the quantities are assumed to be functions of the prices.

1.55 Chapter 16 starts by considering a set of 20 axioms, but only a selection of them is given here by way of illustration.

T1: *positivity* – the price index and its constituent vectors of prices and quantities should be positive.

T3: *identity test* – if the price of every product is identical in both periods, then the price index should equal unity, no matter what the quantity vectors are.

T5: *proportionality in current prices* – if all prices in period t are multiplied by the positive number λ, then the new price index should be λ times the old price index; i.e., the price index function is (positively) homogeneous of degree one in the components of the period t price vector.

T10: *invariance to changes in the units of measurement* (commensurability test) – the price index does not change if the units in which the products are measured are changed.

T11: *time reversal test* – if all the data for the two periods are interchanged, then the resulting price index should equal the reciprocal of the original price index.

T14: *mean value test for prices* – the price index lies between the highest and the lowest price relatives.

T16: *Paasche and Laspeyres bounding test* – the price index lies between the Laspeyres and Paasche indices.

T17: *monotonicity in current prices* – if any period t price is increased, then the price index must increase.

1.56 Some of the axioms or tests can be regarded as more important than others. Indeed, some of the axioms seem so inherently reasonable that it might be assumed that any index number actually in use would satisfy them. For example, test T10, the commensurability test, says that if the unit of quantity in which a product is measured is changed, say, from a gallon to a litre, the index must be unchanged. One index that does not satisfy this test is the *Dutot* index, which is defined as the ratio of the arithmetic means of the prices in the two periods. As explained later, this is a type of elementary index that is in fact widely used in the early stages of CPI calculation.

1.57 Consider, for example, the average price of salt and pepper. Suppose it is decided to change the unit of measurement for pepper from grams to ounces while leaving the units in which salt is measured (for example, kilos) unchanged. As an ounce is equal to 28.35 grams, the absolute value of the price of pepper increases by over 28 times, which effectively increases the weight of pepper in the Dutot index by over 28 times.

1.58 When the products covered by an index are heterogeneous and measured in different physical units, the value of any index that does not satisfy the commensurability test depends on the purely arbitrary choice of units. Such an index must be unacceptable conceptually. If the prices refer to a strictly homogeneous set of products that all use the same unit of measurement, the test becomes irrelevant.

1.59 Another important test is T11, the time reversal test. In principle, it seems reasonable to require that the same result should be obtained whichever of the two periods is chosen as the price reference period: in other words, whether the change is measured forwards in time, i.e., from 0 to t, or backwards in time from t to 0. The Young index fails this test because an arithmetic average of a set of price relatives is not equal to the reciprocal of the arithmetic average of the reciprocals of the price relatives. The fact that the *conceptually* arbitrary decision to measure the change in prices forwards from 0 and t gives a different result from measuring backwards from t to 0 is seen by many users as a serious disadvantage. The failure of the Young index to satisfy the time reversal test needs to be taken into account by statistical offices.

1.60 Both the Laspeyres and Paasche fail the time reversal test for the same reasons as the Young index. For example, the formula for a Laspeyres calculated backwards from t to 0, P_{BL}, is:

$$P_{BL} = \frac{\sum_{i=1}^{n} p_i^0 q_i^t}{\sum_{i=1}^{n} p_i^t q_i^t} \equiv \frac{1}{P_P} \qquad (1.14)$$

This index is identical to the reciprocal of the (forwards) Paasche, not to the reciprocal of the (forwards) Laspeyres. As already noted, the (forwards) Paasche tends to register a smaller increase than the (forwards) Laspeyres so that the Laspeyres index cannot satisfy the time reversal test. The Paasche index also fails the time reversal test.

1.61 In contrast, the Lowe index satisfies the time reversal test *provided* that the quantities q_i^b remain fixed when the price reference period is changed from 0 to t. The quantities of a Laspeyres index are, however, those of the price reference period *by definition*, and must change whenever the price reference period is changed. The basket for a forwards Laspeyres is different from that for a backwards Laspeyres, and the Laspeyres fails the time reversal test in consequence.

1.62 Similarly, the Lowe index is transitive whereas the Laspeyres and Paasche indices are not. Assuming that a Lowe index uses a fixed set of quantities, q_i^b, whatever the price reference period, it follows that

$$Lo^{0,t} = Lo^{0,t-k} Lo^{t-k,t}$$

where $Lo^{0,t}$ is the Lowe index for period t with period 0 as the price reference period. The Lowe index that compares t directly with 0 is the same as that calculated indirectly as a chain index through period $t-k$.

1.63 If, on the other hand, the Lowe index is defined in such a way that quantities vary with the price reference period, as in the index $\sum p^{t+1}q^{t-j}/\sum p^t q^{t-j}$ considered earlier, the resulting chain index is not transitive. The chain Laspeyres and chain Paasche indices are special cases of this index.

1.64 In the real world, the quantities do change and the whole point of chaining is to enable the *quantities* to be continually updated to take account of the changing universe of products. Achieving transitivity by arbitrarily holding the quantities constant, especially over a very long period of time, does not compensate for the potential biases introduced by using out-of-date quantities.

Ranking of indices using the first axiomatic approach

1.65 In Chapter 16 it is shown not only that the Fisher price index satisfies all the 20 axioms listed but also, more remarkably, that it is the only possible index that can satisfy all 20 axioms. Thus, on the basis of this particular set of axioms, the Fisher clearly dominates other indices.

1.66 In contrast to Fisher, the other two symmetric (and superlative) indices defined in equations (1.11) and (1.12) above do not emerge so well from the 20 tests. In Chapter 16, it is shown that the Walsh price index fails four tests while the Törnqvist index fails nine tests. Nevertheless, the Törnqvist and the Fisher may be expected to approximate each other quite closely numerically when the data follow relatively smooth trends, as shown in Chapter 19.

1.67 One limitation of the axiomatic approach is that the list of axioms is inevitably somewhat arbitrary. Some axioms, such as the Paasche and Laspeyres bounding test failed by both Törnqvist and Walsh, could be regarded as dispensable. Additional axioms or tests can be envisaged, and two further axioms are considered below. Another problem with a simple application of the axiomatic approach is that it is not sufficient to know which tests are failed. It is also necessary to know how badly an index fails. Failing badly one major test, such as the commensurability test, might be considered sufficient to rule out an index, whereas failing several minor tests marginally may not be very disadvantageous.

Some further tests

1.68 Consider a further symmetry test. Reversing the roles of prices and quantities in a price index yields a quantity index of the same functional form as the price index. The *factor reversal test* requires that the product of this quantity index and the original price index should be identical with the change in the value of the aggregate in question. The test is important if, as stated earlier, price and quantity indices are intended to enable changes in the values of aggregates over time to be factored into their price and quantity components in an economically meaningful way. Another interesting result given in Chapter 16 is that the Fisher index is the only price index to satisfy four minimal tests: T1 (positivity), T11 (time reversal test), T12 (quantity reversal test) and T21 (factor reversal test). As the factor reversal test implicitly assumes that the prices and quantities must refer either to period 0 or to period t, it is not relevant to a Lowe index in which three periods are involved, b, 0 and t.

1.69 As shown earlier, the product of the Laspeyres price (quantity) index and the Paasche quantity (price) index is identical with the change in the total value of the aggregate in question. Thus, Laspeyres and Paasche indices may be said to satisfy a weak version of the factor reversal test in that dividing the value change by a Laspeyres (Paasche) price index does lead to a meaningful quantity index, i.e., the Paasche (Laspeyres), even though the functional forms of the price and quantity indices are not identical.

1.70 Another test discussed in Chapter 16 is the *additivity test*. This is more important from the perspective of quantity than price indices. Price indices may be used to deflate value changes to obtain implicit quantity changes. The results may be presented for sub-aggregates such as broad categories of household consumption. Just as expenditure aggregates at current prices are, by definition, obtained simply by summing individual expenditures, it is reasonable to expect that the changes in the sub-aggregates of a quantity index should add up to the changes in the totals – the additivity test. Quantity indices such as Laspeyres and Paasche that use a common set of prices to value quantities in both periods must satisfy the additivity test. Similarly, the Lowe quantity index defined as $\sum p^j q^t/\sum p^j q^0$ is also additive. The Geary–Khamis quantity index (see Annex 4) used to make international comparisons of real consumption and gross domestic product (GDP) between countries is an example of such a Lowe quantity index. It uses an arithmetically weighted average of the prices in the different countries as the common price vector p^j to compare the quantities in different countries.

1.71 Similarly, an average of the prices in two periods can be used to value the quantities in intertemporal indices. If the quantity index is also to satisfy the time reversal test, the average must be symmetrical. The *invariance to proportional changes in current prices test* (which corresponds to test T7 listed in Chapter 16, except that the roles of prices and quantities are reversed) requires that a quantity index depend only on the *relative*, not the absolute, level of the prices in each period. The Walsh quantity index satisfies this test, is additive and satisfies the time reversal test as well. It emerges as a quantity index with some very desirable properties.

1.72 Although the Fisher index itself is not additive, it is possible to decompose the overall *percentage change* in a Fisher price, or quantity, index into additive components that reflect the percentage change in each price or quantity. A similar multiplicative decomposition is possible for a Törnqvist price or quantity index.

The stochastic approach and a second axiomatic approach

1.73 Before considering a second axiomatic approach, it is convenient to take the stochastic approach to price indices. The stochastic approach treats the observed price *changes* or *relatives* as if they were a random sample drawn from a defined universe whose mean can be interpreted as the general rate of inflation. There can, however, be no single unique rate of inflation. Many possible universes can be defined, depending on which particular sets of expenditures or transactions the user is interested in. Clearly, the sample mean depends on the choice of universe from which the sample is drawn. Specifying the universe is similar to specifying the scope of a CPI. The stochastic approach addresses issues such as the appropriate form of average to take and the most efficient way to estimate it from a sample of price relatives, once the universe has been defined.

1.74 The stochastic approach is particularly useful when the universe is reduced to a single type of product. Because of market imperfections, there may be considerable variation in the prices at which the same product is sold in different outlets and also in the price changes observed. In practice, statistical offices have to estimate the average price change for a single product from a sample of price observations. Important methodological issues are raised, which are discussed in some detail in Chapter 7 and Chapter 20.

The unweighted stochastic approach

1.75 In Chapter 16, the unweighted stochastic approach to index number theory is explained. If simple random sampling has been used, equal weight may be given to each sampled price relative. Suppose each price relative can be treated as the sum of two components: a common inflation rate and a random disturbance with a zero mean. Using least squares or maximum likelihood, the best estimate of the common inflation rate is the unweighted *arithmetic* mean of price relatives, an index formula known as the *Carli* index. This index is the unweighted version of the Young index and is discussed further below, in the context of elementary price indices.

1.76 If the random component is multiplicative, not additive, the best estimate of the common inflation rate is given by the unweighted *geometric* mean of price relatives, known as the *Jevons* index. The Jevons index may be preferred to the Carli on the grounds that it satisfies the time reversal test, whereas the Carli does not. As explained below, this fact may be decisive when determining the functional form to be used to estimate the elementary indices compiled in the early stages of CPI calculations.

The weighted stochastic approach

1.77 As explained in Chapter 16, a *weighted* stochastic approach can be applied at an aggregative level covering sets of different products. As the products may be of differing economic importance, equal weight should not be given to each type of product. The products may be weighted on the basis of their share in the total value of the expenditures, or other transactions, in some period or periods. In this case, the index (or its logarithm) is the expected value of a random sample of price relatives (or their logarithms) whose probability of selection is proportional to the expenditure on that type of product in some period, or periods. Different indices are obtained depending on which expenditure weights are used and on whether the price relatives or their logarithms are used.

1.78 Suppose a sample of price relatives is randomly selected with the probability of selection proportional to the expenditure on that type of product in the price reference period 0. The expected price change is then the Laspeyres price index for the universe. Other indices may, however, also be obtained using the weighted stochastic approach. Suppose both periods are treated symmetrically and the probabilities of selection are made proportional to the arithmetic mean expenditure shares in both periods 0 and t. When these weights are applied to the logarithms of the price relatives, the expected value of the logarithms is the Törnqvist index, also known as the Törnqvist–Theil index. From an axiomatic viewpoint, the choice of a symmetric average of the expenditure shares ensures that the time reversal test is satisfied, while the choice of the arithmetic mean, as distinct from some other symmetric average, may be justified on the grounds that the fundamental proportionality in current prices test, T5, is thereby satisfied.

1.79 By focusing on price changes, the Törnqvist index emerges as an index with some very desirable properties. This suggests a second axiomatic approach to indices, in which the focus is shifted from the individual prices and quantities used in the traditional axiomatic approach, to price changes and values shares.

A second axiomatic approach

1.80 A second axiomatic approach is examined in Chapter 16 in which a price index is defined as a function of the two sets of prices, or their ratios, and two sets of values. Provided the index is invariant to changes in units of measurement, i.e., satisfies the commensurability test, it makes no difference whether individual prices or their ratios are specified. A set of 17 axioms is postulated which are similar to the 20 axioms considered in the first axiomatic approach.

1.81 It is shown in Appendix 16.1 that the Törnqvist, or Törnqvist–Theil, is the only price index to satisfy all 17 axioms, just as the Fisher price index is the only index to satisfy all 20 tests in the first approach. However, the Törnqvist index does not satisfy the factor reversal test, so that the implicit quantity index obtained by deflating the change in value by the Törnqvist price index is not the Törnqvist quantity index. The implicit quantity index is therefore not "best" in the sense of satisfying the 17 axioms when these are applied to the quantity, rather than price, indices.

1.82 Zero prices may cause problems for indices based on price ratios, and especially for geometric averages of price ratios. In particular, if any price tends to zero, one test that may be applied is that the price index ought not to tend to zero or plus infinity. The Törnqvist does not meet this test. It is therefore proposed in Chapter 16 that when using the Törnqvist index, care should be taken to bound the prices away from zero in order to avoid a meaningless index number.

1.83 Finally, Chapter 16 examines the axiomatic properties of the Lowe and Young indices. The Lowe index emerges quite well from the axiomatic approach, satisfying both the time reversal and circularity tests. On the other hand, the Young index, like the Laspeyres and Paasche indices, fails both tests. As already explained, however, the attractiveness of the Lowe index depends more on how relevant the fixed quantity weights are to the two periods being compared, that is on the positioning of period *b*, than its axiomatic properties.

1.84 Although the "best" indices emerging from the two axiomatic approaches, namely Fisher and Törnqvist, are not the same, they have much in common. As already noted, they are both symmetric indices and they are both superlative indices. Although their formulae are different, they may be expected to behave in similar ways and register similar price movements. The same *type* of indices keep emerging as having desirable properties whichever approach to index theory is adopted, a conclusion that is reinforced by the economic approach to index numbers, which is explained in Chapter 17.

Cost of living index

1.85 Approaching the consumer price index from the standpoint of economic theory has led to the development of the concept of a cost of living index (COLI). The theory of the COLI was first developed by Konus (1924). It rests on the assumption of optimizing behaviour on the part of a rational consumer. The COLI for such a consumer has been defined succinctly as the ratio of the minimum expenditures needed to attain the given level of utility, or welfare, under two different price regimes. A more precise definition and explanation are given in Chapter 17.

1.86 Whereas a Lowe index measures the change in the cost of purchasing a fixed basket of goods and services resulting from changes in their prices, a COLI measures the change in the *minimum* cost of maintaining a given level of utility, or welfare, that results from changes in the prices of the goods and services consumed.

1.87 A COLI is liable to possible misinterpretation because households' welfare depends on a variety of physical and social factors that have no connection with prices. Events may occur that impinge directly on welfare, such as natural or man-made disasters. When such events occur, households may need to increase their consumption of goods and services in order to compensate for the loss of welfare caused by those events. Changes in the costs of consumption triggered by events *other than changes in prices* are irrelevant for a CPI that is not merely intended to measure changes in the prices of consumer goods and services but is generally interpreted by users as measuring price changes, and only price changes. In order to qualify as a CPI, a COLI must therefore hold constant not only the consumer's preferences but all the non-price factors that affect the consumer's welfare and standard of living. If a CPI is intended to be a COLI it must be *conditional* on:

– a particular level of utility or welfare;
– a particular set of consumer preferences;
– a particular state of the physical and social environment.

Of course, Lowe indices are also conditional as they depend on the particular basket of goods and services selected.

1.88 Lowe indices and COLIs have in common the fact that they may both be defined as the ratios of expenditures in two periods. However, whereas, by definition, the quantities are fixed in Lowe indices, they vary in response to changes in relative prices in COLIs. In contrast to the fixed basket approach to index theory, the economic approach explicitly recognizes that the quantities consumed are actually dependent on the prices. In practice, rational consumers may be expected to adjust the *relative* quantities they consume in response to changes in *relative* prices. A COLI assumes that a consumer seeking to minimize the cost of maintaining a given level of utility will make the necessary adjustments. The baskets of goods and services in the numerator and denominator of a COLI are not therefore exactly the same.

1.89 The observed expenditure of a rational consumer in the selected base period may be assumed to be the minimum expenditure needed to achieve the level of utility enjoyed in that period. In order to calculate a COLI based on that period, it is necessary to know what would be the minimum expenditure needed to attain precisely the same level of utility if the prices prevailing were those of the second period, other things remaining equal. The quantities purchased under these assumed conditions are likely to be *hypothetical*. They will not be the quantities actually consumed in the second period if other factors, including the resources available to the consumer, have changed.

1.90 The quantities required for the calculation of the COLI in at least one of the periods are not likely to be observable in practice. The COLI is not an operational index that can be calculated directly. The challenge is therefore to see whether it is possible to find methods of estimating a COLI indirectly or at least to find upper and lower bounds for the index. There is also considerable interest in establishing the relationship between a COLI and Lowe indices, including Laspeyres and Paasche, that can be calculated.

Upper and lower bounds on a cost of living index

1.91 It follows from the definition of a Laspeyres index that, if the consumer's income were to change by the same proportion as the change in the Laspeyres index, the consumer must have the possibility of purchasing the same basket of products as in the base period. The consumer cannot be worse off. However, if *relative* prices have changed, a utility-maximizing consumer would not continue to purchase the same quantities as before. The consumer would be able to achieve a *higher level* of utility by substituting, at least marginally, products that have become relatively cheaper for those that have become dearer. As a COLI measures the change in the minimum expenditures needed to maintain a constant level of utility, the COLI based on the first period will increase by less than the Laspeyres index.

1.92 By a similar line of reasoning, it follows that when relative prices change, the COLI based on the second period must increase by more than the Paasche index. As explained in more detail in Chapter 17, the Laspeyres index provides an upper bound to the COLI based on the first period and the Paasche a lower bound to the COLI based on the second period. It should be noted that there are two different COLIs involved here: one based on the first period and the other based on the second period. In general, however, the two COLIs are unlikely to differ much.

1.93 Suppose that the theoretical target index is a COLI, but that, for practical reasons, the CPI is actually calculated as a Lowe index in which the quantities refer to some period b that precedes the price reference period 0. One important conclusion to be drawn from this preliminary analysis is that as the Lowe may be expected to exceed the Laspeyres, assuming long-term price trends and substitution, while the Laspeyres may in turn be expected to exceed the COLI, the widely used Lowe index may be expected to have an upward bias. This point has had a profound influence on attitudes towards CPIs in some countries. The bias results from the fact that, by definition, fixed basket indices, including Laspeyres, do not permit any substitution between products in response to changes in relative prices. It is therefore usually described as "substitution bias". A Paasche index would be expected to have a downward substitution bias.

Some special cases

1.94 The next step is to establish whether there are special conditions under which it may be possible to measure a COLI exactly. In Chapter 17 it is shown that if the consumer's preferences are homothetic – that is, each indifference curve has the same shape, each being a uniform enlargement, or contraction, of each other – then the COLI is independent of the utility level on which it is based. The Laspeyres and Paasche indices provide upper and lower bounds to the *same* COLI.

1.95 One interesting special case occurs when the preferences can be represented by the so-called "Cobb–Douglas" function in which the cross-elasticities of demand between the various products are all unity. Consumers adjust the relative quantities they consume inversely in proportion to the changes in relative prices so that expenditure shares remain constant. With Cobb–Douglas preferences, the geometric Laspeyres provides an exact measure of the COLI. As the expenditure shares remain constant over time, all three *geometric* indices – Young, Laspeyres and Paasche – coincide with each other and with the COLI. Of course, the arithmetic versions of these indices do not coincide in these circumstances, because the baskets in periods b, 0 and t are all different as substitutions take place in response to changes in relative prices.

1.96 One of the more famous results in index number theory is that if the preferences can be represented by a homogeneous quadratic utility function, the Fisher index provides an exact measure of the COLI (see Chapter 17). Even though consumers' preferences are unlikely to conform exactly with this particular functional form, this result does suggest that, in general, the Fisher index is likely to provide a close approximation to the underlying unknown COLI and certainly a much closer approximation than either the arithmetic Laspeyres or Paasche indices.

Estimating COLIs by superlative indices

1.97 The intuition – that the Fisher index approximates the COLI – is corroborated by the following line of reasoning. Diewert (1976) noted that a homogeneous quadratic is a flexible functional form that can provide a second-order approximation to other twice-differentiable functions around the same point. He then described an index number formula as *superlative* when it is exactly equal to the COLI based on a certain functional form *and* when that functional form is flexible, e.g., a homogeneous quadratic. The derivation of these results, and further explanation, is given in detail in Chapter 17. In contrast to the COLI based on the true but unknown utility function, a superlative index is an actual index number that can be calculated. The practical significance of these results is that they provide a theoretical justification for expecting a superlative index to provide a fairly close approximation to the underlying COLI in a wide range of circumstances.

1.98 *Superlative indices as symmetric indices.* The Fisher is by no means the only example of a superlative index. In fact, there is a whole family of superlative indices. It is shown in Chapter 17 that any quadratic mean of order r is a superlative index for each value of $r \neq 0$. A quadratic mean of order r price index P^r is defined as follows:

$$P^r \equiv \frac{\sqrt[r]{\sum_{i=1}^{n} s_i^0 \left(\frac{p_i^t}{p_i^0}\right)^{r/2}}}{\sqrt[r]{\sum_{i=1}^{n} s_i^t \left(\frac{p_i^0}{p_i^t}\right)^{r/2}}} \tag{1.15}$$

where s_i^0 and s_i^t are defined as in equations (1.2) and (1.3) above.

1.99 The symmetry of the numerator and denominator of equation (1.15) should be noted. A distinctive feature of equation (1.15) is that it treats the price changes and expenditure shares in both periods symmetrically, whatever value is assigned to the parameter r. Three special cases are of interest:

- when $r = 2$, equation (1.1) reduces to the Fisher price index;
- when $r = 1$ it is equivalent to the Walsh price index;
- in the limit as $r \rightarrow 0$, it equals the Törnqvist index.

These indices were introduced earlier as examples of indices that treat the information available in both periods *symmetrically*. Each was originally proposed long before the concept of a superlative index was developed.

1.100 *Choice of superlative index*. Chapter 17 addresses the question of which superlative formula to choose in practice. As each may be expected to approximate to the same underlying COLI, it may be inferred that they ought also to approximate to each other. The fact that they are all symmetric indices reinforces this conclusion. These conjectures tend to be borne out in practice by numerical calculations. So long as the parameter r does not lie far outside the range 0 to 2, superlative indices tend to be very close to each other. In principle, however, there is no limit on r and it has recently been shown that as r becomes larger, the formula tends to assign increasing weight to the extreme price relatives and the resulting superlative indices may diverge significantly from each other. Only when the absolute value of r is small, as in the case of the three commonly used superlative indices (Fisher, Walsh and Törnqvist), is the choice of superlative index unimportant.

1.101 Both the Fisher and the Walsh indices date back nearly a century. The Fisher index owes its popularity to the axiomatic, or test, approach, which Fisher himself was instrumental in developing. As already noted, it dominates other indices using the first axiomatic approach, while the Törnqvist dominates using the second axiomatic approach outlined above. The fact that the Fisher and the Törnqvist are both superlative indices whose use can be justified on economic grounds suggests that, from a theoretical point of view, it may not be possible to improve on them for CPI purposes.

Representativity bias

1.102 The fact that the Walsh index is a Lowe index that is also superlative suggests that the bias in other Lowe indices depends on the extent to which their quantities deviate from those in the Walsh basket. This can be viewed from another angle.

1.103 As the quantities in the Walsh basket are *geometric* averages of the quantities in the two periods, equal importance is assigned to the *relative*, as distinct from the *absolute*, quantities in both periods. The Walsh basket may therefore be regarded as being the basket that is most representative of *both* periods. If equal importance is attached to the consumption patterns in the two periods, the optimal basket for a Lowe index ought to be the most representative basket. The Walsh index then becomes the conceptually preferred target index for a Lowe index.

1.104 Suppose that period b, for which the quantities are actually used in the Lowe index, lies midway between 0 and t. In this case, assuming fairly smooth trends in the relative quantities, the actual basket in period b is likely to approximate to the most representative basket. Conversely, the further away that period b is from the midpoint between 0 and t, the more the relative quantities of b are likely to diverge from those in the most representative basket. In this case, the Lowe index between periods 0 and t that uses period b quantities is likely to exceed the Lowe index that uses the most representative quantities by an amount that becomes progressively larger the further back in time period b is positioned. The excess constitutes "bias" if the latter index is the target index. The bias can be attributed to the fact that the period b quantities tend to become increasingly unrepresentative of a comparison between 0 and t the further back period b is positioned. The underlying economic factors responsible are, of course, exactly the same as those that give rise to bias when the target index is the COLI. Thus, certain kinds of indices can be regarded as biased without invoking the concept of a COLI. Conversely, the same kinds of indices tend to emerge as being preferred, whether or not the objective is to estimate a cost of living bias.

1.105 If interest is focused on short-term price movements, the target index is an index between consecutive time periods t and $t+1$. In this case, the most representative basket has to move forward one period as the index moves forward. Choosing the most representative basket implies chaining. Similarly, chaining is also implied if the target index is a COLI between t and $t+1$. In practice, the universe of products is continually changing as well. As the most representative basket moves forward, it is possible to update the set of products covered, as well as take account of changes in the relative quantities of products that were covered previously.

Data requirements and calculation issues

1.106 As superlative indices require price and expenditure data for both periods, and as expenditure data are usually not available for the current period, it is not feasible to calculate a superlative CPI, at least at the time that a CPI is first published. In practice, CPIs tend to be Lowe indices with fixed quantities or annually updated chain Lowe indices. In the course of time, however, the requisite expenditure data may become available, enabling a superlative CPI to be calculated subsequently. Users will find it helpful for superlative CPIs to be published retrospectively as they make it possible to evaluate the properties and behaviour of the official index. Superlative CPIs can be treated as supplementary indices that complement, rather than replace, the original indices, if the policy is not to revise the official index.

1.107 Chapter 17 notes that, in practice, CPIs are usually calculated in stages (see also Chapters 9 and 20) and addresses the question of whether indices calculated this way are consistent in aggregation: that is, have the same values whether calculated in a single operation or in two stages. The Laspeyres index is shown to be exactly consistent, but the superlative indices are not. The widely used Fisher and Törnqvist indices are nevertheless shown to be approximately consistent.

Allowing for substitution

1.108 Chapter 17 examines one further index proposed recently, the Lloyd–Moulton index, P_{LM}, defined as follows:

$$P_{LM} \equiv \left\{ \sum_{i=1}^{n} s_i^0 \left(\frac{p_i^t}{p_i^0} \right)^{1-\sigma} \right\}^{\frac{1}{1-\sigma}} \qquad \sigma \neq 1 \qquad (1.16)$$

The parameter σ, which must be non-negative, is the elasticity of substitution between the products covered. It reflects the extent to which, on average, the various products are believed to be substitutes for each other. The advantage of this index is that it may be expected to be free of substitution bias to a reasonable degree of approximation, while requiring no more data than a Lowe or Laspeyres index. It is therefore a practical possibility for CPI calculation, even for the most recent periods, although it is likely to be difficult to obtain a satisfactory, acceptable estimate of the numerical value of the elasticity of substitution, the parameter used in the formula.

Aggregation issues

1.109 It has been assumed up to this point that the COLI is based on the preferences of a single representative consumer. Chapter 18 examines the extent to which the various conclusions reached above remain valid for CPIs that are actually compiled for groups of households. The general conclusion is that essentially the same relationships hold at an aggregate level, although some additional issues arise which may require additional assumptions.

1.110 One issue is how to weight individual households. Aggregate indices that weight households by their expenditures are called "plutocratic", while those that assign the same weight to each household are called "democratic". Another question is whether, at any one point of time, there is a single set of prices or whether different households face different prices. In general, when defining the aggregate indices it is not necessary to assume that all households are confronted by the same set of prices, although the analysis is naturally simplified if there is only a single set.

1.111 A plutocratic aggregate COLI assumes that each individual household minimizes the cost of attaining a given level of utility when confronted by two different sets of prices, the aggregate COLI being defined as the ratio of the aggregate minimum costs over all households. As in the case of a single household, it is recognized that the aggregate COLI that is appropriate for CPI purposes must be *conditional* on the state of a particular set of environmental variables, typically those in one or other of the periods compared. The environment must be understood in a broad sense to refer not only to the physical environment but also to the social and political environment.

1.112 Like the index for a single representative consumer, an aggregate COLI cannot be calculated directly, but it may be possible to calculate aggregate Laspeyres and Paasche indices that bound their respective COLIs from above or below. If there is only one single set of national prices, the aggregate plutocratic Laspeyres index reduces to an ordinary aggregate Laspeyres index. As the aggregate plutocratic Laspeyres and Paasche can, in principle, be calculated, so can the aggregate plutocratic Fisher index. It is argued in Chapter 18 that this should normally provide a good approximation to the aggregate plutocratic COLI.

1.113 Chapter 18 finally concludes that, in principle, both democratic and plutocratic Laspeyres, Paasche and Fisher indices could be constructed by a statistical agency, provided that information on household-specific price relatives and expenditures is available for both periods. If expenditure information is available only for the first period, then only the Laspeyres democratic and plutocratic indices can be constructed. The data requirements are rather formidable, however. The required data are unlikely to be available for *individual* households in practice and, even if they were to be, they could be subject to large errors.

Illustrative numerical data

1.114 Chapter 19 presents some numerical examples using an artificial data set. The purpose is not to illustrate the methods of calculation as such, but rather to demonstrate how different index number formulae can yield very different numerical results. Hypothetical but economically plausible prices, quantities and expenditures are given for six commodities over five periods of time. In general, differences between the different formulae tend to increase with the variance of the price relatives. They also depend on the extent to which the prices follow smooth trends or fluctuate.

1.115 The numerical results are striking. For example, the Laspeyres index over the five periods registers an increase of 44 per cent while the Paasche falls by 20 per cent. The two commonly used superlative indices, Törnqvist and Fisher, register increases of 25 per cent and 19 per cent respectively, an index number spread of only 6 points compared with the 64-point gap between the Laspeyres and Paasche. When the indices are chained, the chain Laspeyres and Paasche indices register increases of 33 per cent and 12 per cent respectively, reducing the gap between the two indices from 64 to 21 points. The chained Törnqvist and Fisher indices register increases of 22.26 per cent and 22.24 per cent respectively, being virtually identical numerically. These results show that the choice of index formula and method does matter.

Seasonal products

1.116 As explained in Chapter 22, the existence of seasonal products poses some intractable problems and serious challenges for CPI compilers and users. Seasonal products are products that are either:

– not available during certain seasons of the year; or
– are available throughout, but their prices or quantities are subject to regular fluctuations that are synchronized with the season or time of the year.

There are two main sources of seasonal fluctuations: the climate and custom. Month-to-month movements in a CPI may sometimes be so dominated by seasonal influences that it is difficult to discern the underlying trends in prices. Conventional seasonal adjustment programmes may be applied, but these may not always be satisfactory. The problem is not confined to interpreting movements in the CPI, as seasonality creates serious problems for the compilation of a CPI when some of the products in the basket regularly disappear and reappear, thereby breaking the continuity of the price series from which the CPI is built up. There is no panacea for seasonality. A consensus on what is best practice in this area has not yet been formed. Chapter 22 examines a number of different ways in which the problems may be tackled using an artificial data set to illustrate the consequences of using different methods.

1.117 One possibility is to exclude seasonal products from the index, but this may be an unacceptable reduction in the scope of the index, as seasonal products can account for a significant proportion of total household consumption. Assuming seasonal products are retained, one solution is to switch the focus from month-to-month movements in the index to changes between the same month in successive years. In some countries, it is common for the media and other users, such as central banks, to focus on the annual rate of inflation between the most recent month and the same month in the previous year. This year-on-year figure is much easier to interpret than month-to-month changes, which can be somewhat volatile, even in the absence of seasonal fluctuations.

1.118 This approach is extended in Chapter 22 to the concept of a rolling year-on-year index that compares the prices for the most recent 12 months with the corresponding months in the price reference year. The resulting *rolling year indices* can be regarded as seasonally adjusted price indices. They are shown to work well using the artificial data set. Such an index can be regarded as a measure of inflation for a year that is centred around a month that is six months earlier than the last month in the rolling index. For some purposes, this time lag may be disadvantageous, but in Chapter 22 it is shown that under certain conditions the current month year-on-year monthly index, together with the previous month's year-on-year monthly index, can successfully predict the rolling year index that is centred on the current month. Of course, rolling year indices and similar analytic constructs are not intended to replace the monthly or quarterly CPI but to provide supplementary information that can be extremely useful to users. They can be published alongside the official CPI.

1.119 Various methods of dealing with the breaks in price series caused by the disappearance and reappearance of seasonal products are examined in Chapter 22. However, this remains an area in which more research needs to be done.

Elementary price indices

1.120 As explained in Chapters 9 and 20, the calculation of a CPI proceeds in stages. In the first stage, *elementary price indices* are estimated for the *elementary expenditure aggregates* of a CPI. In the second stage, these elementary indices are aggregated, or averaged, to obtain higher-level indices using the elementary expenditure 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. As explained in Chapter 6, statistical offices usually select a set of representative products within each aggregate and then collect samples of their prices from a number of different outlets. The elementary aggregates serve as strata for sampling purposes.

1.121 The prices collected at the first stage are typically not prices observed in actual transactions between different economic units, but the prices at which the products are offered for sale in retail outlets of one kind or another. In principle, however, a CPI measures changes in the prices paid by households. These prices may actually vary during the course of a month, which is typically the time period to which the CPI relates. In principle, therefore, the first step should be to average the prices at which some product is sold during the period, bearing in mind that the price may vary even for the same product sold in the same outlet. In general, this is not a practical possibility. However, when the outlet is an electronic point of sale at which all the individual products are "scanned" as they are sold, the values of the transactions are actually recorded, thereby making it possible to calculate an average price instead of simply recording the offer price at a single point of time. Some use of scanner data is already made for CPI purposes and it may be expected to increase over the course of time.

1.122 Once the prices are collected for the representative products in a sample of outlets, the question arises of what is the most appropriate formula to use to estimate an elementary price index. This topic is considered in Chapter 20. It was comparatively neglected until a number of papers in the 1990s provided much clearer insights into the properties of elementary indices and their relative strengths and weaknesses. The quality of a CPI depends heavily on the quality of the elementary indices which are the building blocks from which CPIs are constructed.

1.123 Prices are collected for the same product in the same outlet over a succession of time periods. An elementary price index is therefore typically calculated from two sets of matched price observations. Here it is assumed that there are no missing observations and no changes in the quality of the products sampled so that the two sets of prices are perfectly matched. The treatment of new and disappearing products, and of quality change, is a separate and complex issue in its own right. It is outlined below, and discussed in detail in Chapters 7, 8 and 21.

Weights within elementary aggregates

1.124 In most cases, the price indices for elementary aggregates are calculated without the use of explicit expenditure weights. Whenever possible, however, weights should be used that reflect the relative importance of the sampled items, even if the weights are only approximate. In many cases, the elementary aggregate is simply the lowest level at which any reliable weighting information is available. In this case, the elementary index has to be calculated without the use of weights. Even in this case, however, it should be noted that when the items are selected with probabilities proportional to the size of some relevant variable such as sales, for example, weights are implicitly introduced by the sampling selection procedure.

1.125 For certain elementary aggregates, information about sales of particular items, market shares and regional weights may be used as explicit weights within an elementary aggregate. Weights within elementary aggregates may be updated independently, and possibly more often than the elementary aggregates themselves (which serve as weights for the higher-level indices).

1.126 For example, assume that the number of suppliers of a certain product, such as petrol, is limited. The market shares of the suppliers may be known from business survey statistics and can be used as weights in the calculation of an elementary aggregate price index for petrol. As another example, prices for water may be collected from a number of local water supply services where the population in each local region is known. The relative size of the population in each region may then be used as a proxy for the relative consumption expenditures to weight the price in each region to obtain the elementary aggregate price index for water.

Interrelationships between different elementary index formulae

1.127 Useful insights into the properties of various formulae that have been used, or considered, for elementary price indices may be gained by examining the mathematical interrelationships between them. Chapter 20 provides a detailed analysis of such relationships. As it is assumed that there are no explicit weights available, the various formulae considered all make use of unweighted averages: that is, *simple* averages in which the various items are *equally* weighted. There are two basic options for an elementary index:

- some kind of simple average of the price ratios or relatives;
- the ratio of some kind of simple average of the prices in the two periods.

In the case of a geometric average, the two methods coincide, as the geometric average of the price ratios or relatives is identical to the ratio of the geometric average prices.

1.128 Using the first of the above options, three possible elementary price indices are:

- a simple arithmetic average of the price relatives, known as the *Carli* index, or P_C; the Carli is the unweighted version of the Young index;
- a simple geometric average of the price relatives, known as the *Jevons* index, or P_J; the Jevons is the unweighted version of the geometric Young index;
- a simple harmonic average of the price relatives, or P_H.

As noted earlier, for any set of positive numbers the arithmetic average is greater than, or equal to, the geometric average, which in turn is greater than, or equal to, the harmonic average, the equalities holding only when the numbers are all equal. It follows that $P_C \geq P_J \geq P_H$.

1.129 It is shown in Chapter 20 that the gaps between the three indices widen as the variance of the price relatives increases. The choice of formula becomes more important the greater the diversity of the price movements. P_J can be expected to lie approximately halfway between P_C and P_H.

1.130 Using the second of the options, three possible indices are:

- the ratio of the simple arithmetic average prices, known as the *Dutot* index, or P_D;
- the ratio of the simple geometric averages, again the Jevons index, or P_J;
- the ratio of the simple harmonic averages, or P_H.

The ranking of *ratios* of different kinds of average are not predictable. For example, the Dutot, P_D, could be greater or less than the Jevons, P_J.

1.131 The Dutot can also be expressed as a weighted average of the price relatives in which the prices of period 0 serve as the weights:

$$P_D \equiv \frac{\sum_{i=1}^{n} p_i^t \big/ n}{\sum_{i=1}^{n} p_i^0 \big/ n} = \frac{\sum_{i=1}^{n} p_i^0 \left(\frac{p_i^t}{p_i^0}\right)}{\sum_{i=1}^{n} p_i^0} \qquad (1.17)$$

As compared with the Carli, which is a simple average of the price relatives, the Dutot gives more weight to the price relatives for the products with high prices in period 0. It is nevertheless difficult to provide an economic rationale for this kind of weighting. Prices are not expenditures. If the products are homogeneous, very few quantities are likely to be purchased at high prices if the same products can be purchased at low prices. If the products are heterogeneous, the Dutot should not be used anyway, as the quantities are not commensurate and not additive.

1.132 While it is useful to establish the interrelationships between the various indices, they do not actually help decide which index to choose. However, as the differences between the various formulae tend to increase with the dispersion of the price relatives, it is clearly desirable to define the elementary aggregates in such a way as to try to minimize the variation in the price movements within each aggregate. The less variation there is, the less difference the choice of index formula makes. As the elementary aggregates also serve as strata for sampling purposes, minimizing the variance in the price relatives within the strata will also reduce the sampling error.

Axiomatic approach to elementary indices

1.133 One way to decide between the various elementary indices is to exploit the axiomatic approach outlined earlier. A number of tests are applied to the elementary indices in Chapter 20.

1.134 The Jevons index, P_J, satisfies all the selected tests. It dominates the other indices in the way that the Fisher tends to dominate other indices at an aggregative level. The Dutot index, P_D, fails only one, the commensurability test. This failure is critical, however. It reflects the fundamental point made earlier that when the quantities are not additive from an economic viewpoint, the prices are also not additive and hence cannot be meaningfully averaged. However, P_D performs well when the sampled products are homogeneous. The key issue for the Dutot is therefore how heterogeneous are the products within the elementary aggregate. If the products are not sufficiently homogeneous for their quantities to be additive, the Dutot should not be used.

1.135 Although the Carli index, P_C, has been widely used in practice, the axiomatic approach shows it to have some undesirable properties. In particular, as the unweighted version of the Young index, it fails the time reversal and transitivity tests. This is a serious disadvantage, especially as elementary indices are often monthly chain indices. A consensus has emerged that the Carli may be unsuitable because it is liable to have a significant upward bias. This is illustrated by numerical example in Chapter 9. Its use is not sanctioned for the Harmonized Indices of Consumer Prices used within the European Union. Conversely, the harmonic average of the price relatives, P_H, is liable to have an equally significant downward bias; anyway, it does not seem to be used in practice.

1.136 Based on the axiomatic approach, the Jevons emerges as the preferred index, but its use may not be appropriate in all circumstances. If one observation is zero, the geometric mean is zero. The Jevons is sensitive to extreme falls in prices; it may be necessary to impose upper and lower bounds on the individual price relatives when using the Jevons.

Economic approach to elementary indices

1.137 The economic approach to elementary indices is explained in Chapter 20. The sampled products for which prices are collected are treated as if they constituted a basket of goods and services purchased by rational utility-maximizing consumers. The objective is then to estimate a conditional cost of living index covering the set of products in question.

1.138 It should be noted, however, that the differences between the prices of the sampled products do not necessarily mean that the products are qualitatively different. If markets were perfect, relative prices should reflect relative costs of production and relative utilities. In fact, price differences may occur simply because of market imperfections. For example, exactly the same products may be bought and sold at different prices in different outlets simply because consumers lack information about the prices charged in other outlets. Producers may also practise price discrimination, charging different customers different prices for exactly the same products. Price discrimination is widespread in many service industries. When the price differences are a result of market imperfections, consumers cannot be expected to react to changes in the relative prices of products in the same way as they would if they were well informed and had free choice.

1.139 In any case, assuming there is no information about quantities or expenditures within an elementary aggregate, it is not possible to calculate any kind of superlative index. So the conditional cost of living index at the level of an elementary aggregate can be estimated only on the assumption that certain special conditions apply.

1.140 There are two special cases of some interest. The first case is where the underlying preferences are so-called Leontief preferences. With these preferences *relative* quantities remain fixed whatever the relative prices. No substitutions are made in response to changes in relative prices. The cross-elasticities of demand are zero. With Leontief preferences, a Laspeyres index provides an exact measure of the cost of living index. In this case, the Carli calculated for a random sample would provide an estimate of the cost of living index provided that the items were selected with probabilities proportional to the population expenditure shares. It might appear that if the items were selected with probabilities proportional to the population quantity shares, the sample Dutot would provide an estimate of the population Laspeyres. However, assuming that the basket for the Laspeyres index contains a number of heterogeneous products whose quantities are not additive, the quantity shares, and hence the probabilities, are undefined.

1.141 The second case is one already considered above, namely when the preferences can be represented by a Cobb–Douglas function. As already explained, with these preferences, the geometric Laspeyres would provide an exact measure of the cost of living index. In this case, the Carli calculated for a random sample would provide an unbiased estimate of the cost of living index, provided that the items were selected with probabilities proportional to the population expenditure shares.

1.142 On the economic approach, the choice between the sample Jevons and the sample Carli rests on which is likely to approximate the more closely to the underlying COLI: in other words, on whether the demand cross-elasticities are likely to be closer to unity or zero, on average. In practice, the cross-elasticities could take on any value ranging up to plus infinity for an elementary aggregate in which the sampled products were strictly homogeneous, i.e., perfect substitutes. It should be noted that in the limiting case in which the sampled products are homogeneous, there is only a single kind of product and therefore no index number problem: the price index is given by the ratio of the unit values in the two periods. It may be conjectured that, on average, the cross-elasticities are likely to be closer to unity than zero for most elementary aggregates so that, in general, the Jevons index is likely to provide a closer approximation to the cost of living index than the Carli. In this case, the Carli must be viewed as having an upward bias.

1.143 It is worth noting that the use of the Jevons index does not imply, or assume, that expenditure shares remain constant. Obviously, a geometric average of the price relatives can be calculated whatever changes do or do not occur in the expenditure shares, in practice. What the economic approach shows is that *if* the expenditure shares remain constant (or roughly constant), *then* the Jevons can be expected to provide a good estimate of the underlying cost of living index. The insight provided by the economic approach is that the Jevons is likely to provide a closer approximation to the cost of living index than the Carli because a significant amount of substitution is more likely than no substitution, especially as elementary aggregates should be deliberately constructed in such a way as to group together similar items that are close substitutes for each other.

1.144 An alternative to the Jevons, P_J, would be a geometric average of P_C and P_H, an index labelled P_{CSWD} in Chapter 20. This could be justified on grounds of treating the data in both periods symmetrically without invoking any particular assumption about the form of the underlying preferences. It is also shown in Chapter 20 that the geometric average of P_C and P_H is likely to be very close to P_J, so that the latter may be preferred on the grounds that it is a simpler concept and easier to compile.

1.145 It may be concluded that, based on the economic approach, as well as the axiomatic approach, the Jevons emerges as the preferred index in general, although there may be cases in which little or no substitution takes place within the elementary aggregate and the Carli might be preferred. The index compiler must make a judgement on the basis of the nature of the products actually included in the elementary aggregate.

1.146 The above discussion has also thrown light on some of the sampling properties of the elementary indices. If the products in the sample are selected with probabilities proportional to expenditures in the price reference period:

- the sample (unweighted) Carli index provides an unbiased estimate of the population Laspeyres;
- the sample (unweighted) Jevons index provides an unbiased estimate of the population geometric Laspeyres.

These results hold irrespective of what the underlying cost of living index may be.

Concepts, scope and classifications

1.147 The purpose of Chapter 3 of the manual is to define and clarify a number of basic concepts underlying a CPI and to explain the scope of the index: that is, the set of goods and services and the set of households that the index is intended to cover, in principle. Chapter 3 also examines the structure of the classification of consumer goods and services used.

1.148 While the general purpose of a CPI is to measure changes in the prices of *consumption* goods and services, there are a number of concepts that need to be defined precisely before an operational definition of a CPI can be arrived at. The concept of consumption is an imprecise one that can be interpreted in several different ways, each of which may lead to a different CPI. It is also necessary to decide whether the index is meant to cover all consumers, i.e., all households, or just a particular group of households. The scope of a CPI is inevitably influenced by what is intended, or believed, to be the main use of the index. Compilers also need to remember that the index may be used as proxy for a general price index and used for purposes other than those for which it is intended.

1.149 The word "consumer" can be used to refer both to a type of economic unit and to a type of product. To avoid confusion here, the term *consumption* good or service will be used where necessary, rather than *consumer* good or service. A consumption good or service provides utility to its user. It may be defined as *a good or service that members of households use, directly or indirectly, to satisfy their own personal needs and wants*. "Utility" should be interpreted in a broad sense. It is simply the generic, technical term preferred by economists for the benefit or welfare that individuals or households derive from the use of a consumer good or service.

1.150 A CPI is generally understood to be a price index that measures changes in the prices of consumption goods and services acquired and used by households. In principle, more broadly based price indices can be defined whose scope extends beyond consumption goods and services to include the prices of physical assets such as land or dwellings. Such indices may be useful as broad measures of inflation as perceived by households, but most CPIs are confined to consumption goods and services. These may include the prices of the flows of services provided by assets such as dwellings, even though the assets themselves may be excluded. In any case, the prices of financial assets such as bonds, shares or other marketable securities purchased by households are generally regarded as being outside the scope of a CPI.

Acquisitions and uses

1.151 The times at which households acquire and use consumption goods or services are generally not the same. Goods are typically acquired at one point in time and used at some other point in time, or even used repeatedly over an extended period of time. The time of acquisition of a *good* is the moment at which the legal or effective economic ownership of the good passes to the consumer. In a market situation, this is the point at which the purchaser incurs a liability to pay. A *service* is acquired at the time that the producer provides it, no change of ownership being involved. The time at which acquisitions are recorded, and their prices, should also be consistent with the way in which the same transactions are recorded in the expenditure data used for weighting purposes.

1.152 The time at which payment is made may be determined mainly by institutional arrangements and administrative convenience. When payments are not made in cash, there may be a significant lapse of time before the consumer's bank account is debited for a purchase paid for by cheque, by credit card or similar arrangements. The time at which these debits are eventually made is irrelevant for the recording of the acquisitions and the prices. On the other hand, when the acquisition of a good or service is financed by the creation of a new financial asset at the time of acquisition, such as a loan to the purchaser, two economically separate transactions are involved, the purchase/sale of the good or service and the creation of the asset. The price to be recorded is the price payable at the time of acquisition, however the purchase is financed. Of course, the provision of finance may affect the price payable. The subsequent repayments of any debt incurred by the purchaser and the associated interest payments are financial transactions that are quite distinct from the purchase of the good or service whose price has to be recorded. The explicit or implicit interest payments payable on the amount depend on the capital market, the nature of the loan, its duration, the creditworthiness of the purchaser, and so on. These points are explained in more detail in Chapter 3.

1.153 The distinction between the *acquisition* and the *use* of a consumer good or service outlined above has led to two different concepts of a CPI being proposed:

- A CPI may be intended to measure the average change between two time periods in the prices of the consumer goods and services acquired by households.
- Alternatively, a CPI may be intended to measure the average change between two time periods in the prices of the consumer goods and services used by households to satisfy their needs and wants.

The distinction between time of acquisition and time of use is particularly important for durable goods and certain kinds of services.

1.154 *Durable and non-durable goods*. A "non-durable" good might be better described as a *single use* good. For example, food or drink are used once only to satisfy hunger or thirst. Many so-called non-durable consumer goods are in fact extremely durable physically. Households may hold substantial stocks of non-durables, such as many foodstuffs and fuel, for long periods of time before they are used.

1.155 The distinguishing feature of a durable consumption good is that it is durable under use. Consumer durables can be used repeatedly or continuously to satisfy the needs or wants of consumers over a long period of time, possibly many years: for example, furniture or vehicles. For this reason, a durable is often described as providing a flow of services to the consumer over the period it is used (see also Box 14.3 of Chapter 14). There is a close parallel between the definitions of consumer durables and fixed assets. Fixed assets are defined in national accounts as goods that are used repeatedly or continuously over long periods of time in processes of production: for example, buildings or other structures, machinery and equipment.

1.156 A list of the different kinds of consumer durables distinguished in the Classification of Individual Consumption according to Purpose (COICOP) is given in Chapter 3. Of course, some durables last much longer than others, the less durable ones being described as "semi-durables" in COICOP: for example, clothing. It should be noted that dwellings are classified as fixed assets, not durable consumption goods, and are therefore not included in COICOP. Dwellings are used to *produce* housing services. These services are consumed by tenants or owner-occupiers, as the case may be, and are therefore included in COICOP.

1.157 Many services are durable and are also not fully consumed, or used up, at the time they are acquired. Some services bring about long-lasting improvements from which the consumers derive enduring benefits. The condition and quality of life of persons receiving medical treatments such as hip replacements or cataract surgery, for example, are substantially and permanently improved. Similarly, consumers of educational services can derive lifetime benefits from them. Expenditures on education and health also share with durable goods the characteristic that they are also often so costly that they have to be financed by borrowing or by running down other assets.

1.158 Expenditures on durable goods and durable services are liable to fluctuate, whereas using up such goods and services is likely to be a fairly steady process. However, the using up cannot be directly observed and valued. It can only be estimated by making assumptions about the timing and duration of the flows of benefits. Partly because of the conceptual and practical difficulties involved in measuring uses, statistical offices tend to adopt the acquisitions approach to consumer durables in both their national accounts and CPIs.

1.159 *A consumer price index based on the acquisitions approach*. Households may acquire goods and services for purposes of consumption in four main ways. They may:

- purchase them in monetary transactions;
- produce them themselves for their own consumption;
- receive them as payments in kind in barter transactions, particularly as remuneration in kind for work done;
- receive them as free gifts, or transfers, from other economic units.

1.160 The broadest possible scope for goods and services based on the acquisitions approach would be one covering all four categories, irrespective of who bears the costs. It would therefore include all *social transfers in kind* in the form of education, health, housing and other goods and services provided free of charge, or at nominal prices, to individual households by governments or non-profit institutions (NPIs). Total acquisitions are equivalent to the total actual individual consumption of (non-institutional) households, as defined in the SNA (see Chapter 14). *Collective* services provided by governments to the community as whole, such as public administration and defence, are not included and are outside the scope of a CPI.

1.161 From the point of view of the government or NPI that provides and pays for them, social transfers are valued either by the market prices paid for them or by the costs of producing them. From the point of view of the receiving households they have zero or nominal prices. For CPI purposes, the appropriate price is that paid by the household. The price paid by the government belongs in a price index for government expenditures. When households incur zero expenditures, the services provided free carry zero weight in a CPI. However, when governments and NPIs introduce charges for goods or services that were previously provided free, the increase from a zero to a positive price could be captured by a CPI, as explained in Chapter 3.

1.162 *Expenditures versus acquisitions*. Expenditures need to be distinguished from acquisitions. Expenditures are incurred by the economic units that bear the costs. Households do not incur expenditures on social transfers in kind, so the scope of households' expenditures is generally narrower than the scope of their acquisitions. Moreover, not all expenditures are monetary. A *monetary expenditure* occurs when a household pays in cash, by cheque or credit card, or otherwise incurs a financial liability to pay. Only monetary expenditures generate monetary prices that can be observed and recorded for CPI purposes.

1.163 *Non-monetary expenditures* occur when households pay, but in other ways than cash. There are three important categories of non-monetary expenditures:

- In barter transactions, households exchange consumption goods and services among themselves. As the values of the goods and services surrendered as payments constitute negative expenditures, the expenditures should cancel out so that barter transactions between households carry zero weight on aggregate. They can be ignored in practice for CPI purposes.
- When employees are remunerated in kind, they purchase the goods or services, but pay with their labour, not cash. Monetary values can be imputed for the expenditures implicitly incurred by the households.
- Similarly, when households produce goods and services for themselves, they incur the costs, some of which may be monetary in the form of purchased inputs. The monetary values of the implicit expenditures on the outputs produced can be imputed on the basis of the corresponding market prices. If such imputed prices were to be included in the CPI, the prices of the inputs would have to be excluded to avoid double counting.

1.164 *A hierarchy of consumption aggregates*. A hierarchy of possible consumption aggregates may be envisaged, as explained in Chapter 14:

– total acquisitions of goods and services by households;
– *less* social transfers in kind = households' total expenditures;
– *less* non-monetary expenditures = households' monetary expenditures.

The choice of consumption aggregate is a policy matter. For example, if the main reason for compiling a CPI is to measure inflation, the scope of the index might be restricted to household monetary expenditures on consumption, inflation being essentially a monetary phenomenon. Prices cannot be collected for the consumer goods and services involved in non-monetary expenditures, although they can be estimated on the basis of the prices observed in corresponding monetary transactions. The European Union's Harmonized Indices of Consumer Prices, which are specifically intended to measure inflation within the EU, are confined to monetary expenditures.

Unconditional and conditional cost of living indices

1.165 Cost of living indices, or COLIs, are explained in Chapters 15 and 17. As also noted in Chapter 3, the scope of a COLI depends on whether it is conditional or unconditional. The welfare of a household depends not only on the utility derived from the goods and services it consumes, but on the social, political and physical environment in which the household lives. An *unconditional* cost of living index measures the change in the minimum cost of maintaining a given level of welfare in response to changes in any of the factors that affect welfare, whereas a *conditional* cost of living index measures the change in the minimum cost of maintaining a given level of utility or welfare resulting from changes in consumer prices, holding the environmental factors constant.

1.166 An unconditional COLI may be a more comprehensive *cost of living* index than a conditional COLI, but it is not a more comprehensive *price* index. An unconditional index does not include any more price information than a conditional index and it does not give more insight into the impact of price changes on welfare. On the contrary, the impact of the price changes is diluted and obscured the more environmental variables are included within the scope of an unconditional index. In order to qualify as a price index, a COLI must be conditional.

Specific types of transactions

1.167 Given that conceptually, a CPI is an index that measures changes in the prices of consumption goods and services, expenditures on items that are not consumption goods and services fall outside the scope of the CPI; for example, expenditures on assets such as land or bonds, shares and other financial assets. Similarly, payments that do not involve any flows of goods or services in return for the payments are outside the scope; for example, payments of income taxes or social security contributions.

1.168 *Transfers*. A transfer occurs when one economic unit provides a good, service or asset, including money, to another without receiving any counterpart good, service or asset in return. As no good or service is acquired when a household makes a transfer, the transfer must be outside the scope. For this reason, compulsory cash transfers, such as payments of direct taxes on income or wealth, must be outside the scope of a CPI. It is not always clear, however, whether certain payments to government are transfers or purchases of services. For example, payments to obtain certain kinds of licences are sometimes taxes under another name, whereas in other cases the government may provide a service by exercising some kind of supervisory, regulatory or control function. Gifts or donations must be transfers and therefore outside the scope. On the other hand, subscriptions to clubs and societies which provide their members with some kind of service in return are included. Tips and gratuities can be borderline cases. When they are effectively an expected, even obligatory, part of the payment for a service they are not transfers and should be treated as part of the price paid.

1.169 *Undesirable or illegal goods or services.* All goods and services that households *willingly* buy on the market to satisfy their own needs and wants should be included, even if most people might regard them as undesirable or even if they are prohibited by law. Of course, illegal goods and services may have to be excluded in practice because the requisite data cannot be collected.

1.170 *Financial transactions.* Financial transactions occur when one kind of financial asset is exchanged for another, bearing in mind that money is itself a financial asset. For example, the purchase of a bond or share is a financial transaction. Borrowing is a financial transaction in which cash is exchanged, the counterpart being the creation of a financial asset or liability.

1.171 No consumption occurs when a financial transaction takes place, even though financial transactions may be undertaken in order to facilitate future consumption. Financial transactions as such are not covered by CPIs because, by definition, no goods are exchanged, nor services provided, in financial transactions. However, some "financial" transactions may not be entirely financial because they may include an explicit or implicit service charge in addition to the provision of an asset, such as a loan. As a service charge constitutes the purchase of a service by the household, it should be included in a CPI, although it may be difficult to separate out the service charge in some cases. For example, foreign exchange transactions are financial transactions in which one financial asset is exchanged for another. Changes in the price of a foreign currency in terms of the domestic currency resulting from changes in the exchange rate are outside the scope of a CPI. On the other hand, the commission charges associated with the exchange of currencies are included as payments for the services rendered by the foreign exchange dealers.

1.172 Households may borrow in order to make large expenditures on durables or houses, but also to finance large educational or health expenses, or even expensive holidays. Whatever the purpose of the borrowing, the financial transaction in which the loan is contracted is outside the scope of a CPI. The treatment of the interest payable on loans is a separate issue considered below.

1.173 *Composite transactions.* As just noted, some transactions are composite transactions containing two or more components whose treatment may be quite different for CPI purposes. For example, part of a life insurance premium is a financial transaction leading to the creation of a financial claim and is therefore outside the scope, whereas the remainder consists of a service charge which should be covered by a CPI. The two components are not separately itemized, however.

1.174 As explained in Chapter 3, the treatment of payments of nominal interest is difficult because it may have four conceptually quite different components:

– a pure interest payment;
– a risk premium that depends on the creditworthiness of the borrower;
– a service charge payable to the bank, moneylender or other financial institution engaged in the business of making loans;
– a payment to compensate the creditor for the real holding loss incurred on the principal of the loan during inflation.

The fourth component is clearly outside the scope of a CPI as it is a capital flow. Conversely, the third component, the service charge, should clearly be included. The treatment of the first two components is controversial. When there is significant inflation or a very imperfect capital market, payments of nominal interest may be completely dominated by the last two components, both of which are conceptually quite different from the concept of pure interest. For example, the "interest" charged by a village moneylender may be mostly a high service charge. It may be impossible to decompose the various components of nominal interest in practice. The treatment of nominal interest as a whole remains difficult and somewhat controversial.

Household production

1.175 When households engage in production for the market, the associated business transactions are all outside the scope of a CPI. Expenditures incurred for business purposes are excluded, even though they involve purchases of goods and services that might be used to satisfy the personal needs and wants of members of the household instead.

1.176 Households also produce goods and services for their own consumption, mainly service production such as the preparation of meals, the care of children, the sick or the elderly, the cleaning and maintenance of durables and dwellings, the transportation of household members, and so on. Owner-occupiers produce housing services for their own consumption. Households also grow vegetables, fruit, flowers or other crops for their own use.

1.177 Many of the goods or services purchased by households do not provide utility directly but are used as inputs into the production of other goods and services that do provide utility: for example, raw foodstuffs, fertilizers, cleaning materials, paints, electricity, coal, oil, petrol, and so on.

1.178 In principle, a CPI should record changes in the prices of the outputs from these production activities, as it is the outputs rather than the inputs that are actually consumed and provide utility. However, as the outputs are not themselves purchased, no prices can be observed for them. Prices could be imputed for them equal to the prices they would fetch on the market, but this would make a CPI heavily dependent on assumed rather than collected prices. The pragmatic solution recommended in Chapter 3 is to treat all goods and services purchased on the market to be used exclusively as inputs into the production of other goods and services that are directly consumed by households as if they were themselves consumption goods and services. On this principle, goods such as insecticides and electricity are treated as providing utility indirectly and included in CPIs. This is, of course, the solution usually adopted in practice not only for CPIs but also in national accounts, where most expenditures on inputs into household production are classified as final consumption expenditures.

1.179 In some countries, there is an increasing tendency for households to purchase prepared, take-away meals rather than the ingredients. As the prices of such meals cost more than the sum of the ingredients that the households previously purchased, the weight attached to food consumption increases. This partly reflects the fact that the costs of the households' own labour inputs into the preparation of meals were previously ignored. Various kinds of household service activities that were previously outside the scope of a CPI may be brought within the scope if households choose to pay others to perform the services.

1.180 *Subsistence agriculture and owner-occupied housing*. In the case of two important types of production for own consumption within households, namely agricultural production for own consumption and housing services produced by owner-occupiers, the national accounts do actually try to record the values of the outputs produced and consumed rather than the inputs. Similarly, CPIs may also try to price the outputs rather than the inputs in these two cases.

1.181 In principle, the prices of the outputs from own-account agricultural production may be included in CPIs, even though they are imputed. On the other hand, for households relying on subsistence agriculture, the prices of inputs of agricultural materials purchased on the market may be their main exposure to inflation. Two points may be noted. First, the imputed market value of the output should usually be greater than the costs of the purchased inputs, if only because it should cover the costs of the labour inputs provided by the household. Thus, pricing the purchased inputs rather than the outputs may mean that the consumption of own agricultural production in CPIs does not receive sufficient weight. Second, double counting should be avoided. If the imputed prices of the outputs are included, the actual prices of the inputs consumed should not be included as well.

1.182 In the case of owner-occupied housing, the situation is complicated by the fact that the production requires the use of the capital services provided by a large fixed asset in the form of the dwelling itself. Even if the inputs into the production of housing services are priced for CPI purposes, it is still necessary to impute prices for the inputs of capital services (mainly depreciation plus interest) provided by the dwelling. Some countries therefore prefer to impute the prices of the outputs of housing services actually consumed on the basis of the rents payable for the same kind of dwellings rented on the market. The treatment of owner-occupied housing is complex, and somewhat controversial, and is considered in Chapters 3, 9, 10 and 23, among others.

Coverage of households and outlets

1.183 As explained in Chapter 3, households may be either individual persons or groups of persons living together who make common provision for food or other essentials for living. A CPI may be required to cover:

– *either* the consumption expenditures made by households resident in a particular area, usually a country or region, whether the expenditures are made inside or outside the area – this is called the "national" concept of expenditure;
– *or* the consumption expenditures that take place within a particular area, whether made by households resident in that area or residents of other areas – this is called the "domestic" concept.

Adopting the domestic concept may make it more difficult to collect the relevant disaggregated expenditure data in household surveys. A CPI may also be defined to cover a group of countries, such as the European Union.

1.184 Not all kinds of households have to be included. As explained in Chapter 3, some countries choose to exclude particular categories of households such as very wealthy households or households engaged in agriculture. Some countries also compile different indices designed to cover different groups of households, such as households resident in different regions. Another possibility is to compile a general CPI designed to cover all or most households and, in addition, one or more special indices aimed at particular sections of the community, such as households headed by pensioners. The precise coverage of households is a matter of choice. It is inevitably influenced by what are believed to be the main uses of the index. The set of households actually covered by the CPI is described as the "reference population".

Price variation

1.185 Prices for exactly the same good or service may vary between different outlets, while different prices may sometimes be charged to different types of customers. Prices may also vary during the course of the month to which the index relates. Conceptually, it is necessary to distinguish such pure price variation from price differences that are attributable to differences in the quality of the goods or services offered, although it is not always easy to distinguish between the two in practice. The existence of pure price differences reflects some form of market imperfections, such as consumers' lack of information or price discrimination.

1.186 When pure price differences exist, a change in market conditions may make it possible for some households to switch from purchasing at higher prices to purchasing at lower prices, for example if new outlets open that offer lower prices. The resulting fall in the average price paid by households counts as a price fall for CPI purposes, even though the price charged by each individual outlet may not change. If the prices are collected from the outlets and switches in households' purchasing habits remain unobserved, the CPIs are said to be subject to outlet substitution bias, as explained in more detail in Chapter 11. On the other hand, when the price differences reflect differences in the quality of the goods and services sold in the different outlets, switching from outlets selling at higher prices to outlets selling at lower prices simply means that households are choosing to purchase lower-quality goods or services. In itself, this does not imply any change in price.

Classifications

1.187 As explained in Chapter 3, the classification of household expenditures used in a CPI provides the necessary framework for the various stages of CPI compilation. It provides a structure for purposes of weighting and aggregation, and also a basis for stratifying the samples of products whose prices are collected. The goods and services covered by a CPI may be classified in several ways: not simply on the basis of their physical characteristics but also by the purposes they serve and the degree of similarity of their price behaviour. Product-based and purpose-based classifications differ but can usually be successfully mapped onto each other. In practice, most countries use a hybrid classification system in which the breakdown at the highest level is by purpose while the lower-level breakdowns are by product type. This is the case for the recently revised internationally agreed Classification of Individual Consumption according to Purpose (COICOP), which provides a suitable classification for CPI purposes.

1.188 The first level of classification in COICOP consists of 12 divisions covering total consumption expenditures of households. As just noted, the breakdown into divisions is essentially by purpose. At the second level of disaggregation, the 12 *divisions* are divided into 47 *groups* of products, which are in turn divided into 117 *classes* of products at the third level. Chapter 3 provides a listing of ten classes of goods defined as durables in COICOP. It also gives a list of seven classes described as semi-durables, such as clothing, footwear and household textiles.

1.189 The 117 classes at the lowest level of aggregation of COICOP are not sufficiently detailed for CPI purposes. They can be divided into sub-classes using the sub-classes of the internationally agreed Central Product Classification (CPC). Even some of these may require further breakdown in order to arrive at some of the elementary aggregates used for CPI purposes. In order to be useful for CPI purposes, expenditure weights must be available for the various sub-classes or elementary aggregates. From a sampling perspective, it is desirable for the price movements of the individual products within the elementary aggregates to be as homogeneous as possible. The elementary aggregates may also be divided into strata for sampling purposes, on the basis of location or the type of outlet in which the products are sold.

Consumer price indices and national accounts price deflators

1.190 Appendix 3.1 of Chapter 3 explains the differences between the overall CPI and the deflator for total household consumption expenditures in national accounts. In practice, CPIs may be designed to cover only a subset of the households and a subset of the expenditures covered by the national accounts. Moreover, the index number formulae needed for CPIs and national accounts deflators may be different. These differences mean that the overall CPI is generally not the same as the deflator for total household consumption expenditures in the national accounts. On the other hand, the basic price and expenditure data collected and used for CPI purposes are also widely used to build up the price indices needed to deflate the individual components of household consumption in the national accounts.

Expenditure weights

1.191 As already noted, there are two main stages in the calculation of a CPI. The first is the collection of the price data and the calculation of the elementary price indices. The second is the averaging of the elementary price indices to arrive at price indices at higher levels of aggregation up to the overall CPI itself. Expenditure data are needed for the elementary aggregates that can be used as weights in the second stage. These weights are needed whatever index number formula is used for aggregation purposes. Chapter 4 is concerned with the derivation, and sources, of the expenditure weights.

Household expenditure surveys and national accounts

1.192 The principal data source for household consumption expenditures in most countries is a household expenditure survey (HES). An HES is a sample survey of thousands of households that are asked to keep records of their expenditures on different kinds of consumer goods and services over a specified period of time, such as a week or longer. The size of the sample obviously depends on the resources available, but also on the extent to which it is desired to break down the survey results by region or type of household. HESs are costly operations. This manual is not concerned with the conduct of HESs or with general sampling survey techniques or procedures. There are several standard texts on survey methods to which reference may be made. Household expenditure surveys may be taken at specified intervals of time, such as every five years, or they may be taken each year on a continuing basis.

1.193 HESs can impose heavy burdens on the respondents, who have to keep detailed expenditure records of a kind that they would not normally keep, although this may become easier when supermarkets or other retail outlets provide detailed printouts of purchases. HESs tend to have some systematic biases. For example, many households either deliberately, or unconsciously, understate the amounts of their expenditures on certain ‘‘undesirable’’ products, such as gambling, alcoholic drink, tobacco or drugs. Corrections can be made for such biases. Moreover, the data collected in HESs may also need to be adjusted to bring them into line with the concept of expenditure required by the CPI. For example, the imputed expenditures on the housing services produced and consumed by owner-occupiers are not collected in HESs.

1.194 As explained in Chapter 14, the use of the commodity flow method within the supply and use tables of the SNA enables data drawn from different primary sources to be reconciled and balanced against each other. The commodity flow method may be used to improve estimates of household consumption expenditures derived from expenditure surveys by adjusting them to take account of the additional information provided by statistics on the sales, production, imports and exports of consumer goods and services. By drawing on various sources, the household expenditure data in the national accounts may provide the best estimates of aggregate household expenditures, although the classifications used may not be fine enough for CPI purposes. Moreover, because HESs may be conducted only at intervals of several years, the expenditure data in the national accounts may be more up to date, as national accounts are able to draw upon other kinds of more recent data, such as retail sales and the production and import of consumer goods and services. It is important to note, however, that national accounts should not be viewed as if they were an alternative, independent data source to HESs. On the contrary, HESs provide one of the main sources for the expenditure data on household consumption used to compile national accounts.

1.195 Household expenditure surveys in many countries may not be conducted as frequently as might be desired for CPI, or national accounts, purposes. National HESs can be very costly and onerous for the households, as already noted. They may be conducted only once every five or ten years, or even at longer intervals. In any case, conducting and processing HESs is time-consuming, so the results may not be available for CPI purposes until one or two years after the surveys have been conducted. It is for these practical reasons that CPIs in many countries are Lowe indices that use the quantities of some base period b that may precede the time reference period 0 by a few years and period t by many years.

1.196 Some countries conduct continuous HESs not only in order to update their CPI weights but also to improve their national accounts. Of course, the same panel of households does not have to be retained indefinitely; the panel can be gradually rotated by dropping some households and replacing them by others. Countries that conduct continuous expenditure surveys are able to revise and update their expenditure weights each year so that the CPI becomes a chain index with annual linking. Even with continuous expenditure surveys, however, there is a lag between the time at which the data are collected and the time at which the results are processed and ready for use, so that it is never possible to have survey results that are contemporaneous with the price changes. Thus, even when the weights are updated annually, they still refer to some period that precedes the time reference period. For example, when the price reference period is January 2000, the expenditure weights may refer to 1997 or 1998, or both years. When the price reference period moves forward to January 2001, the weights move forward to 1998 or 1999, and so on. Such an index is a chain Lowe index.

1.197 Some countries prefer to use expenditure weights that are the average rates of expenditure over periods of two or three years in order to reduce ‘‘noise’’ caused by errors of estimation (the expenditure surveys are only samples) or erratic consumer behaviour over short periods of time resulting from events such as booms or recessions, stock market fluctuations, oil shocks, or natural or other disasters.

Other sources for estimating expenditure weights

1.198 If expenditures need to be disaggregated by region for sampling or analytical purposes, it is possible to supplement whatever information may be available by region in HESs by using data from population censuses. Another data source may be food surveys. These are special surveys, conducted in some countries, that focus on households' expenditures on food products. They can provide more detailed information on food expenditures than that available from HESs.

1.199 Another possible source of information consists of points of purchase (POP) surveys, which are conducted in some countries. A POP survey is designed to provide information about the retail outlets at which households purchase specified groups of goods and services. Households are asked, for each item, about the amounts spent in each outlet and the names and addresses of the outlets. The main use for a POP survey is for selecting the sample of outlets to be used for price collection purposes.

Collection of price data

1.200 As explained in Chapter 9, there are two levels of calculation involved in a CPI. At the lower level, samples of prices are collected and processed to obtain lower-level price indices. These lower-level indices are the elementary indices, whose properties and behaviour are studied in Chapter 20. At the higher level, the elementary indices are averaged to obtain higher-level indices using expenditures as weights. At the higher level, all the index number theory elaborated in Chapters 15 to 18 comes into play.

1.201 Lower-level indices are calculated for elementary aggregates. Depending on the resources available and procedures adopted by individual countries, these elementary aggregates could be sub-classes or micro-classes of the expenditure classification described above. If it is desired to calculate CPIs for different regions, the sub-classes or micro-classes have to be divided into strata referring to the different regions. In addition, in order to improve the efficiency of the sampling procedures used to collect prices, it will usually be desirable, if feasible, to introduce other criteria into the definitions of the strata, such as the type of outlet. When the sub-classes or micro-classes are divided into strata for data collection purposes, the strata themselves become the elementary aggregates. As a weight needs to be attached to each elementary aggregate in order to calculate the higher-level indices, an estimate of the expenditure within each elementary aggregate must be available. Expenditure or quantity data are typically not available within an elementary aggregate, so the elementary indices have to be estimated from price data alone. This may change if scanner data from electronic points of sale become more available.

1.202 Chapter 5 is concerned with sampling strategies for price collection. Chapter 6 is concerned with the methods and operational procedures actually used to collect prices. In principle, the relevant prices for a CPI should be the purchasers' prices actually paid by households, but it is generally neither practical nor cost-effective to try to collect prices each month or quarter directly from households, even though expenditure data are collected directly from households in household expenditure surveys. In practice, the prices that are collected are not actual transaction prices, but rather the prices at which goods and services are offered for sale in outlets such as retail shops, supermarkets or service providers. However, it may become increasingly feasible to collect actual transactions prices as more goods and services are sold through electronic points of sale that record both prices and expenditures.

Random sampling and purposive sampling

1.203 Given that the prices are collected from the sellers, there are two different sampling problems that arise. The first is how to select the individual products within an elementary aggregate whose prices are to be collected. The second is how to select a sample of outlets selling those products. For some products, it may not be necessary to visit retail outlets to collect prices because there may be only a single price applying throughout the country. Such prices may be collected from the central organization responsible for fixing the prices. The following paragraphs refer to the more common situation in which prices are collected from a large number of outlets.

1.204 As explained in Chapter 5, the universe of products from which the sample is taken has several dimensions. The products may be classified not only on the basis of the characteristics and functions that determine their place in COICOP, but also according to the locations and outlets at which they are sold and the times at which they are sold. The fact that the universe is continually changing over time is a major problem, not only for CPIs but also for most other economic statistics. Products disappear to be replaced by other kinds of products, while outlets close and new ones open. The fact that the universe is changing over time creates both conceptual and practical problems, given that the measurement of price changes over time requires some continuity in the products priced. In principle, the price changes recorded should refer to matched products that are identical in both time periods. The problems created when products are not identical are considered in some detail later.

1.205 In designing the sample for price collection purposes, due attention should be paid to standard statistical criteria to ensure that the resulting sample estimates are not only unbiased and efficient in a statistical sense, but also cost-effective. There are two types of bias encountered in the literature on index numbers, namely *sampling bias* as understood here and the *non-sampling biases* in the form of substitution bias or representativity bias, as discussed in Chapter 10. It is usually clear from the context which type of bias is meant.

1.206 There is a large literature on sampling survey techniques to which reference may be made and which need not be summarized here. In principle, it would be desirable to select both outlets and products using random sampling with known probabilities of selection. This ensures that the sample of products selected is not distorted by subjective factors and enables sampling errors to be calculated. Many countries nevertheless continue to rely heavily on the purposive selection of outlets and products, because random sampling may be too difficult and too costly. Purposive selection is believed to be more cost-effective, especially when the sampling frames available are not comprehensive and not well suited to CPI purposes. It may also be cost-effective to collect a "cluster" of prices on different products from the same outlet, instead of distributing the price collection more thinly over a larger number of outlets.

1.207 Efficient sampling, whether random or purposive, requires comprehensive and up-to-date sampling frames. Two types of frames are needed for CPI purposes: one listing the universe of outlets, and the other listing the universe of products. Examples of possible sampling frames for outlets are business registers, central or local government administrative records or telephone directories. When the sampling frames contain the requisite information, it may be possible to increase efficiency by selecting samples of outlets using probabilities that are proportional to the size of some relevant economic characteristic, such as the total value of sales. Sampling frames for products are not always readily available in practice. Possible frames are catalogues or other product lists drawn up by major manufacturers, wholesalers or trade associations, or lists of products that are specific to individual outlets such as large supermarkets.

1.208 Depending on the information available in the sampling frame, it may be possible to group the outlets into strata on the basis of their location and size, as indicated by sales or employees. When there is information about size, it may be possible to increase efficiency by taking a random sample of outlets with probabilities proportional to size. In practice, however, there is also widespread use of purposive sampling.

1.209 In most countries, the selection of most of the individual items to be priced within the selected outlets tends to be purposive, being specified by the central office responsible for the CPI. The central office draws up lists of products that are deemed to be representative of the products within an elementary aggregate. The lists can be drawn up in collaboration with managers of wholesale or large retail establishments, or other experts with practical experience and knowledge. The actual procedures are described in more detail in Chapter 6.

1.210 It has been argued that the purposive selection of products is liable to introduce only a negligible amount of sampling bias, although there is not much conclusive evidence on this matter. In principle, random sampling is preferable and it is also quite feasible. For example, the United States Bureau of Labor Statistics makes extensive use of random selection procedures to select both outlets and products within outlets. When the selection of products is delegated to the individual price collectors, it is essential to ensure that they are well trained and briefed, and closely supervised and monitored.

Methods of price collection

1.211 The previous section focused on the sampling issues that arise when prices have to be collected for a large number of products from a large number of outlets. This section is concerned with some of the more operational aspects of price collection.

1.212 *Central price collection*. Many important prices can be collected directly by the central office responsible for the CPI from the head office of the organization responsible for fixing the prices. When prices are the same throughout the country, collection from individual outlets is superfluous:

- Some tariffs or service charges are fixed nationally and apply throughout the country. This may be the case for public utilities such as water, gas and electricity, postal services and telephone charges, or public transport. The prices or charges can be obtained from the relevant head offices.
- Some national chains of stores or supermarkets may charge the same prices everywhere, in which case the prices can be obtained from their head offices. Even when national chains do not charge uniform prices, there may be only a few minor regional differences in the prices and all the relevant information may be obtainable centrally.
- Many of these prices determined centrally may change very infrequently, perhaps only once or twice or year, so they do not have to be collected monthly. Moreover, many of these prices can be collected by telephone, fax or email and may not require visits to the head offices concerned.

1.213 *Scanner data*. One important new development is the increasing availability in many countries of large amounts of very detailed "scanner" data obtained from electronic points of sale. Such data are collated by commercial databases. Scanner data are up to date and comprehensive. An increasingly large proportion of all goods sold are being scanned as they pass through electronic points of scale.

1.214 The potential benefits of using scanner data are obviously considerable and could ultimately have a significant impact on the way in which price data are collected for CPI purposes. Not enough experience is yet available to provide general guidelines about the use of scanner data. Clearly, statistical offices should monitor developments in this field closely and explore the possibility of exploiting this major new source of data. Scanner data also increase the scope for using improved methods of quality adjustment, including hedonic methods, as explained in Chapter 7.

1.215 *Local price collection*. When prices are collected from local outlets, the individual products selected for pricing can be determined in two ways. One way is for a specific list of individual products to be determined in advance by the central office responsible for the CPI. Alternatively, the price collector can be given the discretion to choose from a specified range of products. The collector may use some kind of random selection procedure, or select the products that sell the most or are recommended by the shop owner or manager. An individual product selected for pricing in an individual outlet may be described as a sampled product. It may be a good or a service.

1.216 When the list of products is determined in advance by the central office, the objective is usually to select products that are considered to be representative of the larger group of products within an elementary aggregate. The central office also has to decide how loosely or tightly to describe, or specify, the representative products selected for pricing. In theory, the number of different products that might be identified is to some extent arbitrary, depending on the number of economic characteristics that are deemed to be relevant or important. For example, "beef" is a generic term for a group of similar but nevertheless distinct products. There are many different cuts of beef, such as minced beef, stewing steak or rump steak, each of which can be considered a different product and which can sell at very different prices. Furthermore, beef can also be classified according to whether it is fresh, chilled or frozen, and cross-classified again according to whether it comes from domestic or imported animals, or from animals of different ages or breeds.

1.217 Tightening the specifications ensures that the central office has more control over the items actually priced in the outlets, but it also increases the chance that some products may not actually be available in some outlets. Loosening the specifications means that more items may be priced but leaves the individual price collectors with more discretion with regard to the items actually priced. This could make the sample as a whole less representative.

Continuity of price collection

1.218 A CPI is intended to measure pure price changes. The products whose prices are collected and compared in successive time periods should ideally be perfectly *matched*; that is, they should be identical in respect of their physical and economic characteristics. When the products are perfectly matched, the observed price changes are *pure* price changes. When selecting representative products, it is therefore necessary to ensure that enough of them can be expected to remain on the market over a reasonably long period of time in exactly the same form or condition as when first selected. Without continuity, there would not be enough price changes to measure.

1.219 Having identified the items whose prices are to be collected, the normal strategy is to continue pricing exactly those same items for as long as possible. Price collectors can do this if they are provided with very precise, or tight, specifications of the items to be priced.

Alternatively, they must keep detailed records themselves of the items that they have selected to price.

1.220 The ideal situation for a price index would be one in which all the products whose prices are being recorded remain on the market indefinitely without any change in their physical and economic characteristics, except of course for the timing of their sale. It is worth noting that many theorems in index number theory are derived on the assumption that exactly the same set of goods and services is available in both the time periods being compared. Most products, however, have only a limited economic life. Eventually, they disappear from the market to be replaced by other products. As the universe of products is continually evolving, the representative products selected initially may gradually account for a progressively smaller share of total purchases and sales. As a whole, they may become less and less representative. As a CPI is intended to cover all products, some way has to be found to accommodate the changing universe of products. In the case of consumer durables whose features and designs are continually being modified, some models may have very short lives indeed, being on the market for only a year or less before being replaced by newer models.

1.221 At some point the continuity of the series of price observations may have to be broken. It may become necessary to compare the prices of some products with the prices of other new ones that are very similar but not identical. Statistical offices must then try to eliminate from the observed price changes the estimated effects of the changes in the characteristics of the products whose prices are compared. In other words, they must try to adjust the prices collected for any changes in the quality of the products priced, as explained in more detail below. At the limit, a completely new product may appear that is so different from those existing previously that quality adjustment is not feasible and its price cannot be directly compared with that of any previous product. Similarly, a product may become so unrepresentative or obsolete that it has to be dropped from the index because it is no longer worth trying to compare its price with those of any of the products that have displaced it.

Resampling

1.222 One strategy to deal with the changing universe of products would be to resample, or reselect, at regular intervals the complete set of items to be priced. For example, with a monthly index, a new set of items could be selected each January. Each set of items would be priced until the following January. Two sets have to be priced each January in order to establish a link between each set of 12 monthly changes. Resampling each year would be consistent with a strategy of updating the expenditure weights each year.

1.223 Although resampling may be preferable to maintaining an unchanged sample or selection, it is not used much in practice. Systematically resampling the entire set of products each year would be difficult to manage and costly to implement. Moreover, it does not provide a complete solution to the problem of the changing universe of products, as it does not capture price changes that occur at the moment of time when new products or new qualities are first introduced. Many producers deliberately use the time when products are first marketed to make significant price changes.

1.224 A more practical way in which to keep the sample up to date is to rotate it gradually by dropping certain items and introducing new ones. Items may be dropped for two reasons:

- The product is believed by the price collector or central office to be no longer representative. It appears to account for a steadily diminishing share of the total expenditures within the basic categories in question.
- The product may simply disappear from the market altogether. For example, it may have become obsolete as a result of changing technology or unfashionable because of changing tastes, although it could disappear for other reasons.

1.225 At the same time, new products or new qualities of existing products appear on the market. At some point, it becomes necessary to include them in the list of items priced. This raises the general question of the treatment of quality change and the treatment of new products.

Adjusting prices for quality changes

1.226 The treatment of quality change is perhaps the greatest challenge facing CPI compilers. It is a recurring theme throughout this manual. It presents both conceptual and practical problems for compilers of CPIs. The whole of Chapter 7 is devoted to the treatment of quality change, while Chapter 8 addresses the closely related topic of new goods and item substitution.

1.227 When a sampled product is dropped from the list of products priced in some outlet, the normal practice is to find a new product to replace it in order to ensure that the sample, or selection, of sampled products remains sufficiently comprehensive and representative. If the new product is introduced specifically to replace the old one, it is necessary to establish a link between the series of past price observations on the old item and the subsequent series for the new item. The two series of observations may, or may not, overlap in one or more periods. In many cases, there can be no overlap because the new quality, or model, is only introduced after the one which it is meant to replace is discontinued. Whether or not there is an overlap, the linking of the two price series requires some estimate of the change in quality between the old product and the product selected to replace it.

1.228 However difficult it is to estimate the contribution of the change in quality to the change in the observed price, it must be clearly understood that some estimate has to be made either explicitly or, by default, implicitly. The issue cannot be avoided or bypassed. All statistical offices have limited resources and many may not have the capacity to undertake the more elaborate explicit adjustments for quality change described in

Chapter 7. Even though it may not be feasible to undertake an explicit adjustment through lack of data or resources, it is not possible to avoid making some kind of implicit adjustment. Even apparently "doing nothing" necessarily implies some kind of implicit adjustment, as explained below. Whatever the resources available to them, statistical offices must be conscious of the implications of the procedures they adopt.

1.229 Three points are stressed in the introductory section of Chapter 7:

- The pace of innovation is high, and possibly increasing, leading to continual changes in the characteristics of products.
- There is not much consistency between countries in the methods they use to deal with quality change.
- A number of empirical studies have demonstrated that the choice of method does matter, as different methods can lead to very different results.

Evaluation of the effect of quality change on price

1.230 It is useful to try to clarify why one would wish to adjust the observed price change between two items that are similar, but not identical, for differences in their quality. A change in the quality of a good or service occurs when there is a change in some, but not most, of its characteristics. For purposes of a CPI, a quality change must be evaluated from the consumer's perspective. As explained in Chapter 7, the evaluation of the quality change is essentially an estimate of the additional amount that a consumer is willing to pay for the new characteristics possessed by the new quality. This additional amount is not a price increase because it represents the monetary value of the additional satisfaction or utility that is derived from the new quality. Of course, if the old quality is preferred to the new one, consumers would only be willing to buy the new quality if its price were lower.

1.231 Consider the following hypothetical experiment in which a new quality appears alongside an old one. Assume that the two products are substitutes and that the consumer is familiar with the characteristics of the old and the new qualities. Use lower case p to refer to prices of the old quality and upper case P for the prices of the new quality. Suppose that both qualities are offered to the consumer at the same price, namely the price P_t at which the new quality is actually being sold in period t. The consumer is then asked to choose between them and prefers the new quality.

1.232 Suppose next that the price of the old quality is progressively reduced until it reaches p_t^*, at which point the consumer becomes indifferent between purchasing the old quality at p_t^* and the new quality at P_t. Any further decrease below p_t^* causes the consumer to switch back to the old quality. The difference between P_t and p_t^* is a measure of the additional value that the consumer places on the new quality as compared with the old quality. It measures the maximum amount that the consumer is willing to pay for the new quality over and above the price of the old quality.

1.233 Let p_{t-1} denote the actual price at which the old quality was sold in period $t-1$. For CPI purposes, the price increase between the two qualities is not the observed difference $P_t - p_{t-1}$ but $p_t^* - p_{t-1}$. It is important to note that p_t^*, the hypothetical price for the old quality in period t, is directly comparable with the actual price of the old quality in period $t-1$ because both refer to the same identical product. The difference between them is a *pure* price change. The difference between P_t and p_t^* is not a price change but an evaluation of the difference in the quality of the two items in period t. The actual price of the new quality in period t needs to be multiplied by the ratio p_t^*/P_t in order to make the comparison between the prices in periods $t-1$ and t a comparison between products of equal quality in the eyes of the consumer. The ratio p_t^*/P_t is the required quality adjustment.

1.234 Of course, it is difficult to estimate the quality adjustment in practice, but the first step has to be to clarify conceptually the nature of the adjustment that is required in principle. In practice, producers often treat the introduction of a new quality, or new model, as a convenient opportunity at which to make a significant price change. They may deliberately make it difficult for consumers to disentangle how much of the observed difference in price between the old and the new qualities represents a price change.

1.235 Chapter 7 explains the two possibilities open to statistical offices. One possibility is to make an explicit adjustment to the observed price change on the basis of the different characteristics of the old and new qualities. The other alternative is to make an implicit adjustment by making an assumption about the pure price change; for example, on the basis of price movements observed for other products. It is convenient to take the implicit methods first.

Implicit methods for adjusting for quality changes

1.236 *Overlapping qualities*. Suppose that the two qualities overlap, both being available on the market at time t. If consumers are well informed, have a free choice and are collectively willing to buy some of both at the same time, economic theory suggests that the ratio of the prices of the new to the old quality should reflect their relative utilities to consumers. This implies that the difference in price between the old and the new qualities does not indicate any change in price. The price changes up to period t can be measured by the prices for the old quality, while the price changes from period t onwards can be measured by the prices for the new quality. The two series of price changes are linked in period t, the difference in price between the two qualities not having any impact on the linked series.

1.237 When there is an overlap, simple linking of this kind may provide an acceptable solution to the problem of dealing with quality change. In practice, however, this method is not used very extensively because the requisite data are seldom available. Moreover, the conditions may not be consistent with those assumed in the theory. Even when there is an overlap, consumers may not have had time to acquire sufficient knowledge of the characteristics to be able to evaluate the relative qualities properly, especially when there is a substantial change in quality. Not all consumers may have access to both qualities. When the new quality first appears, the market is liable to remain in disequilibrium for some time, as it takes time for consumers to adjust their consumption patterns.

1.238 There may be a succession of periods in which the two qualities overlap before the old quality finally disappears from the market. If the market is temporarily out of equilibrium, the relative prices of the two qualities may change significantly over time so that the market offers alternative evaluations of the relative qualities depending on which period is chosen. When new qualities that embody major new improvements appear on the market for the first time, there is often a tendency for their prices to fall relatively to older qualities before the latter eventually disappear. In this situation, if the price series for the old and new qualities are linked in a single period, the choice of period can have a substantial effect on the overall change in the linked series.

1.239 The statistician has then to make a deliberate judgement about the period in which the relative prices appear to give the best representation of the relative qualities. In this situation, it may be preferable to use a more complex linking procedure which uses the prices for both the new and the old qualities in several periods in which they overlap. However, the information needed for this more complex procedure will never be available if price collectors are instructed only to introduce a new quality when an old one is dropped. In this case, the timing of the switch from the old to the new can have a significant effect on the long-term change in the linked series. This factor must be explicitly recognized and taken into consideration.

1.240 If there is no overlap between the new and the old qualities, the problems just discussed do not arise as no choice has to be made about when to make the link. Other and more difficult problems nevertheless take their place.

1.241 *Non-overlapping qualities.* In the following sections, it is assumed that the overlap method cannot be used because there is a discontinuity between the series of price observations for the old and new qualities. Again, using lower case p for the old quality and upper case P for the new, it is assumed that the price data available to the index compiler take the following form:

$$\ldots, p_{t-3}, p_{t-2}, p_{t-1}, P_t, P_{t+1}, P_{t+2}, \ldots$$

The problem is to estimate the pure price change between $t-1$ and t in order to have a continuous series of price observations for inclusion in the index. Using the same notation as above:

- price changes up to period $t-1$ are measured by the series for the old quality;
- the change between $t-1$ and t is measured by the ratio p_t^*/p_{t-1} where p_t^* is equal to P_t *after* adjustment for the change in quality;
- price changes from period t onwards are measured by the series for the new quality.

1.242 The problem is to estimate p_t^*. This may be done explicitly by one of the methods described later. Otherwise, one of the implicit methods has to be used. These may be grouped into three categories:

- The first solution is to assume that $p_t^*/p_{t-1} = P_t/p_{t-1}$ or $p_t^* = P_t$. No change in quality is assumed to have occurred, so the whole of the observed price increase is treated as a pure price increase. In effect, this contradicts the assumption that there has been a change in quality.
- The second is to assume that $p_t^*/p_{t-1} = 1$, or $p_t^* = p_{t-1}$. No price change is assumed to have occurred, the whole of the observed difference between p_{t-1} and P_t being attributed to the difference in their quality.
- The third is to assume that $p_t^*/p_{t-1} = I$, where I is an index of the price change for a group of similar products, or possibly a more general price index.

1.243 The first two possibilities cannot be recommended as default options to be used automatically in the absence of any adequate information. The use of the first option can only be justified if the evidence suggests that the extent of the quality change is negligible, even though it cannot be quantified more precisely. "Doing nothing", in other words ignoring the quality change completely, is equivalent to adopting the first solution. Conversely, the second can only be justified if the evidence suggests that the extent of any price change between the two periods is negligible. The third option is likely to be much more acceptable than the other two. It is the kind of solution that is often used in economic statistics when data are missing.

1.244 Elementary indices are typically based on a number of series relating to different sampled products. The particular linked price series relating to the two qualities is therefore usually just one out of a number of parallel price series. What may happen in practice is that the price observations for the old quality are used up to period $t-1$ and the prices for the new quality from t onwards, the price change between $t-1$ and t being omitted from the calculations. In effect, this amounts to using the third option: that is, estimating the missing price change on the assumption that it is equal to the average change for the other sampled products within the elementary aggregate.

1.245 It may be possible to improve on this estimate by making a careful selection of the other sampled products whose average price change is believed to be more similar to the item in question than the average for the group of sampled products as a whole. This procedure is described in some detail in Chapter 7, where it is illustrated with a numerical example and described as "targeting" the imputation or estimation.

1.246 The general method of estimating the price on the basis of the average change for the remaining group of products is widely used. It is sometimes described as the "overall" class mean method. The more refined targeted version is the "targeted" mean method. In general, one or other method seems likely to be preferable to either of the first two options listed above, although each case must be considered on its individual merits.

1.247 While the class mean method seems a sensible practical solution, it may nevertheless give biased results, as explained in Chapter 7. The introduction of a new quality is precisely the occasion on which a producer may choose to make a significant price change. Many of the most important price changes may be missed if, in effect, they are assumed to be equal to the average price changes for products not subject to quality change.

1.248 It is necessary, therefore, to try to make an explicit adjustment for the change in quality, at least when a significant quality change is believed to have occurred. Again there are several methods that may be used.

Explicit quality adjustments

1.249 *Quantity adjustments*. The quality change may take the form of a change in the physical characteristics of the product that can easily be quantified, such as change in weight, dimensions, purity, or chemical composition of a product. It is generally a considerable oversimplification to assume that the quality of a product changes in proportion to the size of some single physical characteristic. For example, most consumers are very unlikely to rate a refrigerator that has three times the capacity of a smaller one as being worth three times the price of the latter. Nevertheless it is clearly possible to make some adjustment to the price of a new quality of different size to make it more comparable with the price of an old quality. There is considerable scope for the judicious, or common sense, application of relatively straightforward quality adjustments of this kind. A thorough discussion of quality adjustments based on "size" is given in Chapter 7.

1.250 *Differences in production or option costs*. An alternative procedure may be to try to measure the change in quality by the estimated change in the costs of producing the two qualities. The estimates can be made in consultation with the producers of the goods or services, if appropriate. This method, like the first, is only likely to be satisfactory when the changes take the form of relatively simple changes in the physical characteristics of the good, such as the addition of some new feature, or option, to an automobile. It is not satisfactory when a more fundamental change in the nature of the product occurs as a result of a new discovery or technological innovation. It is clearly inapplicable, for example, when a drug is replaced by another more effective variant of the same drug that also happens to cost less to produce.

1.251 Another possibility for dealing with a quality change that is more complex or subtle is to seek the advice of technical experts. This method is especially relevant when the general consumer may not have the knowledge or expertise to be able to assess or evaluate the significance of all of the changes that may have occurred, at least when they are first made.

1.252 *The hedonic approach*. Finally, it may be possible to systematize the approach based on production or option costs by using econometric methods to estimate the impact of observed changes in the characteristics of a product on its price. In this approach, the market prices of a set of different qualities or models are regressed on what are considered to be the most important physical or economic characteristics of the different models. This approach to the evaluation of quality change is known as *hedonic analysis*. When the characteristics are attributes that cannot be quantified, they are represented by dummy variables. The regression coefficients measure the estimated marginal effects of the various characteristics on the prices of the models and can therefore be used to evaluate the effects of changes in those characteristics, i.e., changes in quality, over time.

1.253 The hedonic approach to quality adjustment can provide a powerful, objective and scientific method of evaluating changes in quality for certain kinds of products. It has been particularly successful in dealing with computers. The economic theory underlying the hedonic approach is examined in more detail in Chapter 21. The application of the method is explained in some detail in Chapter 7. Products can be viewed as bundles of characteristics that are not individually priced, as the consumer buys the bundle as a single package. The objective is to try to "unbundle" the characteristics to estimate how much they contribute to the total price. In the case of computers, for example, three basic characteristics are the processor speed, the size of the RAM and the hard drive capacity. An example of a hedonic regression using these characteristics is given in Chapter 7.

1.254 The results obtained by applying hedonics to computer prices have had a considerable impact on attitudes towards the treatment of quality change in CPIs. They have demonstrated that for goods where there are rapid technological changes and improvements in quality, the size of the adjustments made to the market prices of the products to offset the changes in the quality can largely determine the movements of the elementary price index. For this reason, the manual contains a thorough treatment of the use of hedonics. Chapter 7 provides further analysis, including a comparison showing that the results obtained by using hedonics and matched models can differ significantly when there is a high turnover of models.

1.255 It may be concluded that statistical offices must pay close attention to the treatment of quality change and try to make explicit adjustments whenever possible. The importance of this topic can scarcely be over-emphasized. The need to recognize and adjust for changes in quality has to be impressed on price collectors. Failure to pay proper attention to quality changes can introduce serious biases into a CPI.

Item substitution and new goods

1.256 As noted above, ideally price indices would seek to measure pure price changes between matched products that are identical in the two periods compared. However, as explained in Chapter 8, the universe of products that a CPI has to cover is a dynamic universe that is gradually changing over time. Pricing matched products constrains the selection of products to a static universe of products given by the intersection of the two sets of products existing in the two periods compared.

This static universe, by definition, excludes both new products and disappearing products, whose price behaviour is likely to diverge from that of the matched products. Price indices have to try to take account of the price behaviour of new and disappearing products as far as possible.

1.257 A formal consideration and analysis of these problems are given in Appendix 8.1 to Chapter 8. A replacement universe is defined as one that starts with the base period universe but allows new products to enter as replacements as some products disappear. Of course, quality adjustments of the kind discussed above are needed when comparing the prices of the replacement products with those of the products that they replace.

1.258 One way in which to address the underlying problem of the changing universe is by sample rotation. This requires a completely new sample of products to be drawn to replace the existing one. The two samples must overlap in one period that acts as the link period. This procedure can be viewed as a systematic exploitation of the overlap method of adjusting for quality change. It may not therefore deal satisfactorily with all changes in quality that occur, because the relative prices of different goods and services at a single point of time may not provide satisfactory measures of the relative qualities of all the goods and services concerned. Nevertheless, frequent sample rotation helps by keeping the sample up to date and may reduce the extent to which explicit quality adjustments are required. Sample rotation is expensive, however.

New goods and services

1.259 The difference in quality between the original product and the one that it replaces may become so great that the new quality is better treated as a new good, although the distinction between a new quality and a new good is inevitably somewhat arbitrary. As noted in Chapter 8, a distinction is also drawn in the economics literature between evolutionary and revolutionary new goods. An evolutionary new good or service is one that meets existing needs in much more efficient or new ways, whereas a revolutionary new good or service provides completely new kinds of services or benefits. In practice, an evolutionary new good can be fitted into some sub-class of the product or expenditure classification, whereas a revolutionary new good will require some modification to the classification in order to accommodate it.

1.260 There are two main concerns with new goods or services. The first relates to the timing of the introduction of the new product into the index. The second relates to the fact that the mere availability of the new product on the market may bring a welfare gain to consumers, whatever the price at which it is sold initially. Consider, for example, the introduction of the first antibiotic drug, penicillin. The drug provided cures for conditions that previously might have been fatal. The benefit might be virtually priceless to some individuals. One way of gauging how much benefit is gained by the introduction of a new good is to ask how high its price would have to be to reduce the demand for the product to zero. Such a price is called the "demand reservation price". It could be very high indeed in the case of a new life-saving drug. If the demand reservation price could be estimated, it could be treated as the price in the period just before the new product appeared. The fall between the demand reservation price and the price at which the product actually makes its first appearance could be included in the CPI.

1.261 In practice, of course, statistical offices cannot be expected to estimate demand reservation prices with sufficient reliability for them to be included in a CPI. The concept is nevertheless useful because it highlights the fact that the mere introduction of a new good may bring a significant welfare gain that could be reflected in the CPI, especially if it is intended to be a COLI. In general, any enlargement of the set of consumption possibilities open to consumers has the potential to make them better off, other things being equal.

1.262 It is often the case that new goods enter the market at a higher price than can be sustained in the longer term, so their prices typically tend to fall relatively over the course of time. Conversely, the quantities purchased may be very small initially but increase significantly. These complications make the treatment of new products particularly difficult, especially if they are revolutionary new goods. Because of both the welfare gain from the introduction of a new product and the tendency for the price of a new good to fall after it has been introduced, it is possible that important price reductions may fail to be captured by CPIs because of the technical difficulties created by new products. Chapter 8 concludes by expressing concern about the capacity of CPIs to deal satisfactorily with the dynamics of modern markets. In any case, it is essential that statistical offices are alert to these issues and adopt procedures that take account of them to the maximum extent possible, given the data and resources available to them.

Calculation of consumer price indices in practice

1.263 Chapter 9 provides a general overview of the ways in which CPIs are calculated in practice. The methods used in different countries are by no means all the same, but they have much in common. There is clearly interest from users as well as compilers in knowing how most statistical offices set about calculating their CPIs. The various stages in the calculation process are illustrated by numerical examples. The chapter is descriptive and not prescriptive, although it does try to evaluate the strengths and weaknesses of existing methods. It makes the point that because of the greater insights into the properties and behaviour of indices gained in recent years, it is now recognized that not all existing practices are necessarily optimal.

1.264 As the various stages involved in the calculation process have, in effect, already been summarized in the preceding sections of this chapter, it is not proposed to repeat them all again in this section. It may be useful, however, to give an indication of the nature of the contents of Chapter 9.

Elementary price indices

1.265 Chapter 9 starts by describing how the elementary aggregates are constructed by working down from groups, classes and sub-classes of COICOP, or some equivalent expenditure classification. It reviews the principles underlying the delineation of the elementary aggregates themselves. Elementary aggregates are intended to be as homogeneous as possible, not merely in terms of the physical and economic characteristics of the products covered but also in terms of their price movements.

1.266 Chapter 9 then considers the consequences of using alternative elementary index formulae to calculate the elementary indices. It proceeds by means of a series of numerical examples that use simulated price data for four different products within an elementary aggregate. The elementary indices themselves, and their properties, have already been explained above. An elementary price index may be calculated either as a chain index or as a direct index; that is, either by comparing the price each month, or quarter, with that in the immediately preceding period or with the price in the fixed price reference period. Table 9.1 of Chapter 9 uses both approaches to illustrate the calculation of three basic types of elementary index, Carli, Dutot and Jevons. It is designed to highlight a number of their properties. For example, it shows the effects of "price bouncing" in which the same four prices are recorded for two consecutive months, but the prices are switched between the four products. The Dutot and Jevons indices record no increase but the Carli index registers an increase. Table 9.1 also illustrates the differences between the direct and the chain indices. After six months, each of the four prices is 10 per cent higher than at the start. Each of the three direct indices records a 10 per cent increase, as also do the chained Dutot and Jevons indices because they are transitive. The chained Carli, however, records an increase of 29 per cent, which is interpreted as illustrating the systematic upward bias in the Carli formula resulting from its failure to satisfy the time reversal test.

1.267 It is noted in Chapter 9 that the chaining and direct approaches have different implications when there are missing price observations, quality changes and replacements. The conclusion is that the use of a chain index can make the estimation of missing prices and the introduction of replacement items easier from a computational point of view.

1.268 Chapter 9 also examines the effects of missing price observations, distinguishing between those that are temporarily missing and those that have become permanently unavailable. Table 9.2 contains a numerical example of the treatment of the temporarily missing prices. One possibility is simply to omit the product whose price is missing for one month from the calculation of indices that compare that month with the preceding and following months, and also with the base period. Another possibility is to impute a price change on the basis of the average price for the remaining products, using one or other of the three types of average. The example is a simplified version of the kind of examples that are used in Chapter 7 to deal with the same problem.

1.269 Tables 9.3 and 9.4 illustrate the case in which one product disappears permanently to be replaced by another product. In Table 9.3 there is no overlap between the two products and the options considered are again to omit the products or to impute price changes for them based on averages for the other products. Table 9.4 illustrates the situation in which the products overlap in one month.

1.270 Chapter 9 also considers the possibility that there may be some expenditure weights available within an elementary aggregate, in which case it may be possible to calculate a Laspeyres or a geometric Laspeyres index, these being the weighted versions of the Carli and the Jevons.

Higher-level indices

1.271 Later sections of Chapter 9 illustrate the calculation of the higher-level indices using the elementary price indices and the weights provided by the elementary expenditure aggregates. It is at this stage that the traditional index number theory that was summarized earlier in this chapter and is explained in detail in Chapters 15 to 19 comes into play.

1.272 At the time the monthly CPI is first calculated, the only expenditure weights available must inevitably refer to some earlier period or periods of time. As explained earlier in this chapter, this predisposes the CPI to some form of Lowe or Young index in which the quantities, or expenditures, refer to some weight reference period b which precedes the price reference period 0. Such indices are often loosely described as Laspeyres type indices, but this description is inappropriate. At some later date, however, estimates may become available of the expenditures in both the price reference period 0 and the current period t, so that retrospectively the number of options open is greatly increased. It then becomes possible to calculate both Laspeyres and Paasche type indices, and also superlative indices such as Fisher or Törnqvist. There is some interest in calculating such indices later, if only to see how the original indices compare with the superlative indices. Some countries may wish to calculate retrospective superlative indices for that reason. Although most of the discussion in Chapter 9 focuses on some type of Lowe index because the official index first published will inevitably be of that type, this should not be interpreted as implying that such an index is the only possibility in the longer term.

1.273 *Production and maintenance of higher-level indices*. In practice, the higher-level indices up to and including the overall CPI are usually calculated as Young indices; that is, as weighted averages of the elementary price indices using weights derived from expenditures in some earlier weight reference period. This is a relatively straightforward operation, and a numerical example is given in Table 9.5 of Chapter 9 in which, for simplicity, the weight and price reference periods are assumed to be the same. Table 9.6 illustrates the case in which weight and price reference periods are not the same, and the weights are price updated between weight reference period b and the price reference period 0. It illustrates the point that statistical offices have two options when a new price reference period is introduced: they can either preserve the relative quantities of the weight reference period or they can preserve the relative expenditures, but they cannot do both. Price updating preserves the quantities.

1.274 The introduction of new weights is a necessary and integral part of the compilation of a CPI over the long run. Weights have to be updated sooner or later, some countries preferring to update their weights each year. Whenever the weights are changed, the index based on the new weights has to be linked to the index based on the old weights. Thus, the CPI inevitably becomes a chain index over the long term. An example of the linking is given in Table 9.7. Apart from the technicalities of the linking process, the introduction of new weights, especially if carried out at intervals of five years or so, provides an opportunity to undertake a major review of the whole methodology. New products may be introduced into the index, classifications may be revised and updated, while even the index number formula might be changed. Annual chaining facilitates the introduction of new products and other changes on a more regular basis, but in any case some ongoing maintenance of the index is needed whether it is annually chained or not.

1.275 Chapter 9 concludes with a section on data editing, a process that is very closely linked to the actual calculation of the elementary prices indices. Data editing comprises two steps: the detection of possible errors and outliers, and the verifying and correction of the data. Effective monitoring and quality control are needed to ensure the reliability of the basic price data fed into the calculation of the elementary prices indices, on which the quality of the overall index depends.

Organization and management

1.276 The collection of price data is a complex operation involving extensive fieldwork by a large number of individual collectors. The whole process requires careful planning and management to ensure that data collected conform to the requirements laid down by the central office with overall responsibility for the CPI. Appropriate management procedures are described in Chapter 12 of this manual.

1.277 Price collectors should be well trained to ensure that they understand the importance of selecting the right products for pricing. Inevitably, price collectors are bound to use their own discretion to a considerable extent. As already explained, one issue of crucial importance to the quality and reliability of a CPI is how to deal with the slowly evolving set of products with which a price collector is confronted. Products may disappear and have to be replaced by others, but it may also be appropriate to drop some products before they disappear altogether, if they have become unrepresentative. Price collectors need to be provided with appropriate training and very clear instructions and documentation about how to proceed. Clear instructions are also needed to ensure that price collectors collect the right prices when there are sales, special offers or other exceptional circumstances.

1.278 As just noted, the price data collected have also to be subjected to careful checking and editing. Many checks can be carried out by computer, using standard statistical control methods. It may also be useful to send out auditors to accompany price collectors and monitor their work. The various possible checks and controls are explained in detail in Chapter 12.

1.279 Improvements in information technology should obviously be exploited to the fullest extent possible. For example, collectors may use hand-held computers and transmit their results electronically to the central office.

Publication and dissemination

1.280 As noted above and in Chapter 2, the CPI is an extremely important statistic whose movements can influence the central bank's monetary policy, affect stock markets, influence wage rates and social security payments, and so on. There must be public confidence in its reliability, and in the competence and integrity of those responsible for its compilation. The methods used to compile it must therefore be fully documented, transparent and open to public scrutiny. Many countries have an official CPI advisory group consisting of both experts and users. The role of such a group is not just to advise the statistical office on technical matters but also to promote public confidence in the index.

1.281 Users of the index also attach great importance to having the index published as soon as possible after the end of each month or quarter, preferably within two or three weeks. There are also many users who do not wish the index to be revised once it has been published. Thus there is likely to be some trade-off between the timeliness and the quality of the index.

1.282 Publication should be understood to mean the dissemination of the results in any form. In addition to publication in print, or hard copy, the results should be released electronically and be available through the Internet on the web site of the statistical office.

1.283 As explained in Chapter 13, good publication policy goes beyond timeliness, confidence and transparency. The results should be made available to all users, in both the public and the private sectors, at the same time and according to a publication schedule announced in advance. There should be no discrimination among users in the timing of the release of the results. The results should not be subject to governmental scrutiny as a condition for their release, and should be seen to be free from political or other pressures.

1.284 There are many decisions to be taken about the degree of detail in the published data and the different ways in which the results may be presented. Users need to be consulted about these questions. These issues are discussed in Chapter 13. As they do not affect the actual calculation of the index, they need not be pursued further at this point.