Abbreviations

ABS	Antilock brake system; Australian Bureau of Statistics
AF	Acre foot
ANZIC	Australian and New Zealand Standard Industrial Classification
ATM	Automated teller machine
BPM5	*Balance of Payments Manual, Fifth Edition*
BEA	Bureau of Economic Analysis
BLS	U.S. Bureau of Labor Statistics
CAPI	Computer-assisted personal interviews
CATI	Computer-assisted telephone interviews
CD	Compact disk
CD-ROM	Compact disk-read-only memory
CD-RW	Compact disk-rewritable
c.i.f.	Cost-insurance-freight
CIR	Current Industrial Report
COFOG	Classification of the Functions of Government
COICOP	Classification of Individual Consumption by Purpose
COL	Cost of living
COPNI	Classification of the Purposes of Nonprofit Institutions Serving Households
COPP	Classification of the Purposes of Producers
CPA	Classification of Products by Activity, also known as PRODCOM (Eurostat)
CPC	Central Product Classification
CPI	Consumer price index
CSWD	Carruthers, Sellwood, Ward, Dalén price index
DRAM	Dynamic random-access memory
DRG	Diagnostic-Related Group
DRP	Disaster Recovery Plan
e-	Electronic (e-business, e-commerce, e-mail, etc.)
EC	European Commission
ECB	European Central Bank
ECI	Employment cost index
EDI	Electronic data interchange
EFQM	European Foundation for Quality Management
ESMR	Enhanced specialized mobile radio
EU	European Union
Eurostat	Statistical Office of the European Communities
FEPI	Final expenditure price index
FIOPI	Fixed-input output price index
FISIM	Financial Intermediation Services Implicitly Measured

f.o.b.	Free on board
FOIPI	Fixed-output input price index
FPI	Final uses price index
FPPI	Farm product price index
GB	Gigobytes
GDDS	General Data Dissemination System (IMF)
GDP	Gross domestic product
GPI	Global price index; government price index
GPS	Global positioning system
HBS	Household Budget Survey
HICPs	Harmonized Indices of Consumer Prices (Eurostat)
HP	Hodrick- Prescott; horsepower
HPI	Household consumption price index
HS	Harmonized Commodity Description and Coding System
ICP	Implicit characteristic price
ICPI	Intermediate consumption price index
IDI	Implicit deflator index
ILO	International Labour Office /International Labour Organization
IMF	International Monetary Fund
I/O	Input/output
IPP	International Price Program
ISIC	International Standard Industrial Classification of All Economic Activities
ISO	International Standards Organization
IT	Information technology
IWGPS	Inter-Secretariat Working Group on Price Statistics
KPI	Fixed capital formation price index
LIFO	Last in, first out
LKAU	Local kind of activity unit
LPG	Liquefied propane gas
MHz	Megahertz
MPI	Import price index
MSA	Metropolitan Statistical Area
NACE	General Industrial Classification of Economic Activities within the European Communities
NAFTA	North American Free Trade Association
NAICS	North American Industrial Classification System
1993 SNA	Commission of the European Communities (Eurostat), International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations, and World Bank, 1993, *System of National Accounts 1993* (Brussels/Luxembourg, New York, Paris, and Washington)
NPI	Inventory price index
NPISH	Nonprofit institution serving households
OECD	Organisation for Economic Co-operation and Development
OLS	Ordinary least squares
Ottawa Group	International Working Group on Price Indices

P_C	Carli price index
P_{CSWD}	Carruthers, Sellwood, Ward, and Dalén price index
P_D	Dutot price index
P_{DR}	Drobisch index
P_F	Fisher price index
P_{GL}	Geometric Laspeyres price index
P_{GP}	Geometric Paasche price index
P_H	Harmonic average of price relatives
P_{IT}	Implicit Törnqvist price index
P_J	Jevons price index
P_{JW}	Geometric Laspeyres price index (weighted Jevons index)
P_{KB}	Konüs and Byushgens price index
P_L	Laspeyres price index
P_{LM}	Lloyd-Moulton price index
P_{Lo}	Lowe price index
P_{ME}	Marshall-Edgeworth price index
P_P	Paasche price index
P_{RH}	Ratio of harmonic mean prices
P_T	Törnqvist price index
P_W	Walsh price index
P_Y	Young price index
PC	Personal computer
PCE	Personal consumption expenditures
PCS	Personal communications service
PDA	Personal digital assistant
PMC	Profit-maximizing center
PPI	Producer price index
PPP	Purchasing power parity
PPS	Probability proportional to size
PR	Price relative
PRODCOM	Product/commodity classification system for the European Community
RAM	Random-access memory
R_H	Ratio of harmonic average prices
RMSE	Root mean square error
ROSC	Reports on the Observance of Standards and Codes
rpm	Revolutions per minute
RSA	Residential Service Area
SAF	Seasonal adjustment factors
SDDS	Special Data Dissemination Standard (IMF)
SEHI	Superlative and exact hedonic indices
SIC	Standard Industrial Classification
SITC	Standard International Trade Classification
SMI	Supply markup index
SNA	*System of National Accounts*
SPI	Supply price index
SSR	Structured Schedule Review
SUT	Supply and use table
TEG-PPI	Technical Expert Group for the Producer Price Index

UN	United Nations
UNECE	UN Economic Commission for Europe
VAT	Value-added tax
Voorburg Group	International Working Group on Service Sector Statistics
VPI	Valuables price index
WD	Wheel drive
WLS	Weighted least squares
WPI	Wholesale price index
XPI	Export price index
YPI	Output price index

PART I

Methods, Uses, and Coverage

1. An Introduction to PPI Methodology

1.1 A price index is a measure of the proportionate, or percentage, changes in a set of prices over time. PPIs measure changes in the prices of domestic producer goods and services. Such measures need to distinguish between changes in the volume of domestic production and such changes in nominal terms. Because the prices of different goods and services do not all change at the same rate, a price index can reflect only their average movement. A price index typically assumes a value of unity, or 100, in some base period. The values of the index for other periods of time show the average proportionate, or percentage, change in prices from the base period. Price indices can also measure differences in price levels between different cities, regions, or countries at the same point of time.

1.2 Two basic questions are the focus of this *Manual* and the associated economic literature on price indices:

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

1.3 The answer to the first question depends largely on the purposes for which the index is to be used. Separate price indices can be compiled for different flows of goods and services, such as household production, government production, investment, or foreign trade flows. *Output* PPIs, which measure changes in the prices of goods and services produced by *businesses*, are the primary concern of this *Manual*. However, businesses do not all sell the same set of goods and services. Thus, there can be more than one output PPI depending on the particular set of goods and services selected. As well as considering the problems involved in measuring output prices, this *Manual* will also consider the problems associated with constructing *input* PPIs, used for deflating the value of intermediate inputs used in production. An intermediate input is an input that is used by one establishment or production unit but is the output of another establishment. Of interest to economists is deflating changes in value added over time, and weighted averages of the differences between output and intermediate input price indices, *value-added* PPIs, may ideally serve this purpose.

1.4 Once the appropriate set of prices (and, if weights are available, related quantities and revenue information) are collected, the second question concerns the choice of formula to average the price movements. Two standard methods are available to measure sectoral and overall price changes over time: compile an average of price changes or compile a ratio of average prices. This is summarized below and considered in detail in Chapters 15–20.

1.5 This chapter provides a general introduction to, and review of, the methods of PPI compilation. It provides a summary of the relevant theory and practice of index number compilation that helps reading and understanding the detailed chapters that follow, some of which are inevitably quite technical. The chapter describes the various steps involved in PPI compilation, starting with the basic concepts, definitions, and purposes of PPIs. It then discusses the sampling procedures and survey methods used to collect and process the price data, and finishes with the eventual calculation and dissemination of the final index.

1.6 In an introductory presentation of PPI methods of the kind given in this chapter, it is necessary to start with the basic concept of a PPI and the underlying index number theory. This includes the properties and behavior of the various kinds of index numbers that might be used for PPI purposes. Only after deciding the type of index and its coverage based on these theoretical considerations is it possible to go on to determine the best way in which to estimate the index in practice, taking account of the resources available. As noted in the Reader's Guide, however, the detailed presentation of the relevant index theory appears in later chapters of the *Manual* because the theory can become technically complex when pursued in some depth.

The exposition in this chapter does not therefore follow the same order as the chapters in the *Manual*.

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

- The uses and origins of PPIs;
- Basic index number theory, including the axiomatic and economic approaches to PPIs;
- Elementary price indices and aggregate PPIs;
- The transactions, activities, and establishments covered by PPIs;
- The collection and processing of the prices, including adjusting for quality change;
- The actual calculation of the PPI;
- Potential errors and bias;
- Organization, management, and dissemination policy; and
- An appendix providing an overview of the steps necessary for developing a PPI.

1.8 Not all of the topics treated in the *Manual* are included in this chapter. The objective of this general introduction is to provide a summary presentation of the core issues with which readers need to be acquainted before tackling the detailed chapters that follow. It is not the purpose of this introduction to provide a comprehensive summary of the entire contents of the *Manual*. Some special topics, such as the treatment of certain products whose prices cannot be directly observed, are not considered here because they do not affect general PPI methodology.

A. The Uses and Origins of PPIs

1.9 Four of the principal price indices in the system of economic statistics—the PPI, the CPI, and the export and import price indices—are well known and closely watched indicators of macroeconomic performance. They are direct indicators of the purchasing power of money in various types of transactions and other flows involving goods and services. As such, they are also used to deflate nominal measures of goods and services produced, consumed, and traded to provide measures of volumes. Consequently, these indices are important tools in the design and conduct of the monetary and fiscal policy of the government, but they are also of great utility in informing economic decisions throughout the private sector. They do not, or should not, comprise merely a collection of unrelated price indicators but provide instead an integrated and consistent view of price developments pertaining to production, consumption, and international transactions in goods and services.

1.10 In the system of price statistics, PPIs serve multiple purposes. The precise way in which they are defined and constructed can very much depend on by whom and for what they are meant to be used. PPIs can be described as indices designed to measure the average change in the price of goods and services either as they leave the place of production or as they enter the production process. A monthly or quarterly PPI with detailed product and industry data allows monitoring of short-term price inflation for different types or through different stages of production. Although PPIs are an important economic indicator in their own right, a vital use of PPIs is as a deflator of nominal values of output or intermediate consumption for the compilation of production volumes and for the deflation of nominal values of capital expenditure and inventory data for use in the preparation of national accounts.[1]

1.11 Beyond their use as inflationary indicators or as deflators, certain frameworks for PPIs provide insight into the interlinkages between different price measures. One such framework is aggregation of stage-of-processing indices. This concept classifies goods and services according to their position in the chain of production—that is, primary products, intermediate goods, and finished goods. This method allows analysts to track price inflation through the economy. For example, changes in prices in the primary stage could feed through into the later stages, so the method gives an indicator of future inflation further down the production chain. However, each product is allocated to only one stage in the production chain even though it could occur in several stages. This topic will be considered in Chapter 2 and again in Chapter 14.

1.12 A further method for analysis is to aggregate by stage of production, in which each product is allocated to the stage in which it is used. This differs from stage of processing because a product is included in each stage to which it contributes and is not assigned solely to one stage. The classification of products to the different stages is usually

[1]PPIs are used for this purpose because the volumes underlying the nominal values are not directly measurable.

achieved by reference to input-output tables, and, in order to avoid multiple counting, the stages are not aggregated. There is a growing interest in this type of analysis. For example, these types of indices are already compiled on a regular basis in Australia.[2] This topic will also be considered in Chapters 2 and 14.

1.13 As explained in Chapter 2, PPIs have their beginnings in the development of the wholesale price index (WPI) dating back to the late 19th century. Laspeyres and Paasche indices, which are still widely used today, were first proposed in the mid-19th century. They are explained below. The concepts of the fixed-input *output* price index and the fixed-output *input* price index were introduced in the mid- to late 20th century. These two concepts provide the basic framework for the economic theory of the PPI presented in Chapters 15 and 17.

1.14 Initially, one of the main reasons for compiling a WPI was to measure price changes for goods sold in primary markets before they reached the final stage of production at the retail market level. Thus the WPI was intended to be a general purpose index to measure the price level in markets other than retail. The WPI has been replaced in most countries by PPIs because of the broader coverage provided by the PPI in terms of products and industries and the conceptual concordance between the PPI and the *System of National Accounts*, discussed in more detail in Chapter 14. It is this concordance that makes components of the PPI useful as deflators for industrial outputs and product inputs in the national accounts. In addition, the overall PPI and PPIs for specific products are used to adjust prices of inputs in long-term purchase and sales contracts, a procedure known as "escalation."

1.15 These varied uses often increase the demand for PPI data. For example, using the PPI as an indicator of general inflation creates pressure to extend its coverage to include more industries and products. While many countries initially develop a PPI to cover industrial goods produced in mining and manufacturing industries, the PPI can logically be extended to cover all economic activities, as noted in Chapters 2 and 14.

[2]See, for example, Australian Bureau of Statistics (ABS) (2003 and other years); available via the Internet: www.abs.gov.au.

B. Some Basic Index Number Formulas

1.16 The first question is to decide on the kind of index number to use. The extensive list of references given at the end of this *Manual* reflects the large literature on this subject. Many different mathematical formulas have been proposed over the past two centuries. Nevertheless, there is now a broad consensus among economists and compilers of PPIs about what is the most appropriate type of formula to use, at least in principle. While the consensus has not settled for a single formula, it has narrowed to a very small class of *superlative* indices. A characteristic feature of these indices is that they treat the prices and quantities in both periods being compared symmetrically. They tend to yield very similar results and behave in very similar ways.

1.17 However, when a monthly or quarterly PPI is first published, it is invariably the case that there is not sufficient information on the quantities and revenues in the current period to make it possible to calculate a symmetric, or superlative, index. It is necessary to resort to second-best alternatives in practice, but in order to be able to make a rational choice between the various possibilities, it is necessary to have 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.

1.18 The *Manual* provides a comprehensive, thorough, rigorous, and up-to-date discussion of relevant index number theory. Several chapters from Chapter 15 onward are devoted to a detailed explanation of index number theory from both a statistical and an economic perspective. The main points are summarized in the following sections. Many propositions or theorems are stated without proof in this chapter because the proofs are given or referenced in later chapters to which the reader can easily refer in order to obtain full explanations and a deeper understanding of the points made. There are numerous cross-references to the relevant sections in later chapters.

B.1 Price indices based on baskets of goods and services

1.19 The purpose of an index number may be explained by comparing the *values* of producer's revenues from the production of goods and services in two time periods. Knowing that revenues have increased by 5 percent is not very informative if we do not know how much of this change is due to changes in the *prices* of the goods and services and how much to changes in the *quantities* produced. *The purpose of an index number is to decompose proportionate or percentage changes in value aggregates into their overall price and quantity change components.* A PPI is intended to measure the price component of the change in producer's revenues. One way to do this is to measure the change in the value of an aggregate by holding the quantities constant.

B.1.1 Lowe indices

1.20 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 producing a fixed 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 Section B.2 of Chapter 15. Most statistical offices make use of some kind of Lowe index in practice. It is described in some detail in Sections D.1 and D.2 of Chapter 15.

1.21 In principle, any set of goods and services could serve as the basket. The basket does *not* have to be restricted to the basket actually produced in one or other of the two periods compared. For practical reasons, the basket of quantities used for PPI purposes usually has to be based on a survey of establishment revenues conducted in an earlier period than either of the two periods whose prices are compared. For example, a monthly PPI may run from January 2000 onward, with January 2000 = 100 as its price reference period, but the quantities may be derived from an annual revenue survey made in 1997 or 1998, or even spanning both years. Because it takes a long time to collect and process revenue data, there is usually a considerable time lag before such data can be introduced into the calculation of PPIs. The basket may also refer to a year, whereas the index may be compiled monthly or quarterly.

1.22 Let there be n products in the basket with prices p_i and quantities q_i. Let period b be the period to which the quantities refer and periods 0 and t be the two periods whose prices are being compared. In practice, it is invariably the case that $b \leq 0 < t$ when the index is first published, and this is assumed here. However, b could be any period, including one between 0 and t, if the index is calculated some time after t. The Lowe index is defined in equation (1.1).

$$(1.1)\ P_{Lo} \equiv \frac{\sum_{i=1}^{n} p_i^t q_i^b}{\sum_{i=1}^{n} p_i^0 q_i^b} \equiv \sum_{i=1}^{n} \left(p_i^t / p_i^0 \right) s_i^{0b} ,$$

$$\text{where} \quad s_i^{0b} = \frac{p_i^0 q_i^b}{\sum_{i=1}^{n} p_i^0 q_i^b} .$$

The Lowe 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 revenue shares s_i^{0b} as weights. They 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 revenue 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 relative between b and 0, namely p_i^0 / p_i^b. The concept of the *base period* is somewhat ambiguous with a Lowe index, since either b or 0 might be interpreted as being the base period. To avoid ambiguity, b is described as the *weight reference period* and 0 as the *price reference period.*

1.23 Lowe indices are widely used for PPI purposes.

B.1.2 Laspeyres and Paasche indices

1.24 Any set of quantities could be used in a Lowe index, but there are two special cases that figure prominently in the literature and are of considerable importance from a theoretical point of view. When the quantities are those of the first of the two periods whose prices are being compared—

that is, when b = 0—the *Laspeyres* index is obtained, and when quantities are those of the second 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.

1.25 The formula for the Laspeyres price index, P_L, is given in equation (1.2).

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

where s_i^0 denotes the share of the value of product i in the total output of goods and services in period 0: that is, $p_i^0 q_i^0 / \sum p_i^0 q_i^0$.

1.26 As can be seen from equation (1.2), and as explained in more detail in Chapter 15, the Laspeyres index can be expressed in two alternative ways that are algebraically identical: first, as the ratio of the values of the basket of producer goods and services produced in period 0 when valued at the prices of periods t and 0, respectively; second, as a weighted arithmetic average of the ratios of the individual prices in periods t and 0 using the value shares in period 0 as weights. The individual price ratios, (p_i^t/p_i^0), are described as price *relatives*. Statistical offices often calculate PPIs using the second formula by recording the percentage changes in the prices of producer goods and services sold and weighting them by the total value of output in the base period 0.

1.27 The formula for the Paasche index, P_P, is given in equation (1.3).

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

where s_i^t denotes the actual share of the expenditure on commodity i in period t: that is, $p_i^t q_i^t / \sum p_i^t q_i^t$. The Paasche index can also be expressed in two alternative ways, either as the ratio of two value aggregates or as a weighted average of the price relatives, the average being a *harmonic* average that uses the revenue shares of the later period t as weights. However, 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.

1.28 If the objective is simply to measure the price change between the two periods considered in isolation, there is no reason to prefer the basket of the earlier period to that of the later period, or vice versa. Both baskets are equally relevant. Both indices are equally justifiable, or acceptable, from a conceptual point of view. In practice, however, PPIs are calculated for a succession of time periods. A time series of monthly Laspeyres PPIs based on period 0 benefits from requiring only a single set of quantities (or revenues), those of period 0, so that *only the prices* have to be collected on a regular monthly basis. A time series of Paasche PPIs, on the other hand, requires data on *both* prices *and* quantities (or revenues) in each successive period. Thus, it is much less costly, and time consuming, to calculate a time series of Laspeyres indices than a time series of Paasche indices. This is a *decisive practical* advantage of Laspeyres (as well as Lowe) indices over Paasche indices and explains why Laspeyres and Lowe indices are used much more extensively than Paasche indices. A monthly Laspeyres or Lowe PPI can be published as soon as the price information has been collected and processed, since the base-period weights are already available.

B.1.3 Decomposing current-value changes using Laspeyres and Paasche indices

1.29 Laspeyres and Paasche quantity indices are defined in a similar way to the price indices, simply by interchanging the ps and qs in formulas (1.2) and (1.3). They summarize changes over time in the flow of quantities of goods and services produced. 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 revenues 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_L and Q_L denote the Laspeyres price and quantity indices and let P_P and Q_P denote the Paasche price and quantity indices. As shown in Chapter 15, $P_L \times Q_P \equiv V$ and $P_P \times Q_L \equiv V$.

1.30 Suppose, for example, a time series of industry output in the national accounts is to be deflated to measure changes in output at constant prices over time. If it is desired to generate a series of output values at constant base-period prices (whose movements are identical with those of the Laspeyres volume index), the output at current prices must be deflated by a series of Paasche price indices. Laspeyres-type PPIs would not be appropriate for the purpose.

B.1.4 Ratios of Lowe and Laspeyres indices

1.31 The Lowe index is transitive. The ratio of two Lowe indices using the same set of q^bs 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:

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

1.32 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.

1.33 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,

$$(1.5)\quad P_{Lo} = \frac{\sum_{i=1}^{n} p_i^t q_i^b}{\sum_{i=1}^{n} p_i^0 q_i^b} = \frac{\sum_{i=1}^{n} p_i^t q_i^b \Big/ \sum_{i=1}^{n} p_i^b q_i^b}{\sum_{i=1}^{n} p_i^0 q_i^b \Big/ \sum_{i=1}^{n} p_i^b q_i^b} = \frac{P_L^t}{P_L^0} .$$

B.1.5 Updated Lowe indices

1.34 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 the Lowe index for period $t + 1$ with price reference period t. Thus,

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

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

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

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

1.35 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 revenue 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 revenue 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.

B.1.6 Relationships between fixed-basket indices

1.36 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. Conversely, if the weighted price and quantity changes are *positively* correlated, then the Paasche index exceeds the Laspeyres. The proof is given in Appendix 15.1 of Chapter 15.

1.37 This has different implication for consumers and producers. The theory of consumer behavior indicates that consumers typically react to price changes by substituting goods or services that have become *relatively* cheaper for those that have become *relatively* dearer. Thus they purchase smaller quantities of the higher-priced products and more of lower-priced ones. This is known as the *substitution effect*, and it implies a negative correlation between the price and quantity relatives. In this case the Laspeyres CPI would be greater than the Paasche CPI, with the gap between them tending to widen over time.[3] That the Laspeyres tends to rise faster than the Paasche is a matter of concern to many analysts and CPI users because it suggests that the widely used Laspeyres index may have an upward bias.

1.38 The theory of the firm indicates the opposite behavior on the part of producers. As prices for particular products begin to rise, producers will shift production away from lower-priced, less profitable products toward the higher-priced, more profitable ones. This type of substitution by producers implies a positive correlation between price and quantity relatives. In this case the Paasche PPI would be greater than the Laspeyres PPI, with the gap between them widening over time. That the Paasche tends to rise faster than the Laspeyres is a matter of concern to many analysts and PPI users because it suggests that the widely used Laspeyres index may have a downward bias, a point taken up later.

1.39 In practice, however, statistical offices often do not calculate Laspeyres or Paasche indices but instead 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 Section D.1 of Chapter 15 that *if there are persistent long-term trends in relative prices and if the substitution effect for purchasers is dominant, the Lowe index will tend to exceed the Laspeyres, and therefore also the Fisher and the Paasche.* Assuming that the time period b is prior to the time period 0, the ranking under these conditions will be:

Lowe $\geq$ Laspeyres $\geq$ Fisher $\geq$ 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*.

1.40 The positioning of period b is crucial. Given the assumptions about long-term price trends and substitution, a Lowe index will tend to increase (decrease) as period b is moved backward (forward) 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 the passage of time. Period b can then be moved forward. If b is positioned midway between 0 and t, the quantities are likely to be equirepresentative 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 Section D.2 of Chapter 15.

1.41 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 PPI. However, the amount by which such a PPI exceeds some conceptually preferred target index, such as the economic index discussed in Section E below, 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 an upward bias.[4] A large bias may undermine the credibility and acceptability of the index.

[3]If the revenue shares—that is, the weights associated with the price relatives—happen to be the same in both periods, the Laspeyres must be greater than the Paasche because a weighted arithmetic average is always greater than a harmonic average with the same weights. In order to maintain the revenue shares intact, the substitution of the quantities in response to changes in relative prices must be perfect.

[4]Of course, if producers are price takers from the market and the demand shifts dominate, then producers will respond by increasing the quantities produced of goods with higher relative prices. The correlation between prices and quantities in this instance will be positive, and the relationship among the indices will be:

Paasche $\geq$ Fisher $\geq$ Laspeyres $\geq$ Lowe,

and the bias interpreted as downward.

B.2 The Young index

1.42 Instead of holding constant the quantities of period b, a statistical office may calculate a PPI 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 an another index number pioneer. The Young index is defined in Section D.3 of Chapter 15 as follows:

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

1.43 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 usually is more current than the weight reference period b because of the time needed to collect and process and the revenue data. In that case, a statistical office has the choice of assuming that *either* the quantities of period b remain constant *or* the revenue shares in period b remain constant. Both cannot remain constant if prices change between b and 0. If the revenue shares actually remained constant between periods b and 0, the quantities would have had to change inversely in response to the price changes. In this case the elasticity of substitution is 1; for example, the proportionate decline in quantity is equal to the proportionate increase in prices.

1.44 Section D.3 of Chapter 15 shows that the Young index is equal to the Laspeyres index plus the covariance between the difference in annual shares pertaining to year b and month 0 shares $\left(s_i^b - s_i^0\right)$ and the deviations in relative prices from their means $(r - r_i^*)$. Normally, the weight reference period b precedes the price reference period 0. In this case, if the elasticity of substitution is larger than 1—for example, the proportionate decline in quantity is greater than the proportionate increase in prices—the covariance will be positive. Under these circumstances the Young index will exceed the Laspeyres index.[5] Alternatively, if the elasticity of substitution is less than 1, the covariance will be negative and the Young will be less than the Laspeyres.

1.45 As explained later, the Young index fails some critical index number tests discussed in Section C of this chapter and in Chapter 16, Section C.

B.2.1 Geometric Young, Laspeyres, and Paasche indices

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

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

where s_i^b is defined as above. The geometric Laspeyres is the special case in which $b = 0$: that is, the revenue shares are those of the price reference period 0. Similarly, the geometric Paasche uses the revenue shares of period t. Note 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.

1.47 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 revenue shares, the geometric Laspeyres and Paasche indices coincide. In this case, the ranking of the indices must be:

> ordinary Laspeyres ≥ geometric Laspeyres and Paasche ≥ the ordinary Paasche.

1.48 The indices are, respectively, arithmetic, geometric, and harmonic averages of the same price relatives that all use the same set of weights.

1.49 The geometric Young and Laspeyres indices have the same information requirements as their

[5]This occurs because products with the large relative price increases ($r - r_i$* is positive) would also experience declining shares between periods b and 0 ($s_i^b - s_i^0$ is positive), thus having a positive influence on the covariance. In addition, products with small relative price increases ($r - r_i$* is negative) would experience increasing shares between b and 0 ($s_i^b - s_i^0$ is negative), thus having a positive influence on the covariance.

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 PPI 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.

B.3 Symmetric indices

1.50 When the base and current periods are far apart, the index number spread between the numerical values of a Laspeyres and a Paasche price index is liable to be quite large, especially if *relative* prices have changed a lot (as shown Appendix 15.1 and illustrated numerically in Chapter 19). Index number spread is a matter of concern to users because, conceptually, there is no good reason to prefer the weights of one period to those of the other. In these circumstances, it seems reasonable to take some kind of *symmetric average* of the two indices. More generally, it seems intuitive to prefer indices that treat both of the periods symmetrically instead of relying exclusively on the weights of only one of the periods. It will be shown later that this intuition can be backed up by theoretical arguments. There are many possible symmetric indices, but there are three in particular that command much support and are widely used.

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

$$(1.10)\quad P_F \equiv \sqrt{P_L \times P_P}\ .$$

1.52 The second is the *Walsh price index, P_W*, a pure price index in which the quantity weights are *geometric* averages of the quantities in the two periods; that is

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

The averages of the quantities need to be *geometric* rather than arithmetic for the *relative* quantities in both periods to be given equal weight.

1.53 The third index is the *Törnqvist price index, P_T*, defined as a *geometric* average of the price relatives weighted by the average *revenue* shares in the two periods:

$$(1.12)\quad P_T = \prod_{i=1}^{n} \left(p_i^t / p_i^0\right)^{\sigma_i}\ ,$$

where σ_i is the arithmetic average of the share of revenue on product i in the two periods, and

$$(1.13)\quad \sigma_i = \frac{s_i^t + s_i^0}{2}\ ,$$

where the s_i s are defined as in equation (1.2) and above.

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

B.4 Fixed-base versus chain indices

B.4.1 Fixed-basket indices

1.55 This topic is examined in Section F of 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 whose prices are being compared. The base period whose quantities are used has to be updated sooner or later, and the new index series linked to the old. Linking is inevitable in the long run.

1.56 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 $j = -1$, [that is, $t - (-1) = t + 1$], it becomes a chain Paasche.

1.57 The PPIs in some countries are, in fact, annual chain Lowe indices of this general type, the quantities referring to some year, or years, that precedes 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 revenues for 1998;
- The 12 indices from January 2001 to January 2002 are then based on price updated revenues for 1999; and so on with annual weight updates.

The revenues 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 later, using current revenue 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.

1.58 A chain index between two periods 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 spreads among the Lowe, Laspeyres, and Paasche indices, thereby making the movements in the index less dependent on the choice of index number formula.

1.59 However, if there are fluctuations in the prices and quantities in the intervening periods, 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 does not return to 100. It will 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 the 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 analyzed in Chapter 22. While a number of countries update their revenue weights annually, the 12 monthly indices within each year are not chain indices but Lowe indices using fixed annual quantities.

B.4.2 The Divisia index

1.60 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 pioneered by Divisia. As shown in Section E of Chapter 15, the Divisia index may be derived mathematically by differentiating value (that is, price times quantity) with respect to time to obtain two components: a relative value-weighted price change and relative value-weighted quantity change. These two components are defined to be price and quantity indices, respectively. The Divisia index 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 index. The Divisia index itself offers no practical guidance about the kind of index number formula to choose for the individual links in a chain index.

C. The Axiomatic Approach to Index Numbers

1.61 The *axiomatic approach* to index numbers is explained in Chapter 16. It seeks to decide the most appropriate formula for an index by specifying a number of axioms, or tests, that the index ought to satisfy. It throws light on the properties possessed by different kinds of indices, some of which are by no means intuitively obvious. Indices that fail to satisfy certain basic or fundamental axioms may be rejected completely because they are liable to behave in unacceptable ways. The axiomatic approach is also used to rank indices on the basis of their desirable, and undesirable, properties.

1.62 Twenty axioms or tests (T) are initially considered in Chapter 16. Only a selection of them are 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; that is, the price index function is (positively) homogeneous of degree 1 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.

- T12—*Quantity Reversal Test*: If the quantity vectors for the two periods are interchanged, then the price index remains invariant.

- 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.

1.63 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 milk is measured in liters instead of pints, the index must be unchanged. One index that does not satisfy this test is the ratio of the arithmetic means of the prices in the two periods (the *Dutot* index). This is a type of elementary index that is widely used in the early stages of PPI calculation. This is discussed in more detail in Chapter 20, Sections C and F.

1.64 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. Because 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. 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. However, when the prices refer to a strictly homogeneous set of products that all use the same unit of measurement, the test becomes irrelevant. In practice, products may differ in terms of their quality characteristics, and there is a sense in which this variation in quality is similar to variation in the units of measurement. While the quality of individual products may not change, the price changes of the higher-price varieties of, say, types of pepper, when aggregated, will be given more emphasis in the calculation.

1.65 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 forward in time, from 0 to t, or backward 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. This follows from the general algebraic result that the reciprocal of the arithmetic average of a set of numbers is the *harmonic* average of the reciprocals, not the arithmetic average of the reciprocals. The fact that the *conceptually* arbitrary decision to measure the change in prices forward from 0 and t gives a different result from measuring backward 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.

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

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

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

1.67 On the other hand, 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. However, the quantities of a Laspeyres index are those of the price reference period, *by definition*, and *must change* whenever the price reference period is changed. The basket for a forward Laspeyres is different from that for the backward Laspeyres, and the Laspeyres fails the time reversal test as a consequence.

1.68 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

$$P_{Lo}{}^{0,\,t} = P_{Lo}{}^{0,\,t-k} \cdot P_{Lo}{}^{t-k,\,t}$$

where $P_{Lo}{}^{0,\,t}$ 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 $t–k$.

1.69 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.

1.70 In reality, 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.

C.1 Ranking indices using the axiomatic approach

1.71 In Section B.6 of Chapter 16, it is shown not only that the Fisher price index satisfies all the 20 axioms initially listed in the chapter but also, more remarkably, that it is the *only* possible index that can satisfy all 20 axioms. Thus, on the basis of this set of axioms, the Fisher clearly dominates other indices.

1.72 In contrast to Fisher, the other two symmetric indices defined in equations (1.11) and (1.12) above do not emerge too well from the 20 tests. In Section B.7 of Chapter 16, it is shown that the Walsh price index fails four tests, whereas the Törnqvist index fails nine tests. Although the Törnqvist index does not perform well on these tests, especially compared with Fisher, it should be remembered that the Törnqvist index and Fisher index may, nevertheless, be expected to approximate each other quite closely when the data follow relatively smooth trends, as shown in Chapter 19.

1.73 The Lowe index with fixed quantities emerges quite well from the axiomatic approach. In particular, in contrast to the Laspeyres, Paasche, and Young indices, it satisfies the time reversal test. As already explained, however, the attractiveness of the Lowe index depends very much on the positioning of period b that supplies the quantity weights, rather than its axiomatic properties.

1.74 One limitation of the axiomatic approach is that the list of axioms itself is inevitably arbitrary to some extent. Some axioms, such as the Paasche and Laspeyres bounding test failed by both Törnqvist and Walsh, could be regarded as contrived and dispensable. In particular, many of the test properties have an arithmetic basis, whereas the Törnqvist index is a geometric average. 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. Badly failing 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.

C.1.2 Some further tests

1.75 Consider a further symmetry test. It is reasonable that reversing the roles of prices and quantities in a price index should yield a quantity index of the same formula as the price index. A formula that is good enough for a price index should be equally good for a quantity index. The *factor reversal test* requires that the product of such a quantity index and the original price index should be identical with the change in the value of the aggregate in question. This test is important if, as stated at the outset of this chapter, 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 remarkable result derived from the axiomatic approach, and given in Section B.6 of Chapter 16, is that the Fisher index is the only price index to satisfy four minimal tests: T1 (positivity), T11 (time reversal), T12 (quantity reversal), and T21 (factor reversal).[6] Because 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.*

1.76 It was shown earlier that 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. Because Laspeyres and Paasche have different functional forms, this implies that they fail the factor reversal test. However, 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 or Paasche price index does lead to a meaningful quantity index, even though its formula is not identical with that of the price index.

1.77 Another test, discussed in Section C.8 of Chapter 16, is the *additivity test*. A good property for an index is that the changes in the subaggregates add up to the changes in the totals. This is more important from the perspective of quantity indices than it is for price indices. Price indices may be used to deflate value changes to obtain implicit quantity changes. The results may be presented for subaggregates such as output by industry or product groups. Just as output aggregates at current prices are, by definition, obtained simply by summing individual output values or revenues, it is reasonable to expect that the changes in the subaggregates of a quantity index should add up to the changes in the totals—the additivity test. Quantity indices that use a common set of prices to value quantities in both periods must satisfy the additivity test. Similarly, if the Lowe quantity index is defined as $\sum p^j q^t / \sum p^j q^0$ it is also additive. The Geary-Khamis quantity index used to make international comparisons of real consumption and 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 $\boldsymbol{p}^j$ to compare the quantities in different countries.

1.78 An alternative solution is to use some *average* of the prices in two periods to value the quantities. 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.[7]

1.79 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.

---

[6]See Funke and Voeller (1978, p. 180).

[7]Additivity is a property that is attractive in a national accounts context, where many aggregates are actually defined by processes of addition and subtraction. It is also useful when comparing national accounts data for different countries using purchasing power parities (PPPs), a type of international price index. (See *CPI Manual*, International Labour Organization and others [2004, Annex 4].)

D. The Stochastic Approach

1.80 The stochastic approach treats the observed price relatives as if they were a random sample drawn from a defined universe whose mean can be interpreted as the general rate of inflation. However, there can be no single unique rate of inflation. There are many possible universes that can be defined, depending on which particular sets of industries, products, or transactions the user is interested in. Clearly, the sample mean depends on the choice of universe from which the sample is drawn. The stochastic approach does not help decide on the choice of universe. It 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.

1.81 The stochastic approach becomes particularly useful when the universe is reduced to a single type of product. When there are market imperfections, there may be considerable variation within a country in the prices at which a single product is sold in different establishments and also in their movements over time. 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 5 on sampling issues and Chapter 20 on elementary indices. The main points are summarized in Section I below.

D.1 The unweighted stochastic approach

1.82 In Section C.2 of Chapter 16, the unweighted stochastic approach to index number theory is explained. If simple random sampling has been used to collect prices, 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 estimators, 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 can be regarded as the unweighted version of the Young index. This index is discussed further in Section I below on elementary price indices.

1.83 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 later, this fact may be decisive when deciding on the formula to be used to estimate the elementary indices compiled in the early stages of PPI calculations.

D.2 The weighted stochastic approach

1.84 As explained in Section F of Chapter 16, a *weighted* stochastic approach can be applied at an aggregative level covering sets of different products. Because 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 output, 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), with the probability of any individual sampled product being selected being proportional to the output of that type of product in some period or periods. Different indices are obtained depending on which revenue weights are used and whether the price relatives or their logarithms are used.

1.85 Suppose a sample of price relatives is randomly selected, with the probability of selecting any particular type of product being proportional to the revenue of that type of product in period 0. The expected price change is then the Laspeyres price index for the universe. However, other indices may 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 revenue 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. From an axiomatic viewpoint, the choice of a symmetric average of the revenue 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.

1.86 The examples of the Laspeyres and Törnqvist indices just given show that the stochastic approach in itself does not determine the form of the index number. There are several stochastic indices to choose from, just as there are many possible universes. However, as already noted, the elementary prices from which most aggregate price indices are constructed usually have to be based on samples of prices, and the stochastic approach may provide useful guidance on how best to estimate them.

E. The Economic Approach

1.87 The economic approach differs from the previous approaches in an important respect: quantities are no longer assumed to be independent of prices. If, for example, it is assumed that firms behave as revenue maximizers, it follows that they would produce more of products with above-average price changes in, say, period 1 compared with period 0. As a result, the revenue shares in period 1 from such products will increase, and therefore, their weights. This behavioral assumption about the firm, as it switches production to higher-priced products, allows something to be said about what "true" indices should be and the suitability of different index number formulas. For example, the Laspeyres index uses fixed period-0 revenue shares to weight its price relatives and ignores the substitution of production toward products with higher relative price changes in period 1. It will thus understate aggregate price changes—be biased downward against its true index. The Paasche index uses fixed period 1 weights and ignores the initial revenue shares in period 0. It will thus overstate aggregate price changes—be biased upward against its true index.

1.88 The economic approach can be seen to be very powerful, since it has identified a type of bias in Laspeyres and Paasche indices not apparent from other approaches: *substitution bias*. Laspeyres and Paasche indices ignore the change in weights as producers substitute their production toward products with above-average price increases. Yet the nature of the bias arises from an assumption about the behavior of producers—that they are revenue maximizers. Consider an alternative assumption: that producers respond to demand changes prompted by purchasers buying less of products with relatively high price changes. Products whose price increases are, for example, above average will see a falloff in demand leading to a falloff in production. In this case the revenue shares or weights of products with above-average price increases will fall in period 1, and the fixed period-0 weighted Laspeyres will overstate aggregate price changes—it will be upward biased. This compares with the Paasche index, which will understate aggregate price changes—it will be downward biased. It is shown in Chapter 17 that Laspeyres and Paasche indices can under certain conditions act as bounds on a more generally applicable "true" economic theoretic index. The axiomatic approach in Section C led to an index number formula that used an average of the Laspeyres and Paasche indices, and, even at this early stage in the discussion, the economic approach seems to provide further support.

1.89 The economic approach also identifies the circumstances under which the conventionally used Laspeyres index is appropriate. This would require that the firm does not change its production configuration in response to relative price changes, at least over the short term of the price index comparisons. Economic theory thus argues that the Laspeyres index may be appropriate for industries in which quantities are known not to respond to relative price changes over the period of the price comparisons. But it is more likely that this will be the exception rather than the norm, and the theory points to a requirement for a more generally applicable index number formula.

1.90 The PPI indices considered here include output, input, and value-added price indices (deflators), and different assumptions arise in their formulation from economic theory. In the output case, an assumption is made that firms act to maximize revenues, from a given input base. Firms substitute toward products with relatively high price increases. For the input price index, the concern is to minimize the costs of purchased intermediate goods. Firms substitute away from input products with relatively high price increases. For the value-added deflator, the unusual use of negative weights for the inputs is considered. The economic approach, as shown in Chapter 17, demonstrates that:

- A substitution bias can exist when using Laspeyres and Paasche formulas.
- The nature of the bias depends on the behavioral assumptions of the firm, which will vary between industries and the type of PPI index required—input or output PPI.
- Laspeyres and Paasche indices act as bounds on their true indices and, under certain conditions,

also are bounds for a more generally applicable true index.

It follows that some symmetric average of these bounds is justified from economic theory.

1.91 The approach from economic theory is thus first to develop theoretical index number formulas based on what are considered to be reasonable models of economic behavior by the producer. This approach is very different from the others considered here. A mathematical representation of the production activity—whereby capital and labor conjoin to turn intermediate inputs into outputs—is required. Also, an assumption of optimizing behavior (cost minimization or revenue maximization), along with other assumptions, is required so that a theoretical index can be derived that is "true" under these conditions. The economic approach then examines practical index number formulas such as Laspeyres, Fisher, and Törnqvist, and considers how they compare with "true" formulas defined under different assumptions. Three theoretical formulations will be examined—each, in principle, requiring different assumptions about the optimizing behavior of the firm. None can be practically calculated (for reasons that will be explained). The first approach to an economic theoretical producer price index is the concept of the fixed-input *output price* index. This index is a ratio of hypothetical revenues over the two periods being compared, say periods 0 and 1, that the revenue-maximizing establishment could realize, where the technology and inputs to work with were fixed to be the same for both of the periods. An establishment that, for example, doubles its revenue using a fixed technology and inputs, effectively doubles its prices. The theoretical index is a ratio of revenues, so it incorporates substitution effects as more revenue is obtained as firms substitute toward higher-priced products. The theoretical index wishes to have as its period 1 quantities the results of the firm changing the mix of output it produces in response to relative price changes. But there is a dilemma: only price changes should be reflected, and by allowing quantities to change in this way pure price changes would not be measured. So the theoretical index fixes the amount that can be produced by holding the technology and inputs at some constant level. The firm can change its output mix but must use constant inputs and technology. Note that there is an entire *family* of theoretical price indices depending on which period's reference technology and inputs are held constant: fixed period-0 technology and primary inputs, fixed period-1 technology and primary inputs, or some average of the two.

1.92 Theoretical fixed-output *input price* indices may also be defined. These are the ratio of hypothetical intermediate input costs that the cost-minimizing establishment must pay in order to produce a set of outputs, again with technology and primary inputs fixed to be the same for the comparison in both periods.

1.93 The measurement of GDP using the production approach involves calculating the *value added* by the industry. Value added is the difference between the value of output produced by industries and the value of the intermediate inputs used. The value added by each industry is then summed along with taxes less subsidies on products to provide an estimate of GDP. An important use of the PPI is to deflate the values of outputs and inputs at current-period prices to estimate value added at constant prices. In Chapter 17 the economic approach is first used to define a theoretical *output price index*, *intermediate input price index*, and *value-added deflator* for *a single establishment.* Aggregation is then undertaken over establishments in order to define national counterparts to these establishment price indices in Chapter 18.

E.1 Theoretical output price indices

1.94 The theoretical *output price index* between periods 0 and 1 is the ratio of the maximum revenues that the establishment could attain when faced with period 0 and 1 prices using a fixed, given technology and a fixed set of inputs. Consider a theoretical index in which period 0 technology and inputs are held constant, the theoretical counterpart to the Laspeyres index. What is required for the numerator of the ratio is to generate what the period 1 quantities would be, holding the production process and inputs constant in period 0 after the change in relative prices from the period 0 technology and inputs. This in turn requires a mechanism to generate these hypothetical period 1 quantities from the fixed period-0 technology and inputs. In the economic approach the technology of a firm or industry is described in terms of a production (possibility) function, which tells us the maximum amount of output(s) that can be produced from a given set of inputs. If the values of all the inputs to a firm or industry were given, the production function would be able to generate all possible combinations of output of products from the technology—it would be a mathematical representation of the technology that converts inputs to outputs. The prevailing relative prices would dictate exactly how much of each product is produced. The economic approach to the PPI relies on the assumption of *optimizing behavior* on the part of producers in competitive, price-taking markets so that they respond to relative price changes. In this approach, while actual prices are considered for both periods, the quantities in each period may not be the observed ones. They are generated from a given period's production function (with fixed technology) and level of inputs, using assumptions of maximizing behavior and dictated by relative prices, which may be the ones in another period. This is a powerful analytical framework because it allows us to consider, at least in theory, how quantities would respond to different price regimes (say, period 1 prices) under constant (say, period 0) reference technologies and inputs. They are hypothetical quantities that cannot be observed but are generated in a mathematical model so that their formulation can be compared with real index number formulas based on observable prices and quantities.

1.95 "Pure" price index number formulas (based on observed data) and theoretical indices have in common that they may both be defined as the ratios of revenues in two periods. However, by definition, while the quantities are fixed in pure price indices, they vary in response to changes in relative prices in theoretical indices. In contrast to the axiomatic approach to index theory, the economic approach recognizes that the quantities produced are actually dependent on the prices. In practice, rational producers may be expected to adjust the *relative* quantities they produce in response to changes in *relative* prices. A theoretical PPI assumes that a producer seeking to maximize revenues will make the necessary adjustments. The baskets of goods and services in the numerator and denominator of a theoretical PPI are not, therefore, exactly the same.

E.2 Upper and lower bounds on a theoretical output price index

1.96 The theoretical price index between periods 0 and 1 is the ratio of revenues in those periods using fixed technology and inputs. Consider an index that held the technology and inputs constant in period 0. The revenue generated in period 0 from period 0 prices using period 0 technology and inputs is what actually happened: the denominator of the theoretical ratio is the observed revenue, assuming the producer was optimizing revenue. The numerator is period 1 prices multiplied by the hypothetical quantities that would have been produced using the same period 0 technology and inputs, had period 1 prices prevailed. It is **not**, as in the Laspeyres index, period 1 prices multiplied by the actual quantities produced at period 0 prices using period 0 technology and inputs. Both the theoretical and the Laspeyres indices use the same period 0 technology and inputs, but the theoretical index generates quantities from it as if period 1 prices prevailed, whereas the Laspeyres index uses the actual period 0 quantities. In practice, relative prices may change between the two periods, so the quantities generated will be different. Higher revenue could be achieved by substituting, at least marginally, some products that have relatively high price changes for some that have relatively low ones. The theoretical index based on period 0 technology and inputs will take account of this and will increase by more than the Laspeyres index. The theoretical index will be at least equal to or greater than the Laspeyres, since the producer has the possibility of, at worst, producing the same set of products as in period 0. Being a revenue maximizer, it is assumed the producer will substitute products with relatively high price changes—the Laspeyres index thus incurs a "substitution bias."

1.97 By a similar line of reasoning, it can be shown that when relative prices change, the theoretical output price index based on period 1 technology and inputs will increase by less than the Paasche index. In other words, as shown in Chapter 17, Section B.1, the Laspeyres index provides a lower bound to its (period 0) theoretical index, and the Paasche an upper bound to its (period 1) theoretical index. Note that these inequalities are in the opposite direction to their CPI cost-of-living index counterparts. This is because the optimization problem in the cost-of-living theory is a cost minimization problem as opposed to the present revenue *maximization* problem.

1.98 The practical significance of these results stems from the fact that the Laspeyres and Paasche indices can be calculated directly from the observed prices and quantities, whereas the theoretical indices cannot, thus giving some insight into the bias involved in the use of these two formulas. Suppose the official objective is to estimate a base-period theoretical output price index, but that a Laspeyres index is calculated instead for practical reasons. One important conclusion to be drawn from this preliminary analysis is that the PPI may be expected to have a downward bias. Similarly, a series of Paasche PPIs used to deflate a series of output values at current prices generates a series of values at constant period 0 prices (Laspeyres volume index), which in turn will also suffer from a downward bias. The approach informs us that there are *two* equally valid theoretical economic price indices, and that the bound, while useful, shows only how Laspeyres and Paasche indices compare with their own theoretical counterparts. What we require are *two-sided* bounds on the theoretically justified index.

E.3 Estimating theoretical output indices by superlative indices

1.99 The next step is to establish whether there are special conditions under which it may be possible to exactly measure a theoretical PPI. In Section B.2 of Chapter 17 theoretical indices based on weighted "averages" of the period 0 and period 1 technology and similarly weighted averages of the period 0 and 1 inputs are considered. These theoretical indices deal adequately with *substitution effects*; that is, when an output price increases, the producer's supply increases, holding inputs and the technology constant. Such theoretical indices are argued to generally fall between the Laspeyres (lower bound) and Paasche (upper bound) indices. The Fisher index, as the geometric mean of the Laspeyres and Paasche indices, is the only symmetric average of Laspeyres and Paasche that satisfies the *time reversal test.* Thus, economic theory was used to justify Laspeyres and Paasche bounds, and axiomatic principles led to the Fisher price index as the best symmetric average of these bounds.

1.100 In Section B.3 of Chapter 17 the case for the Törnqvist index number formula is presented. It is assumed that the revenue function takes a specific mathematical form: a translogarithmic function. If the price coefficients of this translog form are equal across the two periods being compared, then the geometric mean of the economic output price index that uses period 0 technology and the period 0 input vector, and the economic output price index that uses period 1 technology and the period 1 input vector, are *exactly equal* to the Törnqvist output price index. The assumptions required for this result are weaker than other subsequent assumptions; in particular, there is no requirement that the technologies exhibit constant returns to scale in either period. The ability to relate an actual index number formula (Törnqvist) to a specific functional form (translog) for the production technology is a powerful analytical device. Statisticians using particular index number formulas are in fact replicating particular mathematical descriptions of production technologies. A good formula should not correspond to a restrictive functional form for the production technology.

1.101 Diewert (1976) described an index number formula to be *superlative* if it is equal to a theoretical price index whose functional form is flexible—it can approximate an arbitrary technology to the second order. That is, the technology by which inputs are converted into output quantities and revenues is described in a manner that is likely to be realistic of a wide range of forms. Relating a class of index number formulas to technologies represented by flexible functional forms is another powerful finding, since it gives credence to this class of index number formulas. Note also that the translog functional form is an example of a *flexible* functional form, so the Törnqvist output price index number formula is *superlative.* In contrast to the theoretical indices, a superlative index is an actual index number that can be calculated. The practical significance of these results is that they give a theoretical justification for expecting a superlative index to provide a fairly close approximation to the unknown, underlying theoretical index in a wide range of circumstances.

1.102 In Section B.4 the Fisher index is revisited from a purely economic approach. An additional assumption is invoked, that outputs are homogeneously separable from other commodities in the production function: if the input quantities vary, the output quantities vary with them, so that the new output quantities are a uniform expansion of the old output quantities. It is shown that a homogeneous quadratic utility function is flexible and corresponds to the Fisher index. The Fisher output price index is therefore also *superlative*. This is one of the more famous results in index number theory. Although it is generally agreed that it is not plausible to assume that a production technology would have this particular functional form, this result does at least suggest that, in general, the Fisher index is likely to provide a close approximation to the underlying unknown theoretical PPI—and certainly a much closer approximation than either the Laspeyres or the Paasche indices can yield on their own.

1.103 This intuition 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 that is exactly equal to a theoretical one based on the underlying aggregator function as *superlative* when that functional form is also flexible—for example, a homogeneous quadratic. The derivation of these results, and further explanation, are given in detail in Section B.3 of Chapter 17. In contrast to the theoretical index itself, 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 unknown underlying theoretical index in a wide range of circumstances.

E.3.1 Superlative indices as symmetric indices

1.104 The Fisher index is not the only example of a superlative index. In fact, there is a whole family of superlative indices. It is shown in Section B.4 of 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:

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

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

1.105 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 revenue shares in both periods symmetrically whatever value is assigned to the parameter r. Three special cases are of interest:

- When r = 2, equation (1.15) reduces to the Fisher price index;
- When r = 1, it is equivalent to the Walsh price index;
- In the limit as $r \to 0$, it equals the Törnqvist index.

1.106 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.

E.3.2 The choice of superlative index

1.107 Section B.5.2 of Chapter 17 addresses the question of which superlative formula to choose in practice. Because each may be expected to approximate to the same underlying theoretical output index, it may be inferred that they ought also to approximate to each other. That they are all symmetric indices reinforces this conclusion. These conjectures tend to be borne out in practice by numerical calculations. It seems that the numerical values of the different superlative indices tend to be very close to each other, but only so long as the value of the parameter r does not lie far outside the range 0 to 2. However, in principle, there is no limit on the value of the parameter r, and in Section B.5.1 of Chapter 17, it is shown that as the value of r becomes progressively 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 very small, as in the case of the three commonly used superlative indices—Fisher, Walsh, and Törnqvist—is the choice of superlative index unimportant.

1.108 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 (1922) himself was instrumental in developing. As shown above, it appears to dominate other indices from an axiomatic viewpoint. That it is also a superlative index whose use can be justified on grounds of economic theory suggests that, from a theoretical point of view, it may be impossible to improve on the Fisher index for PPI purposes.

1.109 However, the Walsh index has the attraction of being not merely a superlative index, but also a conceptually simple *pure* price index based on a fixed basket of goods and services. That the Walsh index is both a superlative and a pure index throws light on the interrelationships between the theoretical output price index and pure price indices. The distinctive feature of a Walsh index is not just that the basket of goods and services is a simple (geometric) average of the quantities in each of the two periods; by being a geometric average, it also assigns equal importance to the *relative*, as distinct from the absolute, quantities. Such an index clearly treats both periods symmetrically.[8] Pure price indices do not have to diverge from the theoretical output price index and are not inherently biased as estimators of the theoretical index. Bias is likely to arise only when the relative quantities used in a pure price index favor one of the periods at the expense of the other, as in a Laspeyres or Paasche index.

E.3.3 Representativity bias

1.110 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.

1.111 Because 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.[9] If equal importance is attached to the production 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.

1.112 Suppose that period b, whose 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 the most representative basket. Conversely, the farther 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 farther 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 farther 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 economic index. Thus, certain kinds of indices can be regarded as biased without invoking the concept of an economic index. Conversely, the same kinds of indices that tend to emerge as being preferred, whether or not the objective is to estimate an economic index.

1.113 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 for the target economic index 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.

E.3.4 Data requirements and calculation issues

1.114 Because superlative indices require price and revenue data for both periods and revenue data are usually not available for the current period, it is not feasible to calculate a superlative PPI, at least at the time that a PPI is first published. In practice, it may be necessary for the official index to be a Laspeyres-type index. However, in the course of time more revenue data may become available, enabling a superlative PPI to be calculated subsequently. Some statistical offices may find it useful to do so, without necessarily revising the original

[8]The Marshall-Edgeworth index (see Chapter 15) uses a simple arithmetic average of the quantities, but the resulting basket will be dominated by the quantities for one or other of the periods if the quantities are larger, on average, in one period than the other. The Marshall-Edgeworth is not a superlative index.

[9]The Walsh basket is the one that minimizes the sum of the squares of the logarithmic deviations between the quantities in the two actual baskets and those in the index basket.

official index. Comparing movements in the official PPI with those in a subsequently calculated superlative version may be helpful in evaluating and interpreting movements in the official PPI. It may be that revenue data can be collected from establishments alongside price data, and this is to be encouraged so that Fisher PPI indices may be calculated in real time for at least some industrial sectors. To the extent that revenue data are available on an annual basis, annual chain Laspeyres indices could be produced initially and Fisher or Törnqvist indices produced subsequently as the new revenue weights become available. The advantage of the annual updating is that chaining helps to reduce the spread between the Laspeyres and Paasche indices.

1.115 Section B.7 of Chapter 17 notes that, in practice, PPIs are usually calculated in stages (see 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 superlative indices are not. However, the widely used Fisher and Törnqvist indices are shown to be approximately consistent.

E.4 Allowing for substitution

1.116 Section B.8 of Chapter 17 examines one further index proposed recently, the Lloyd-Moulton index, P_{LM}, defined as follows:

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

The parameter σ, which must be nonpositive for the output PPI, 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, except for an estimate of the parameter σ, than the Laspeyres index. It is therefore a practical possibility for PPI calculation, even for the most recent periods. However, 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.

E.5 Intermediate input price indices and value-added deflators

1.117 Having considered the theory and appropriate formula for *output* price indices, Chapter 17 turns to *intermediate input* price indices (Section C) and to *value-added* deflators (Section D). The behavioral assumption behind the theory of the output price index was one of producers maximizing a revenue function. An input price index is concerned with the price changes of intermediate inputs, and the corresponding behavioral assumption is the minimization of a conditional cost function. The producer is held to minimize the cost of intermediate inputs in order to produce a set of outputs, given a set of intermediate inputs prices and that primary inputs and technology are fixed. These are fixed so that hypothetical input quantities can be generated from a fixed setup that allows the input quantities in period 1 to reflect the producer buying more of those inputs that have become cheaper. Theoretical intermediate input price indices are defined as ratios of hypothetical intermediate input costs that the cost-minimizing producer must pay in order to produce a fixed set of outputs from technology and primary inputs fixed to be the same for the comparison in both periods. As was the case with the theory of the output price index, theoretical input indices can be derived on the basis of either fixed period-0 technology and primary inputs, or fixed period-1 technology and primary inputs, or some average of the two. The observable Laspeyres index of intermediate input prices is shown to be an *upper bound* to the theoretical intermediate input price index based on period-0 technology and inputs. The observable Paasche index of intermediate input prices is a *lower bound* to its theoretical intermediate input price index based on period-1 fixed technology and inputs. Note that these inequalities are the reverse of the findings for the output price index but that they are analogous to their counterparts in the CPI for the theory of the true cost-of-living index, which is also based on an expenditure (cost) minimization problem.

1.118 Following the analysis for the output price index, a family of intermediate input price indices can be shown to exist based on an *average* of period-0 and period-1 technologies and inputs leading to the result that Laspeyres (upper) and Paasche (lower) indices are bounds on a reasonable theoretical input index. A symmetric mean of the two bounds is argued to be applicable given that

Laspeyres and Paasche indices are equally justifiable, with the Fisher index having support on axiomatic grounds. If the conditional intermediate input cost function takes the form of a translog technology, the theoretical intermediate input price index is exactly given by a Törnqvist index, which is superlative. If separability is invoked, Fisher and Walsh indices are also shown to be superlative, and the three indices closely approximate each other.

1.119 The third index is the *value-added* deflator. The analysis is based on the maximization of a net revenue function, a function that relates output revenue less intermediate input costs to sets of output prices, input prices, and given primary inputs and technology. The results follow those from using a revenue function for the output price index. Laspeyres and Paasche indices are lower and upper bounds on their respective theoretical value-added deflators, and a family of theoretical value-added deflators could be defined that lie between them. The Fisher index again has some support as a symmetric average on axiomatic grounds, although the Törnqvist index is shown, using fairly weak assumptions, to correspond to a flexible translog functional form for the net revenue function and is, therefore, superlative. This finding requires no assumption of the more restrictive constant returns to scale that are necessary for Fisher and Walsh indices, analogous to those for the output price index.

F. Aggregation Issues

1.120 It has been assumed up to this point that the theoretical PPI is based on the technology of a single representative establishment. Chapter 18 examines the extent to which the various conclusions reached above remain valid for PPIs that are actually compiled for industries or the overall economy. The general conclusion is that essentially the same relationships hold at an aggregate level, although some additional issues arise that may require additional assumptions.

1.121 That there are three possible PPIs requires an examination of how they relate to each other. It is thus necessary to consider how the value-added deflator is related to the output price and the intermediate input price indices, and how the output price index and the intermediate input price index can be combined in order to obtain a value-added deflator. It is shown in Chapter 18 that when the Laspeyres output price index is used to separately deflate outputs and the Laspeyres input price index is used to separately deflate inputs—*double deflation*—at each stage of aggregation, the results are the same as when Laspeyres is used to aggregate in one single stage. The separate deflation of inputs by the input price index and outputs by the output price index make up the components of the double-deflated value-added index. The same applies for the Paasche index. However, if superlative price indices are used, there are some small inconsistencies. It was noted previously that unlike superlative indices, Laspeyres and Paasche indices may suffer from serious substitution bias. They may add up, but not to the right number. A value-added deflator equivalent to the separate Laspeyres (Paasche) deflation of output and input indices is shown as a weighted "average" of the Laspeyres (Paasche) output price index and the Laspeyres (Paasche) intermediate input price index, although the weights used to combine the input and output deflators are rather unusual.

1.122 But how do we derive estimates of double-deflated value added? There is an equivalence between a number of methods. Using the product rule, a value ratio divided (deflated) by a Laspeyres value-added deflator generates a Paasche value-added quantity index; or, correspondingly, a value ratio divided by a Paasche value-added deflator generates a Laspeyres value-added quantity index. An alternative approach yielding equivalent results is to take value added in, say, period 0 at period-0 prices and escalate (multiply) it by a series of Laspeyres value-added quantity indices. The resulting series of value added at constant period-0 prices will be identical to the results from separately escalating the value of inputs and outputs by their respective Laspeyres input and output quantity indices and subtracting the (escalated) former from the latter. More usually, estimates of value added at constant prices are derived by deflation. Deflating a series of nominal current-period value added by a series of Paasche value-added indices yields a series of value added at constant prices. This is equivalent to double deflation: the separate deflation of the inputs and output current-period values by their respective input and output separate Paasche price indices, subtracting the former from the latter. Similar equivalence results can be found using the less well-known approach for a comparison between periods 0 and 1 of deflating the period-1 values by a Paasche *quantity* index to provide a measure of current period-0 quantities at period-1 prices. These can be compared with the nominal period-1 value at current prices to provide, for bilateral comparisons, an estimate of quantity change at constant period-1 prices. These results were devised for the establishment, and it is also shown in Chapter 18 that they hold on aggregation for Laspeyres and Paasche indices and fairly closely for the three main superlative indices: Fisher, Törnqvist, and Walsh.

G. Illustrative Numerical Data

1.123 Chapter 19 presents 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 formulas can yield very different numerical results. Hypothetical but economically plausible prices, quantities, and revenues are given for six products over five periods of time. In general, differences between the different formulas 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.

1.124 The numerical results are striking. For example, the Laspeyres index over the five periods registers an increase of 44 percent, whereas the Paasche falls by 20 percent. The two commonly used superlative indices, Törnqvist and Fisher, register increases of 25 percent and 19 percent, 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 register increases of 33 percent and 12 percent, respectively, reducing the gap between the two indices from 64 to 21 points. The chained Törnqvist and Fisher register increases of 22.26 percent and 22.24 percent, respectively, being virtually identical numerically. These results show that the choice of index formula and method does matter.

H. Choice of Index Formula

1.125 By drawing on the index number theory surveyed in Chapters 15–19 it is possible to decide on the type of index number in any given set of circumstances. However, there is little point in asking what is the best index number formula for a PPI. The question is too vague. A precise answer requires a precise question. For example, suppose that the principal concern of most users of PPIs is to have the best measure of the *current rate* of factory gate inflation. The precise question can then be posed: what is the best index number to use to measure the change between periods $t-1$ and t in the prices of the producer goods and services leaving the factory between periods $t-1$ and t?

1.126 The question itself determines both the coverage of the index and the system of weighting. The establishments in question have to be those of the country in question and not, say, those of some foreign country. Similarly, the question refers to establishments in periods $t-1$ and t, not to establishments five or ten years earlier. Sets of establishments five or ten years apart are not all the same, and their inputs and production technologies change over time.

1.127 Because the question specifies goods and services produced in periods $t-1$ and t, the basket of goods and services used should include *all* the quantities produced by the establishments in periods $t-1$ and t, and *only* those quantities. One index that meets these requirements is a pure price index that uses a basket consisting of the total quantities produced in both periods $t-1$ and t. This is equivalent to an index that uses a simple arithmetic mean of the quantities in the two periods, an index known as the Marshall-Edgeworth index. However, this index has a slight disadvantage in that if domestic production is growing, the index gives rather more weight to the quantities produced in period t than those in $t-1$. It does not treat both periods symmetrically. It fails tests T7 and T8 listed in Chapter 16 on the axiomatic approach, the invariance to proportional changes in quantities tests. However, if the arithmetic mean quantities are replaced by the geometric mean quantities, as in the Walsh index, both tests are satisfied. This ensures that the index attaches equal importance to the *patterns* of production, as measured by *relative* quantities produced in both $t-1$ and t.

1.128 The Walsh index therefore emerges as the pure price index that meets all the requirements. It takes account of every single product produced in the two periods. It utilizes all the quantities produced in both periods, and only those quantities. It gives equal weight to the patterns of production in both periods. In practice, it may not be feasible to calculate a Walsh index, but it can be used as the standard by which to evaluate other indices.

1.129 The index theory developed in Chapters 15–17 demonstrates that the Fisher and the Törnqvist indices are equally good alternatives. Indeed, the Fisher may be preferred to the Walsh on axiomatic grounds, given that the two indices will tend to give almost identical results for comparisons between successive time periods.

1.130 As already noted, for practical reasons the PPI is often calculated as a time series of Laspeyres indices based on some earlier period 0. In this case, the published index between $t - 1$ and t may actually be the monthly-change version of the Laspeyres index given in equation (1.4) above. Given that some substitution effect is operative, which seems extremely likely on both theoretic and empirical grounds, it may be inferred, by reasoning along lines explained in Chapter 15, that the monthly-change Laspeyres index will tend to be less than the Walsh index between $t - 1$ and t. If the PPI is intended to measure producer inflation, therefore, the monthly-change Laspeyres could have a downward bias, a bias that will tend to get worse as the current period for the Laspeyres index moves further away from the base period. This is the kind of conclusion that emerges from the index theory presented in Chapters 15 and 16. It is a conclusion with considerable policy and financial implications. It also has practical implications because it provides an argument for rebasing and updating a Laspeyres index as often as resources permit, perhaps on an annual basis as many countries are now doing.

1.131 If the objective of the PPI is to measure the current rate of change in revenues for a fixed, given technology and set of inputs, to be used for output deflation, this translates into asking what is the best estimate of the change in producer output prices. The theory elaborated in Chapter 17 shows that the best estimate will be provided by a superlative index. The three commonly used superlative indices are Fisher, Törnqvist, and Walsh. One or the other of these indices emerges as the theoretically most appropriate formula, whether the objective is to measure the current rate of factory gate inflation or as a deflator. A monthly-change Laspeyres is likely to have the same bias whatever the objective.

1.132 If the objective were to measure price changes over long periods of time—say, 10 or 20 years—the main issue for long-term comparisons is whether to chain or not, or at least how frequently to link.

I. Elementary Price Indices

1.133 As explained in Chapters 9 and 20, the calculation of a PPI typically proceeds in two or more stages. In the first stage, *elementary price indices* are estimated for the *elementary aggregates* of a PPI. In the second stage, these elementary indices are combined to obtain higher-level indices using the elementary aggregate indices with revenue weights. An elementary aggregate consists of the revenue for a small and relatively homogeneous set of products defined within the industrial classification used in the PPI. Samples of prices are collected within each elementary aggregate, so that elementary aggregates serve as strata for sampling purposes.

1.134 Data on the revenues, or quantities, of the different goods and services may not be available within an elementary aggregate. Since it has been shown that it is theoretically appropriate to use superlative formulas, data on revenues should be collected alongside those on prices whenever possible. Given that this may not possible, that there are no quantity or revenue weights, most of the index number theory outlined in the previous sections is not applicable. An elementary price index is a more primitive concept that relies on price data only. It is something calculated when there is no explicit or implicit quantity or revenue data available for weights. Implicit quantity or revenue data may arise from a sampling design whereby the selection of products is with probability proportionate to quantities or sales revenue.

1.135 The question of what is the most appropriate formula to use to estimate an elementary price index is considered in Chapter 20. This topic 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. Since the elementary indices are the building blocks from which PPIs are constructed, the quality of a PPI depends heavily on them.

1.136 As explained in Chapter 6, compilers have to select *representative products* within an elementary aggregate and then collect a sample of prices for each of the representative products, usually from a sample of different establishments. The individual products whose prices are actually collected are described as the *sampled products*. Their prices are collected over a succession of time periods. An elementary price index is therefore typically calculated from two sets of matched price observations. It is assumed in this section 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 that is discussed in detail in Chapters 7, 8, and 21 of the *Manual*.

I.1 Heterogeneity of products within an elementary aggregate

1.137 If a number of different representative products are selected for pricing, the set of products within an elementary aggregate cannot be homogeneous. Even a single representative product may not be completely homogeneous, depending upon how tightly it is specified. This topic is considered in more detail in Chapters 5–7. The degree of heterogeneity of the sampled products must be explicitly taken into account in the calculation of an elementary index.

1.138 When the quantities are not homogeneous, *they cannot be meaningfully added from an economic viewpoint, and their prices should not be averaged*. Consider again the example of salt and pepper, which might be representative products within an elementary aggregate. Pepper is an expensive spice sold in very small quantities such as ounces or grams, whereas salt is relatively cheap and sold in much larger quantities, such as pounds or kilos. A simple arithmetic average of, say, the price of a gram of pepper and the price of a kilo of salt is an arbitrary statistic whose value depends largely on the choice of the quantity units. Choosing the same physical unit of quantity, such as a kilo, for both does not resolve the problem, because both the average price and the change in the average price would be completely dominated by the more expensive product, pepper, even though producers may obtain more revenue from salt. In general, arithmetic averages of prices should be taken only when the corresponding quantities are homogeneous and can be meaningfully added.

I.2 Weighting

1.139 As already noted, it is assumed in this section that there are no quantities or revenues available to weight the prices, or the price relatives, used to calculate an elementary index. If they were available, it would usually be preferable to use them to decompose the elementary aggregate into smaller and more homogeneous aggregates.

1.140 However, some system of weighting may have been implicitly introduced into the selection of the sampled products by the sample design used. For example, the establishments from which the prices are collected may have been selected using probabilities of selection that are proportional to their sales or some other variable.

I.3 Relationships between different elementary index formulas

1.141 Valuable insights into the properties of various formulas that might be used for elementary price indices may be gained by examining the numerical relationships between them, as explained in Section D of Chapter 20. There are two basic options for an elementary index:

- To average the price relatives—that is, the ratios of the matched prices;
- To calculate the ratio of average prices in each period.

1.142 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. Using these three types of average, the ranking of the results obtained by the second method are predictable. It should also be noted that the ratio of geometric averages is identical with the geometric average of the ratios. The two methods give the same results when geometric averages are used.

1.143 As explained in Section C of Chapter 20, there are several elementary price indices that might possibly be used. Using the first of the above options, three possible elementary price indices are:

- The arithmetic average of the price relatives, known as the *Carli* index, or P_C; the Carli is the unweighted version of the Young index.
- The 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.

- The harmonic average of the price relatives, or P_H.

As just noted, $P_C \geq P_J \geq P_H$.

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

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

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

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

$$(1.17) \quad P_D \equiv \frac{\sum_{i=1}^{n} p_i^t \Big/ n}{\sum_{i=1}^{n} p_i^0 \Big/ n} = \frac{\sum_{i=1}^{n} p_i^0 \left(\dfrac{p_i^t}{p_i^0} \right)}{\sum_{i=1}^{n} p_i^0} .$$

As compared with the Carli, which is a simple average of the price relatives, the Dutot index gives more weight to the price relatives for the products with high prices in period 0. However, it is difficult to provide an economic rationale for this kind of weighting. Prices are not revenues. 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, since the quantities are not commensurate and not additive.

1.146 Noting that $P_C \geq P_J \geq P_H$, it is shown in Section D of Chapter 20 that the gaps between these 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. Moreover, both P_D and P_J can be *expected* to lie *approximately* halfway between P_C and P_H. While it is useful to establish the interrelationships between the various indices, they do not actually help decide which index to choose. However, because the differences between the various formulas 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. Since 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.

I.4 The axiomatic approach to elementary indices

1.147 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 Section E of Chapter 20.

1.148 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 can be critical, however. It reflects the point made earlier that when the quantities are not economically commensurate, their prices should not be 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 index should not be used.

1.149 The Carli index, P_C, is widely used, but the axiomatic approach shows that it has some undesirable properties. In particular, as the unweighted version of the Young index, it fails the commodity reversal, the time reversal, and the transitivity tests. These are serious disadvantages, especially when month-to-month indices are chained. 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 (HICPs) used within the European Union. Conversely, the harmonic average of the price relatives, P_H, is liable to have an equally significant downward bias, although it does not seem to be used in practice anyway.

1.150 On the axiomatic approach, the Jevons index, P_J, emerges as the preferred index. However, 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, and it may be necessary to impose upper and lower bounds on the individual price relatives when using the Jevons.

I.5 The economic approach to elementary indices

1.151 The economic approach, explained in Section F of Chapter 20, seeks to take account of the economic behavior of producers and their economic circumstances. Price differences may be observed at the same point of time for two quite different reasons:

- Exactly the same product may be sold by different categories of producers at different prices.
- The sampled products are not exactly the same. The different prices reflect differences in quality.

Both phenomena may occur at the same time.

1.152 *Pure* price differences can occur when the products sold at different prices are exactly the same. Pure price differences imply differing technologies or market imperfections of some kind, such as local monopolies, price discrimination, consumer or producer ignorance, or rationing. If all consumers had equal access, were well informed, and were free to choose, and all producers produced using the same technologies in price-taking markets, all sales would be made at a single price, the lowest on offer.

1.153 On the other hand, if markets were perfect, producers would be prepared to supply at different prices only if the products were qualitatively different. Included in the term "product" are the terms and conditions surrounding the sale, including the level of service and convenience. It is tempting to assume, therefore, that the mere existence of different prices implies that the products *must* be qualitatively different in some way. For example, even units of the same physically homogeneous product produced at different locations or times of the day may be qualitatively different from an economic viewpoint. For example, a service supplied in the center of town in the evening may carry a price premium, due to higher labor costs, even though it is essentially the same service. In this instance the higher price is arguably not a pure price difference. However, the relative prices in different establishments do not necessarily have to match differences in producer inputs and technologies and consumers' preferences and may be, in part, pure price differences. In practice, almost all markets are imperfect to some extent, and pure price differences cannot be assumed away a priori.

1.154 If there is only a single homogeneous product produced by an establishment on a "normal" day, the price differences must be pure. The average price is equal to the *unit value*, defined as the total value sold divided by the total quantity. The unit value is a quantity-weighted average of the different prices at which the product is sold. It changes in response to changes in the mix of quantities sold at different prices as well as to any changes in the prices themselves. In practice, however, the change in the unit value has to be estimated from a sample of prices only. Unit values exist at two levels. The first is for a production run i at the establishment level where a batch of, say, q_i products may be sold for revenue $p_i q_i$, the price recorded being the unit value. There may be more than one production run at different batch sizes, and the unit values may vary with batch size. The recorded "price" for these products may then be the revenue from several batches divided by the quantity supplied, $\Sigma\, p_i q_i \;/\, q_i$. If the mix of batch sizes varies over time, then there will be unit-value bias when dividing the unit value in one period by that in a preceding period. The second aggregation of unit values is across establishments producing the same commodity. Again, any difference in the relative quantities sold from different establishments will lead to unit-value bias if the commodities are not strictly homogeneous.

I.5.1 Sets of homogeneous products

1.155 The economic approach views the products as if they were a sample from a basket produced by a group of rational, revenue-maximizing producers. One critical factor is how much product variation there is within an elementary aggregate, bearing in mind that it should be as narrowly defined as possible, possibly even consisting of a set of homogeneous products.

1.156 If the sampled products are all identical, the observed price differences must be due to establishments using different production technologies and market imperfections such as price discrimination, consumer ignorance, or rationing, or some kind of temporary disequilibrium. Informed producers with unrestricted production possibilities would not sell at a lower price if they had the opportunity to sell exactly the same product at a higher price. It is tempting to assume, therefore, that the products are not really homogeneous and that the observed price differences *must* be due to quality differences of some kind or another, but imperfections in producer and consumer markets are widespread and cannot be assumed away a priori.

1.157 As explained in Section B of Chapter 20, when a single product is sold at different prices, the price of that product for PPI purposes is the unit value, defined as total sales divided by total quantities: that is, the quantity-weighted average price. The price relative for the product is the ratio of the unit values in the two periods. This may be affected by a change in the pattern of products that sell at high and low prices as well as by changes in the individual prices.

1.158 If the representative sampled products are selected with probabilities proportional to the quantities sold at the different prices in the first period, a simple (unweighted) arithmetic average of their prices will provide an estimate of the unit value in the first period. The Dutot index is the ratio of the simple arithmetic average prices in the two periods. However, given that the two sets of prices are perfectly matched—that is, geared to the pattern of production in the first period only—the Dutot cannot take account of any changes in the patterns of production between the two periods and may not provide an unbiased estimate of the ratio of the unit values. As shown in Section F of Chapter 20, the sample Dutot with probabilities proportional to quantities sold in the first period may be expected to approximate to a Laspeyres-type index in which the quantity weights are fixed, by definition. It does not provide a satisfactory estimate of a unit-value index in which the relative quantities do change. Moreover, this approximated Laspeyres-type index is not a conventional Laspeyres index because the quantities do not refer to different products, or even different qualities, but to different quantities of exactly the same product sold at different prices.

1.159 In practice, even though producers' choices may be restricted because of their production technology, buyer-seller relationships, market ignorance, and other market imperfections, they may switch production toward products sold at high prices and away from those at low prices, as market conditions change and restrictions on choice are eased. The Dutot index, based on matched prices, cannot take account of such switches and may tend to understate the *rise* in the unit values for this reason. Alternatively, it may be that the demand side dictates market behavior, with establishments responding to demand by increasing production of low-priced products. When the ratio of the unit values changes because purchasers, or at least some of them, succeed in switching from establishments selling at high prices to establishments selling at low prices, the failure of PPIs to take account of such switches leads to the Dutot index overstating the *fall* in the unit-value index.

I.5.2 Heterogeneous elementary aggregates

1.160 In practice, most elementary aggregates are likely to contain a large number of products that are similar but not identical. Assuming producers are informed and have a perfectly flexible set of production possibilities, the relative prices may then be expected to reflect producers' marginal rates of substitution. Within the same elementary aggregate, the different products will often be close substitutes for each other, often being no more than marginally different qualities of the same generic product, so that the quantities produced may be expected to be quite sensitive to changes in relative prices.

1.161 Using an economic approach, it is possible to ask what is the best estimate of the "true" economic index, for the elementary aggregate. Bearing in mind, however, that no information on quantities and revenues is available within the aggregate, it is necessary to resort to considering certain hypothetical special cases. Suppose that producers react to purchasers' preferences; as demand increases for a relatively low-priced product, producers produce more of it. Assume purchasers have so-called Cobb-Douglas preferences, which imply that the cross-elasticities of substitution between the different products are all unity. The quantity relatives vary inversely with the price relatives, so that their revenue shares and the establishment's revenues remain constant. The true economic index can then be shown to be a weighted geometric average of the price relatives, the weights being the revenue shares—which, as just noted, are the same in both periods. Now, suppose that the products whose prices are sampled are randomly selected with probabilities proportional to their revenue shares in the first period. As shown in Section F of Chapter 20, with this method of selection, the simple geometric average of the sample price relatives—that is, the Jevons index—may be expected to provide an approximation to the underlying economic index.

1.162 However, for PPIs the assumption of unit cross-product elasticities of substitution with equal revenues in both periods is *not* consistent with producer economic theory. Revenue-maximizing producers will produce *more* of the sampled products with above-average price increases, so their share of revenue cannot be expected to be constant. Indeed the Jevons index, in assuming constant revenue shares, will understate price changes under such revenue-maximizing behavioral assumptions. The Jevons index allows implicit quantities to fall as relative prices increase, to maintain equal revenue share, rather than allowing an increase. There is not an accepted unweighted price index number formula that incorporates such substitution behavior, although the Jevons index has been shown to be unsuitable under producer revenue-maximizing assumptions.

1.163 Alternatively, suppose that the production technology is such that, at least in the short term, there is no substitution in response to relative price changes, and the relative quantities remain fixed. In this case, the true economic index would be a Laspeyres-type index. If the products were sampled with probabilities proportional to the revenue shares in the first period, a simple arithmetic average of the price relatives—that is, the Carli index—would approximate to it.[10] However, assuming no substitution is unreasonable and counterfactual in general, although it may occur exceptionally.

1.164 Thus, using the economic approach, under one set of conditions the Jevons index would provide an approximation to the underlying economic index, while under another set of conditions the Carli index would do so. In most cases, the actual conditions seem likely to be closer to those required for the Jevons to estimate the underlying index than for the Carli, since the cross-elasticities of substitution seem much more likely to be close to unity than zero for industries whose pricing behavior is demand driven. Thus, the economic approach provides some support for the use of Jevons rather than Carli, at least in most situations. However, if producer revenue-maximizing behavior is believed to dominate an industry, use of the Jevons index is not supported.

1.165 Another alternative is suggested in Section G of Chapter 20. If products are sampled according to fixed revenue shares in each period, then the resulting sample can be used with the Carli formula (P_C) to estimate the Laspeyres index, and the harmonic mean formula (P_H) to calculate the Paasche index. By taking the geometric average of these two formulas, as suggested by Carruthers, Sellwood, Ward (1980), and Dalén (1992a), a Fisher index would result:

$$(1.18)\; P_{CSWD} = \sqrt{P_C \times P_H}\ .$$

1.166 However, since statistical offices would not have the revenue shares for the current period, an approximation to the Fisher index is obtained by assuming they are not too different from those used in the base period 0. A similar assumption would justify the use of a Jevons index (P_J,) as an approximation to a Törnqvist index. Again recall, that these approximations result when the observations are sampled in proportion to revenue shares.

1.167 One lesson to be drawn is that, when trying to decide on the most appropriate form of the price index for an elementary aggregate, it is essential to pay attention to the characteristics of the products within the aggregate and not rely on a priori generalizations. In particular, the Dutot index should be used only when the products are homogeneous and measured in exactly the same units. When the products are heterogeneous, the choice between the Carli and the Jevons indices turns on the extent to which, and the nature of, substitution behavior that is likely to occur in response to relative price changes. In many cases, the Jevons is likely to be preferred. Because Jevons is also the preferred index on axiomatic grounds, it seems likely to be the most suitable form of elementary index in most situations, although the circumstances underlying its use should be carefully established.

[10] Notice that the Dutot index cannot be used when the products are not homogeneous, since an arithmetic average of the prices of different kinds of products is both arbitrary and economically meaningless. If a Laspeyres index is estimated as a simple average of the price relatives—that is, assuming equal revenue shares—the implied quantities cannot be equal because they vary inversely with the prices.

J. Seasonal Products

1.168 As explained in Chapter 22, the existence of seasonal products poses some intractable problems and serious challenges for PPI 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.

1.169 There are two main sources of seasonal fluctuations: the climate and custom. Month-to-month movements in a PPI may sometimes be so dominated by seasonal influences that it is difficult to discern the underlying trends in prices. Conventional seasonal adjustment programs may be applied, but these may not always be satisfactory. However, the problem is not confined to interpreting movements in the PPI; seasonality creates serious problems for the compilation of a PPI when some of the products in the basket regularly disappear and reappear, thereby breaking the continuity of the price series from which the PPI 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.

1.170 One possibility is to exclude seasonal products from the index, but this may be an unacceptable reduction in the scope of the index, since 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-over-year figure is much easier to interpret than month-to-month changes, which can be somewhat volatile, even in the absence of seasonal fluctuations.

1.171 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 centered 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 Section F of Chapter 22 it is shown that under certain conditions the current month's year-over-year monthly index, together with the previous month's year-over-year monthly index, can successfully predict the rolling-year index that is centered on the current month. Of course, rolling-year indices and similar analytic constructs are not intended to replace the monthly or quarterly PPI but to provide supplementary information that can be extremely useful to users. They can be published alongside the official PPI.

1.172 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.

K. Concepts, Scope, and Classifications

1.173 The purpose of Chapter 3 of the *Manual* is to define and clarify a number of basic concepts underlying a PPI and to explain the scope, or domain, of the index: that is, the set of products and economic activities that the index is intended to cover. The chapter also discusses the various price concepts and types of prices that are used in PPI compilation and examines the structure of the classification systems used in the PPI for products and industries.

1.174 The general purpose of an index of *producer* prices is to measure changes in the prices of goods and services produced by businesses. However, an operational definition of a PPI requires a decision about, first, whether the index will cover output prices or input prices (or both); second, whether the index is meant to cover all production, that is, all economic activities and/or products, or just particular industries and/or product groups; third, for the economic activities included, whether the index should cover just market activities; and, finally, what is the geographic boundary in which the defined production is included. The scope of a PPI is inevitably influenced by what is intended or believed to be its main use, although it should be borne in mind that the index may also be used as a proxy for a general price index and used for purposes other than those for which it is intended.

K.1 Population coverage

1.175 Many decisions must be made to define the scope and coverage of the PPI. These include the economic activities, products, and the types of buyers and sellers to include in the index. The PPI could cover all economic activities in a country, which could be the ultimate goal of the price index. In many countries the PPI is limited to a few industrial activities such as agriculture, mining, manufacturing, and energy supply. These activities represent a good starting point. However, the share of such activities in national economies is becoming smaller, and services such as transport, communication, medical care, trade, and business services are becoming increasingly more important. If the primary purpose of the PPI is an inflation indicator or a deflator for national accounts aggregates, a broad coverage of economic activity is needed.

1.176 A PPI can be compiled and classified both by industry and product. For example, the food slaughtering industry produces meat and leather. Generally, the industrial coverage of the PPI is limited to specific industrial sectors such as mining and manufacturing, and this in turn limits the product coverage. If broad product coverage is a goal, then the PPI would have to cover a larger number of goods- and service-producing sectors. The PPI can also identify products by stage of processing and produce measures of products for final demand, those for intermediate consumption, and those that are primary products.

1.177 A PPI also could cover all domestic production, including exports, or be limited to production for domestic markets only. If it covers all domestic production, then products for export could be separately identified, and an export price index developed. Imports are usually not within the scope of the output PPI because their production is not domestic, but they could be covered in an input price index. (Foreign trade price indices will be the subject of a separate manual.) In addition, the PPI is usually limited to marketed products and thus excludes nonmarket goods and services.

K.2 Price coverage

1.178 The PPI should measure actual transaction prices reflecting revenue received by the producer for goods and services actually sold to customers. These prices would not necessarily be “list” or “book” prices because they should reflect any applicable discounts, rebates, surcharges, etc. that may apply to their customers for the sampled transactions. These would include contract prices, where they exist, and spot market prices. Care must be taken to make sure the prices reflect those at the time the transaction occurs and not those at the time of order, particularly for major durable goods such as airplanes and ships, which have a long production period between order and delivery.

1.179 Average prices are acceptable in the PPI if they represent a strictly *homogeneous* set of product transactions and are for the current time period. Often these two criteria for an average price cannot be met. If average prices are calculated over a large number of transactions with differing quality and/or terms of sale, they are not acceptable in the PPI. Changes in such prices will reflect any changes in the mix of quality characteristics of the products sold as well as any changes in terms of sale. Such changes in the heterogeneous mix of transactions lead to what is often referred to as *unit-value bias* in the measurement of price changes.

1.180 Special care needs to be taken with subsidized prices and intracompany transfer prices. The prices used in the PPI should reflect the revenue received by producers from transactions. Prices for products on which subsidies are received will not reflect the revenue to the producer unless the subsidies are included. This involves making adjustments to the prices as discussed in Section B.3 of Chapter 3. Also, intracompany transfer prices may not reflect actual market prices and may require special treatment as outlined in Section B.4 of Chapter 3.

K.3 The treatment of some specific types of transactions and prices

1.181 The price concept is not always as clear-cut as that for simple homogeneous goods sold in day-to-day transactions. There are a number of conceptually difficult products and industries that present particular problems—agriculture, clothing, steel, ships, automobiles, and banking services, to name a few. Pricing concepts and strategies for these and other special cases are covered in more detail in Chapter 10.

K.4 Statistical units

1.182 The statistical unit in the PPI is usually a single, homogeneous, output-generating entity such as the *establishment*, a concept outlined in the *1993 SNA*. Separate auxiliary, sales, or administrative units are not included. This unit is the decision-making unit for all production operations and maintains records on prices and production activities. In some cases records from a clustering of establishments are sent to a single record-keeping unit, the enterprise, from which prices will have to be collected.

1.183 The rapid rise in electronic commerce (e-commerce), globalization, and outsourcing of production is making the identification of the statistical unit, the producing establishment, more difficult. This is particularly the case with the formation of *virtual corporations*. A virtual corporation is the creation of a partnership among several companies sharing complementary expertise and producing a product with a very short life cycle. With the conclusion of the product's life span, the corporation is disbanded. Also, a considerable volume of business undertaken among corporations is being transacted on the Internet, which is difficult to monitor. These activities will require new approaches to identify and capture such transactions in the PPI.

K.5 Classification

1.184 The classification system provides an organizing structure for the PPI and is the first step in sample surveying. It forms the index structure and defines which industries, products, and aggregate levels will be included. It also determines the publication scheme for the PPI results. International standard classification systems, discussed in Section E.2 of Chapter 3, are available and should be used to group economic activities and products. The use of these classifications provides meaningful series for policymaking and analysis, as well as facilitating international comparisons.

1.185 Industrial classifications group producer units according to their major kind of activity, based mainly on the principal class of goods or services produced—that is, by an output criterion. At the most detailed level of industrial coding (the four-digit level), categories are delineated according to what is in most countries the customary combination of activities undertaken by the statistical units, the establishments. The successively broader levels of classification (three-digit, two-digit, single-digit) combine the statistical units according to character, technology, organization, and financing of production. The major international industrial classifications are the *International Standard Industrial Classification of all Economic Activities (ISIC)*, the *General Industrial Classification of Economic Activities within the European Communities (NACE)*, the *North American Industrial Classification System (NAICS)*, and the *Australian and New Zealand Standard Industrial Classification (ANZIC)*.

1.186 Product classifications group products into somewhat homogeneous categories on the basis of physical properties and intrinsic nature, as well as the principle of industrial origin. Physical properties and intrinsic nature are characteristics that distinguish the product. These include raw materials from which the goods are made, the stage of production and way in which the goods are produced or service rendered, the purpose or use of the products, and the prices at which they are sold. The product categories should be exhaustive and mutually exclusive so that a product belongs to only one category.

1.187 The categories of products (coded, for example, to five digits) can be aggregated to higher-level groupings (four, three, two, and single digits) of products with similar characteristics and uses. There are two primary international product classifications used for PPIs: the *Central Product Classification (CPC)* and the *Eurostat Classification of Products by Activity (CPA* or *PRODCOM)*. In general, each five-digit subclass of the CPC consists of goods and services that are predominantly produced in one specific four-digit class or classes of ISIC Revision 3.

L. Sampling and Collection of Price Data

1.188 As explained in Chapter 9, there are two levels of calculation involved in a PPI. 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 behavior are explained in Chapter 20 and are summarized in Section I above. At the higher level, the elementary indices are averaged to obtain higher-level indices using the relative value of output or revenue as weights. All the index number theory elaborated in Chapters 15–18 comes into play at this higher level.

1.189 Lower-level indices are calculated for elementary aggregates. Depending on the resources available and procedures adopted by individual countries, these elementary aggregates could be subclasses of the industry and product classifications as described in the previous section. If it is desired to calculate PPIs for different regions, the subclasses have to be divided into strata referring to the different regions. In addition, in order to improve the efficiency of the sample estimator, it will usually be desirable, if feasible, to introduce other criteria into the definitions of the strata, such as the size of the establishment. When the subclasses are divided into strata for data collection purposes, the strata themselves become the elementary aggregates. Because a weight needs to be attached to each elementary aggregate in order to calculate the higher-level indices, an estimate of the quantity or value of output for each elementary aggregate should be available, albeit in a preceding period, from separate surveys of establishments, as outlined in Chapter 4 and Section N below. On the other hand, quantity data may not be readily available for all elementary aggregates and may have to be estimated using allocation methods like those described in Chapter 4, Section E.1. It is preferable that the lower-level indices also be compiled using quantity or value weights and that such data are collected at the time of price collection in the initial base period, or, if possible, also in each successive period. This would allow Laspeyres or Fisher indices to be compiled at the lower level, which is advisable for theoretical reasons outlined in Section I above and in Chapter 20. If no weights can be derived, the elementary indices have to be estimated from price data alone, as explained in Chapter 20.

1.190 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. The sampling and collection of prices is at the lower level, sampling being considered first.

L.1 Random sampling and purposive sampling

1.191 Prices are collected for products from establishments in particular industries. The sampling process involves multiple stages of selection. Once the purpose and scope for the PPI have been decided (for example, which single-digit industrial activities will be included), then decisions can be made about the four-digit industries to be included. After the industries have been chosen, then the establishments within industries must be selected and sampled, and then individual (representative) products must be selected or sampled. Finally, individual transactions that represent the sampled products in each sample establishment must be selected. The procedures used for selecting the sample at each stage are important.

1.192 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.[11] 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 stratify the establishments and products by criteria that differentiate them according to their relative price changes, and to further select both establishments 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. However, many countries continue to rely heavily on the purposive selection of establishments 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 for PPI purposes. It may also be cost effective to use "cutoff" sampling procedures, discussed in Chapter 5, Section D.1.2, which are more objective than purposive sampling. Cutoff sampling first establishes a targeted threshold value, and then all establishments/products

[11]There are two types of bias encountered in the literature on index numbers: *sampling bias,* as understood here, and the *nonsampling biases* in the form of substitution bias or bias due to inadequate adjustments for quality change, as discussed in Chapters 11 and 7 of the *Manual*. It is usually clear in context which type of bias is meant.

above this value are selected for the sample. It is a simple means, for example, of selecting the representative four-digit industries within a single-digit category, or products within an establishment.

1.193 The representative sampling of establishments and products requires comprehensive and up-to-date sampling frames. Two separate frames are usually needed for PPI purposes, one listing the universe of establishments and the other listing the universe of products. Examples of possible sampling frames for establishments are business registers, establishment censuses, and central or local government administrative records. When the sampling frames contain the requisite information, it may be possible to increase the efficiency of the sample estimate by selecting samples of establishments using probabilities that are proportional to the size of some relevant economic characteristic, such as the total value of output or sales. Sampling frames for products are usually available from establishment or business censuses and may be supplemented by telephone or price survey visits.

1.194 Depending on the information available in the sampling frame, it may be possible to group the establishments into strata on the basis of region, in addition to industrial activity, to form the elementary indices. When there is information about size, a random sample of establishments may be selected with probabilities proportional to size. An example of this approach is presented in Chapter 5, Section E. Price relatives from preceding periods may further be used as part of the sample allocation, with larger samples being drawn from industrial groups whose variance of price relatives is larger. All of this increases the efficiency of the sample estimate. It would also be possible to use cutoff sampling procedures as a simpler, though less efficient, procedure. Cutoff sampling, unlike random sampling, is open to bias, if the excluded smaller establishments have different price changes to the included larger ones. The extent of the bias depends on the threshold cutoff value and the level of aggregation; some of the bias will be offsetting.

1.195 In most countries, the selection of the individual products to be priced within the selected establishments tends to be purposive, being specified by the central office responsible for the PPI. The central office draws up lists of products that are deemed to be *representative* of the products within an elementary aggregate. However, if detailed output or sales by product are available from a census of establishments, these data can be used to select the sample through probability proportional to size or cutoff sampling.

1.196 It has been argued that the purposive selection of products is liable to introduce only a negligible amount of sampling bias, but this may be no more than speculation or conjecture. In principle, random sampling is preferable, but it may not be feasible for many countries given the additional costs that may be involved. For example, the U.S. Bureau of Labor Statistics (BLS) and the U.K. Office of National Statistics make extensive use of random selection procedures to select both establishments and products within establishments. The last stage of sampling is to select the individual transactions within the establishment to represent the price movements of the selected products. Procedures for selecting transactions are presented in Chapter 5, Section E.3. At this level many countries consult with an official from the establishment to select the most representative transactions for each product. Often selecting those with the largest volume of output or sales does this. Such a procedure is analogous to using cutoff sampling. It is also possible to select a probability sample of transactions if the officials can provide estimates of the relative importance of the transactions.

1.197 As explained in Chapter 5, Section F, the universe of establishments and products, from which the sample is taken, has several dimensions. That the universe is changing over time is a major problem not only for PPIs but also for most other economic statistics. Products disappear, to be replaced by other kinds of products, and establishments close while new ones open. This creates both conceptual and practical problems, given that the measurement of price changes over time requires some continuity in the products priced. The matched-models method requires that the price changes recorded should refer to matched products that are identical in both time periods, so that price changes are not tainted by quality changes. But this matching creates a new problem: new products and new establishments are not introduced, and the sample deteriorates. There are further problems created when products are no longer produced or establishments close, and these are considered in some detail in Chapters 7 and 8, and are outlined in sections L.2.4, L.2.5, and M below.

L.2 Regular price collection

1.198 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 establishments. This section is concerned with some of the operational issues relating to price collection, which are discussed in detail in Chapter 6.

L.2.1 Frequency and timing

1.199 Calculating the PPI entails collecting prices from businesses relating to particular products and time periods. Decisions must be made about the frequency of collection (monthly or quarterly) and the time period covered for the prices (a single point in time, several times during the month, or a monthly average). Usually, price collection is monthly and covers the entire month. However, resource considerations may limit collection to a single point in time.

L.2.2 Product specifications

1.200 For each product in the sample, a detailed list of the specifications needs be collected. These specifications are those that are important in identifying and determining the price and quality characteristics of the detailed transaction. Details such as product name, serial number, description or features, size, units of measure, class of customer, discounts, etc. should be included. The collection of data on such quality characteristics is important to the matched-models method, but it will be seen from Section M below that they can serve as a data source for hedonic regressions, which have a similar function—to price-adjust replacement products of different quality.

L.2.3 Price collection methods

1.201 The aim of survey collection techniques is to facilitate the transmission of price data from businesses to the statistical office in a secure and cost-effective manner, while minimizing the administrative burden of the respondent. In principle, the relevant prices for a PPI should be the basic prices actually received by the establishment. For some products, the prices collected may be *estimated* transaction prices because the transaction sampled did not have sales during the reference period. In addition, it is generally neither practical nor cost-effective to try to collect prices each month or quarter directly from establishments by personal visits. Data can effectively be collected using mail questionnaires, telephone contacts, fax, and electronic media. A range of approaches to PPI data collection are presented in Chapter 6: postal survey, automated telephone response, personal interview, telephone interview, and Internet data provision. All of these methods rely on good questionnaire design, good respondent relations, and good interviewing techniques. The exact methods chosen by countries for particular industries will depend on the special circumstances applicable to each form of collection in their industry/country.

L.2.4 Continuity of price collection

1.202 A PPI 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. Identical economic characteristics include the terms and conditions of sale. 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.

1.203 Having identified the products whose prices are to be collected, the normal strategy is to ask the respondent to continue pricing exactly those same products for as long as possible. The respondents can do this if they are provided with very precise, or tight, specifications of the products to be priced. Alternatively, they must keep detailed records themselves of the products that they have selected to price.

1.204 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.[12] Most products, however, have only a

[12]It is worth noting that many theorems in index number theory are derived on the assumption that exactly the same (continued)

limited economic life. Eventually, they disappear from the market to be replaced by other products. Because the universe of products is continually evolving, the representative products selected initially may gradually account for a progressively smaller share of output and sales. As a whole, they may become less and less representative. Since a PPI is intended to represent all products, some way has to be found to accommodate the changing universe of products. In the case of producer 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.

1.205 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. In 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. Similar issues of course arise for establishments, although the focus here is on products.

L.2.5 Resampling

1.206 One strategy to deal with the changing universe of products would be to resample, or reselect, at regular intervals the complete set of products to be priced. For example, with a monthly index, a new set of products could be selected each January. Each set of products 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 revenue weights each year.

1.207 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 because 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. A more practical way in which to keep the product sample up to date is to rotate it gradually by dropping certain products and introducing new ones. Products may be dropped for two reasons:

- The product is believed by the respondent or central office to be no longer representative. It appears to account for a steadily diminishing share of the total revenue within the product group or industry in question.
- The product may simply disappear from the market altogether. For example, among other reasons, it may have become obsolete due to changing technology, or unfashionable due to changing tastes.

1.208 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 products priced. This raises the general question of the treatment of quality change and the treatment of new products.

M. Adjusting Prices for Quality Changes

1.209 The treatment of quality change is perhaps the greatest challenge facing PPI compilers. It is a recurring theme throughout the *Manual*. It presents both conceptual and practical problems for compilers of PPIs. The whole of Chapter 7 is devoted to the treatment of quality change, and Chapter 8 addresses the closely related topic of new goods and product substitution.

1.210 When a sampled product is no longer produced or is unrepresentive and is dropped from the list of products priced in some establishment, the

---

set of goods and services is available in both the time periods being compared.

normal practice is to find a new product to replace it. This is 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 transaction and the subsequent series for the new transaction. 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 introduced only after the one that 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.

1.211 However difficult it is to estimate the contribution of the changed 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. However, 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 adjustment, as explained below. Whatever the resources available to them, statistical offices must be conscious of the implications of the procedures they adopt.

1.212 Three points are stressed in Section A 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 among 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, since different methods can lead to very different results.

M.1 Evaluation of the effect of quality change on price

1.213 It is useful to try to clarify why one would wish to adjust the observed price change between two products 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 PPI, a quality change must be evaluated from the producer's perspective with regard to the revenue received. As explained in Section B of Chapter 7, the evaluation of the quality change is essentially an estimate of the per-unit change in revenue that a producer will receive for the new characteristics possessed by the new quality using the same technology. This amount is not a price change because it represents the monetary value of the change in the value of production that is involved to produce the new quality. The value can either be estimated on the basis of the value to the user of the new quality, or the production costs from the producer.

1.214 In many cases the concern can be seen with a need to make a quality adjustment to either the original or replacement product's price. The prices of two products need to be compared. They differ in quality, so some of the difference in price is due to quality differences. A quality adjustment in this instance is seen as an adjustment to the price (or price change) of the original or replacement product to remove that part due to quality differences. A quality adjustment can be seen as a coefficient that multiplies the price of, say, the replacement product to make it commensurate, from the producer's point of view, with the price of the original. To take a simple example, suppose that the quantity of some product and its replacement are variable and that quantity k of the replacement is produced using the same technology at the same cost and sold for the same price as quantity j of the original. The producer is indifferent between selling one unit of the original and j/k units of the replacement. To make the price of one unit of the replacement commensurate with the price of one unit of the original, it must be multiplied by k/j. This is the required quality adjustment.

1.215 For example, if two units of the replacement product are equivalent to three of the original, the required quality adjustment to be applied to the price of the replacement product is ⅔. Suppose the revenue from one unit of the replacement is the same as one unit of the original, then the price of the replacement, after adjusting for the change in quality, is only ⅔ that of the price of the original. If one unit of the replacement sells for twice the price of the original, then the quality-adjusted price is (2 × ⅔ =) 1⅓ that of the original: the price increase is 33 percent, not 100 percent. The PPI seeks to record the change between the price of the original and the quality-adjusted price of the replacement.

1.216 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 in which to make a significant price change. They may deliberately make it difficult for purchasers to disentangle how much of the observed difference in price between the old and the new qualities represents a price change.

1.217 For PPI purposes, an explicit quality adjustment is often possible using differences in the costs of production between the two qualities. This approach works as long as production costs are based on the establishment using the same technology. Another 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. The discussion below examines the implicit methods first and then the explicit methods. These approaches are examined in some detail in Sections D and E of Chapter 7.

1.218 When the technology changes, there is no comparable basis for comparing costs between the two qualities, and these procedures break down. An alternative approach would be to use hedonic regression techniques, which are also discussed below and in more detail in Section G of Chapter 7.

M.2 Implicit methods

M.2.1 Overlapping qualities

1.219 Suppose that the two qualities overlap, both being produced at time t. If both are produced and sold in a competitive market, economic theory suggests that the ratio of the prices of the new to the old quality should reflect their relative cost to producers and value to purchasers. 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 onward 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.

1.220 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 to deal with noncomparable replacements 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, the market may not have had time to adjust, particularly when there is a substantial change in quality. When the new quality first appears, the market is liable to remain in disequilibrium for some time. The producers of new qualities may price strategically over the product life cycle to, for example, price-discriminate in the early periods following introduction. There is a case in which the overlap method is used extensively in spite of these difficulties: when the index is rebased or products are rotated. The advantage of refreshing the sample is deemed to outweigh such disadvantages.

1.221 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, it may be that their prices fall relatively to older qualities, before the latter eventually disappear. In general, 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.

1.222 The statistician has then to make a deliberate judgment 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 that uses the prices for both the new and the old qualities in several periods in which they overlap. Such information may be available from the respondent's records, although this requires a good relationship with the respondent and good record-keeping and retrieval systems by the respondent. 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.

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

M.2.2 Nonoverlapping qualities

1.224 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. Adopt the notation that the actual price of the new quality is P_t in period t and the price of the old quality is p_{t-1} in the previous period. Since the new quality is not available in period t, an imputation is made for its price in period 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 producer is needed. The ratio p^*_t / P_t is the required quality adjustment since this ratio provides the estimate of the quality differences at the same point in time. Using lowercase *p*s for the old quality and uppercase *P*s for the new, it is assumed that the price data available to the index compiler take the following form:

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

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; and
- Price changes from period t onward are measured by the series for the new quality.

1.225 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 that 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 noncomparable replacement is deemed comparable.
- 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.

1.226 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 could be justified only if the evidence suggests that the extent of the quality change is negligible, even though it cannot be quantified more precisely. "Doing nothing"—that is, ignoring the quality change completely—is equivalent to adopting the first solution. Conversely, the second could be justified only if 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.

1.227 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 onward, 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.

1.228 It may be possible to improve on this estimate by making a careful selection of the other sampled products to include only those whose average price change is believed to be more similar to the product in question than the average for the group of sampled products as a whole. This procedure is described in some detail in Section D.2 of Chapter 7, where it is illustrated with a numerical example and is described as "targeting" the imputation or estimation.

1.229 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" mean method. The more refined, targeted version is the "targeted" or "class" 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.

1.230 Although the overall mean method superficially seems a sensible practical solution, it may nevertheless give biased results, as explained in Chapter 7. It needs to be repeated that 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 for products not subject to quality change.

1.231 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.

M.3 Explicit quality adjustments

M.3.1 Quantity adjustments

1.232 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, it is very unlikely to rate a refrigerator that has three times the capacity of a smaller one as 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 commonsense, application of relatively straightforward quality adjustments of this kind. A discussion of quality adjustments based on size is given in Section E.2 of Chapter 7.

M.3.2 Differences in production/ option costs

1.233 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 method is explained in Section E.3 of Chapter 7. The estimates can be made in consultation with the producers of the goods or services, if appropriate. This method, like the preceding one, is likely to be satisfactory only when the quality 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 quite unacceptable, for example, when a drug is replaced by another more effective variant of the same drug that also happens to cost less to produce.

1.234 Another possibility when the quality change is more complex or subtle is to seek the advice of technical experts, especially when the respondent 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.

M.3.3 The hedonic approach

1.235 Finally, it may be possible to systematize the production/option cost approach by utilizing econometric methods to estimate the impact of observed changes in the characteristics of a product on its price. 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 may be 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 estimate the effects on price of changes in those characteristics.

1.236 The hedonic approach to quality adjustment can provide a powerful, objective, and scientific method of estimating the effect on price of 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 Section E.4 of Chapter 7. Products can be viewed as bundles of tied characteristics that are not individually priced because the producer sells 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 random-access memory (RAM), and the hard drive capacity. An example of a hedonic regression using these and other characteristics is given in Section E.4 of Chapter 7, the actual numerical results being given in Table 7.3.

1.237 The results obtained by applying hedonics to computer prices have had a considerable impact on attitudes toward the treatment of quality change in PPIs. They have demonstrated that for goods where there is rapid technological change 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. Reference may be made to Section G of Chapter 7 for further analysis, including a comparison showing that the results obtained by using hedonics and matched models can differ significantly when there is a high model turnover.

M.4 Conclusions on quality change

1.238 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 overemphasized. Failure to pay proper attention to quality changes can introduce serious biases into the PPI.

N. Product Substitution and New Goods

N.1 Replacement products

1.239 As noted in the previous section, price indices would, ideally, 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 PPI 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, and in both cases their price behavior is likely to diverge from that of the matched products. Price indices have to try to take account of the price behavior of new and disappearing products so far as possible.

1.240 A formal consideration and analysis of these problems is given in Appendix 8.1 in 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 in the previous section are needed when comparing the prices of the replacement products with those of the products that they replace.

1.241 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 or establishments to be drawn to replace the existing ones. The two samples must overlap in one period that acts as the link period. As noted in Section B.2 of Chapter 8, 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 in 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, however, expensive.

N.2 New goods and services

1.242 The difference in quality between the original product and the one that replaces it 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 Section D of Chapter 8, a distinction is also drawn in the economic 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; 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 subclass of the product or industry classification, whereas a revolutionary new good will require some modification to the classification in order to accommodate it.

1.243 As explained in Section D.2 of Chapter 8, a major concern with new goods or services relates to the timing of the introduction of the new product into the index. It is often the case that new goods enter the market at a higher price than can be sustained in the longer term, so that their prices typically tend to fall over the course of time. Conversely, the quantities sold may be very small initially but may increase significantly over time. These complications make the treatment of new products particularly difficult, especially if they are revolutionary new goods. Because of the tendency for the price of a new good to fall even after it has been introduced, it is possible that important price reductions may fail to be captured by PPIs because of the technical difficulties created by new products. The issues are examined in some detail in Section D of Chapter 8. The chapter concludes by expressing concern about the capacity of PPIs 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.

O. Revenue Weights

1.244 Once the price data have been collected and adjusted as necessary, the next step in the calculation of a PPI is to combine, or average, the elementary price indices to arrive at price indices at higher levels of aggregation up to the overall PPI itself. For this purpose, revenue weights are needed for the various elementary aggregates. These weights are needed whatever index number formula is used for aggregation purposes. Chapter 4 is concerned with the derivation and sources of the revenue weights.

O.1 Establishment censuses and surveys

O.1.1 Establishment or business censuses

1.245 The establishment or business census covers all establishments that have productive activity within the geographic borders of the country. These censuses may be conducted over a span of years, with different economic activities covered at different times during the cycle. For example, a census of agriculture would be conducted one year, a census of industrial activities (mining, manufacturing, and energy supply) completed during the next year, followed by a census of services. In some instances there may be a size cutoff to exclude very small establishments. For example, some countries exclude establishments with fewer than five employees or with some low threshold of annual production, or complete the census using only a sample of small establishments.

1.246 A detailed accounting of annual output in value (at basic prices) and quantity terms by detailed product classification is typically obtained at the enterprise or establishment level. This would include sales and inventories by product, as well as value and quantity of inputs at the prices paid by producers. These data can be used to derive the revenue weights by detailed product classification and establishment. This is an excellent source of weight data, assuming that the coverage of economic activity is essentially complete.

O.1.2 Enterprise or industry surveys

1.247 These surveys differ from censuses primarily in three respects:

- The coverage is limited to a sample of establishments rather than a full enumeration;
- The product detail is limited to higher aggregate levels such as groups, and
- The types of data requested are generally more limited than those requested in a census.

1.248 For example, product information in the census may be obtained at the eight-digit product code level using PRODCOM, with complete detail on product sales and inventories, while in the industry survey data are reported at the six-digit level and are requested only for sales. Also, data may be reported only for the enterprise rather than broken down by establishment.

1.249 Thus, for enterprise or industry surveys, the weights that are available will generally be for higher levels in the aggregation structure, such as product group and industry, rather than detailed product and establishment. The use of these weights for the PPI will depend on how the PPI aggregation structure has been established. If multitier weights (for example, one set of weights for the industry level and above, and another set of weights at the establishment level and below) have been set up, the survey results could be used for aggregation at higher levels, while the weights at lower levels are determined separately. For example, the survey weights could be used for aggregating from the four-digit industry level to higher levels, while sampling weights (that is, sampling fractions from probability selection procedures) could be used at the establishment and product level. In this scheme, the weights at the higher levels would be updated periodically from the industry survey data, while the weights at the lower levels would be updated as the samples of establishments and products are refreshed. This process is discussed in more detail in Chapter 5.

O.2 Other sources for estimating revenue weights

O.2.1 National accounts

1.250 Although much of the same source data described above would also be used in developing the output data for the production account in the national accounts, there can be significant differences. In a number of countries, there may be significant undercoverage in annual industry surveys owing to the exclusion of informal activities. National accountants often make adjustments from a variety of sources for this type of undercoverage or for known biases in the survey data. In such instances, the adjusted national accounts information on output by industry may prove to be a better source of weight information at the industry level than the original survey data.

1.251 The national accounts often provide additional detail on weights particularly if supply and use tables or input/output tables are available. The information on commodity flows for various industries and commodities by type of use is an excellent source of net weight information for development of stage-of-processing indices. One drawback of national accounts data is that the estimates include imputations for nonmarket activities, and such imputed data may not be appropriate for use as weights in an index whose coverage is primarily market activity.

O.2.2 Business register

1.252 Most countries maintain a business register that provides a list of firms that are involved in productive activities. Such registers usually contain information on location, economic activity, size (for example, employment, payrolls, value of annual production, or turnover), contact persons, tax information, etc. The business register could be an alternative source of weight information, particularly if business censuses are not conducted on a regular basis or if annual surveys do not provide sufficient information for establishing weights. This is particularly true if there is an ongoing system for updating and maintaining the information contained in the register, and it contains data at the establishment level.

1.253 There are several shortcomings to the use of these registers for weight information. Often the business register is updated only when a firm begins operations. Unless the register is maintained by purging firms that are no longer in business, it will have superfluous information. The information on the size of the firm also needs to be updated on a regular basis. Much of the information may relate to the time at which the firm was introduced into the register. Also, the business register may comprise a list of *enterprises*, which is not completely suitable for the PPI, where the concern is to obtain information at the *establishment* level. The register will usually be devoid of information on products, which means that additional data collection will be necessary before weights can be established at the product level.

O.2.3 Additional sources of weights

1.254 A wide variety of administrative data on production values may be available from public agencies charged with regulating or monitoring certain economic activities. For example, national, regional, or local governmental bodies regulate many public utilities, communication, and transport activities. Typically, these agencies require detailed annual reports that provide information on production value and/or turnover. These sources also have records of all regulated enterprises/establishments, which can be used as a source for a sampling frame.

1.255 In many countries, data on retail and wholesale turnover are produced on a regular basis from separate surveys. Such data, if maintained at a detailed economic activity level, could serve as a source of weights for wholesale and retail economic activities. This, of course, would depend on whether wholesale and retail trade will be included in the PPI and if the survey information is deemed reliable for use as weights.

1.256 Customs records are a source of information on exports by product and enterprise. If detailed customs records are maintained and available for statistical purposes, information on detailed products by shipping enterprise should be available and provide a source for weights as well as a potential sampling frame for sampling export products.

P. Basic Index Calculations

1.257 Chapter 9 provides a general overview of the ways in which PPIs 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 PPIs. The various stages in the calculation process are illustrated by numerical examples.

1.258 Chapter 9 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 behavior of indices gained in recent years, it is now recognized that not all existing practices are necessarily optimal.

1.259 Because 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. However, it may be useful to give an indication of the nature of the contents of Chapter 9.

P.1 Elementary price indices

1.260 Chapter 9 describes how the elementary price indices are calculated for the elementary aggregates. It reviews the principles underlying the delineation of the elementary aggregates themselves. Elementary aggregates are relatively small groups of products that 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. They may also be broken down by location and establishment type. Samples of prices are collected for a number of representative transactions across establishments within each elementary aggregate in order to estimate the elementary price index for that aggregate, with each elementary price index providing a building block for the construction of the higher-level indices.

1.261 Section B of Chapter 9 considers the consequences of using alternative elementary index formulas 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 and their properties have been explained in some detail in Section I 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 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 these indices' 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 among the four products. The Dutot and Jevons indices record no increase, but the Carli index registers an increase. It also illustrates the differences between the direct and the chain indices. After six months, each of the four prices is 10 percent higher than at the start. Each of the three direct indices records a 10 percent increase, as also do the chained Dutot and Jevons indices because they are transitive. The chained Carli, however, records an increase of 29 percent, which is interpreted as illustrating the systematic upward bias in the Carli formula resulting from its failure to satisfy the time reversal test.

1.262 Section B.3 of Chapter 9 notes that the chaining and direct approaches have different implications when there are missing price observations, quality changes, and replacements. It concludes that the use of a chain index can make the estimation of missing prices and the introduction of replacement products easier from a computational point of view.

1.263 Section B.5 of Chapter 9 examines the effects of missing price observations, distinguishing between those that are temporarily missing and those that have become permanently unavailable. Table 9.3 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.

1.264 The possibility of using other elementary index formulas is considered in Section B.6. The harmonic mean of the price relatives, P_H, and the ratio of the harmonic means, R_H, are examined. The P_H has the inverse properties of the Carli index, P_C, and can therefore be assumed to have an opposite bias. As it is also a rather difficult concept to explain, it is not recommended. The Jevons index, P_J, has attractive axiomatic properties but is advised only when particular patterns of substitution are expected. The geometric mean of the P_C and the P_H, a kind of elementary Fisher index, remains a possibility with some theoretical attractions, though because it provides close results to the Jevons index, P_J, is advised only under the substitution possibilities discussed in Chapter 20.

1.265 Section C of Chapter 9 discusses the issue of consistency in aggregation between lower- and higher-level indices that may arise if different formulas are used at different levels. Consistency of aggregation means that if an index is calculated stepwise, by calculating intermediate indices that are themselves subsequently aggregated, the same result should be obtained as if the calculation were made in a single step without the intermediate indices. This can be an advantage for purposes of presentation. If a Young or Laspeyres index is used for the higher-level indices, including the overall PPI itself, then the Carli index is the required form of elementary index that is consistent with it.[13] Given that the Carli does not emerge as the preferred elementary index from the axiomatic and economic approaches to elementary indices, this creates a dilemma when the Laspeyres or Young formula is used. It is suggested that consistency in aggregation may not be so important if there are different degrees of substitution within elementary aggregates at the lower level, as compared with the degree of substitution between products in different elementary aggregates at the higher.

1.266 It is not necessary to use the same index formula for every elementary index. The characteristics of the price behavior within each elementary aggregate should be examined to identify the most appropriate formula. However, it may be decided to use a single formula throughout if resources are limited and computational procedures need to be kept as simple as possible.

P.2 Calculation of higher-level indices

1.267 Section C of Chapter 9 considers the calculation of the higher-level indices utilizing the elementary price indices and the weights provided for the elementary aggregates. In many instances statistical offices do not use a true Laspeyres index but rather a Lowe or Young index (discussed in Section B.1 above). These two indices use price reference periods and weight reference periods that differ, while in the Laspeyres index the price and weight reference period are one and the same. Typically the weight reference period precedes the price reference period in the version of the Young and Lowe indices used by statistical offices, owing to the time it takes to develop revenue weights from establishment surveys in earlier periods. It is at this stage that the traditional index number theory discussed in Chapters 15–17 comes into play. Since this theory has been explained in detail and in depth in these chapters, which are also summarized in Sections B–E of this chapter, it is not repeated here.

1.268 At the time the monthly PPI is first calculated, the only revenue weights available must inevitably refer to some earlier period or periods. As

[13]Also recall that the Jevons index would be consistent with a geometric Laspeyres at higher levels.

mentioned above, this predisposes the PPI to some form of fixed-basket index (Laspeyres, Lowe, or Young index, or chained Laspeyres index). However, at some later date estimates must become available of the revenues in the current period, so that retrospectively it becomes possible to calculate a Paasche-type index and superlative indices such as Fisher or Törnqvist.[14] 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 this reason. Thus, although most of the discussion in Chapter 9 is based on the assumption that some type of fixed-basket index is being calculated, this should not be interpreted as implying that this is the only possibility in the long term.

P.3 Production and maintenance of higher-level indices

1.269 In practice, the higher-level indices up to and including the overall PPI are often calculated as Young indices: that is, as weighted averages of the elementary price indices using weights derived from revenues 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. This yields a Lowe index with quantities fixed for period *b*. 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 revenues. They cannot do both. Price updating the revenue weights preserves the quantities and produces a Lowe index. A Lowe index with quantities fixed in period *b* might be preferred, because it has better axiomatic properties compared with a Young index with revenue shares from period *b*.

1.270 The weights in the PPI need to be updated periodically or problems will result when a fixed set of weights is used for a very long period of time. For example, the prices of consumer durables, especially when quality-adjusted, have been falling relative to other goods, although the quantities purchased and revenue share have increased. An out-of-date set of weights would give insufficient weight to these falling prices. In the presence of rapid changes in technology or tastes, the weights need to be updated frequently and not allowed to continue for too long.

1.271 Section C.7 of Chapter 9 notes that the introduction of new weights is a necessary and integral part of the compilation of a PPI over the long run. Weights have to be updated sooner or later, and some countries actually update their weights each year. Whenever the weights are changed, the index on the new weights has to be linked to the index on the old weights so that the PPI inevitably becomes a chain index over the long term. Chapter 9 also discusses the techniques for linking series together by developing a set of linking factors (coefficients) that can be used for either forward linking or backward linking. This involves calculating the higher-level indices on both the old and new weights during an overlap period.

1.272 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, and 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.

P.4 Data editing

1.273 Chapter 9 concludes with Section D on data editing. It is included in Chapter 9 because data editing is a process that is closely linked to the actual calculation of the elementary prices indices. Data editing involves 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

[14] In fact, if a Laspeyres index is used and the revenue shares do not change much through time, a geometric Laspeyres index will approximate a Törnqvist index (Chapter 9, Section C.6).

elementary prices indices on which the quality of the overall index depends.

Q. Organization and Management

1.274 The collection of the price data is a complex operation involving extensive work by a large number of statistical office staff and respondents. 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 PPI. Appropriate management procedures are described in Chapter 12 of the *Manual*.

1.275 Price collectors should be well trained to ensure that they understand the importance of selecting the right transactions for pricing on initiation of the sample. 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 PPI is how to deal with the slowly evolving set of products. Products may disappear and have to be replaced by others, but it may also be appropriate to drop some products before they disappear if they have become quite unrepresentative. Price collectors and product analysts need appropriate training and very clear instructions and documentation about how to proceed. Clear instructions are also needed to ensure that price collectors and respondents report the correct prices when there are discounts, special offers, or other exceptional circumstances.

1.276 The price data reported also have to be subjected to careful checking and editing. Computers using standard statistical control methods can carry out many checks. It may also be useful to send out auditors to verify the quality and accuracy of reported price data. The various possible checks and controls are explained in some detail in Chapter 12.

1.277 Improvements in information technology should obviously be exploited to the fullest extent possible. For example, collectors may use establishment websites for price information; establishments can use some form of electronic data transfer to report their prices or use an Internet-based reporting system set up by the statistical office.

R. Publication and Dissemination

1.278 As noted here and in Chapter 2, the PPI is an extremely important statistic whose movements can influence the central bank's monetary policy, affect stock markets, influence wage rates and contract settlements, and so on. There must be public confidence in its reliability and 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 PPI advisory group consisting of both experts and users. Its role is not just to advise the statistical office on technical matters but also to promote public confidence in the index.

1.279 Users of the index also attach great importance to having the index published as soon as possible after the end of the each month or quarter, preferably within two or three weeks. On the other hand, most users do not wish the index to be revised once it has been published, and there can be some trade-off between timeliness and the quality of the index. For example, it would be possible to revise the index subsequently—by calculating a Fisher index when the requisite revenue data become available—without affecting the timeliness of the current index.

1.280 Publication must 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 website of the statistical office.

1.281 As explained in Chapter 13, good publication policy goes beyond timeliness, confidence, and transparency. The results must be made available to all users, within 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 must also not be subject to governmental scrutiny as a condition for their release, and the results must be seen to be free from political or other pressures. There are many decisions to be taken about the degree of detail in the published data and the alternative 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.

Appendix 1.1: An Overview of Steps Necessary for Developing a PPI

1.282 This appendix provides a summary overview of the various steps involved in designing a PPI, deriving the index structure and weighting pattern, designing the sample, establishing price collections, calculating indices, and disseminating the results. It also outlines procedures for ensuring that the price samples, index structure, and weighting pattern remain representative. These issues are discussed in more detail in subsequent chapters.

1.283 In following the steps described below, it is important to be mindful of the practical experience of national statistical agencies, which has led to the identification of several important prerequisites for the construction and compilation of an accurate PPI. That is:

- The prices recorded in the index over time must relate to:
  - (i) product specifications that are representative indicators of price change;
  - (ii) constant-quality products with fixed specifications; and
  - (iii) actual market transactions inclusive of all discounts, rebates, etc;
- The weights need to be representative of the relevant pattern of transactions over the period for which they are used for index aggregation; and
- The aggregation formulas used must be appropriate to the needs of the particular index and not yield significant bias or drift.

Basic Steps in PPI Development

1.284 Ten basic steps can be defined for the design, construction, dissemination, and maintenance of a producer price index. These steps are:

1. Determining the objectives, scope, and conceptual basis of the index;
2. Deciding on the index coverage and classification structure;
3. Deriving the weighting pattern;
4. Designing the sample;
5. Collecting and editing the prices;
6. Adjusting for changes in quality;
7. Calculating the index;
8. Disseminating the indices;
9. Maintaining samples of businesses and product specifications; and
10. Reviewing and reweighting the index.

1.285 A summary of the issues involved with each of these steps is provided in the rest of this appendix.

Step 1. Determining the objectives, scope, and conceptual basis of the index

1.286 Decisions made following close consultation with users (both external users and internal national statistical agency users such as national accounts) about the objectives of the proposed PPI, and hence its scope, have a fundamental influence on the determination of the conceptual basis of the index.

1.287 Uses range from economic policy (for example, inflation analysis), to business applications such as contract price escalation and monitoring of relative performance, industry policy formulation, and volume estimation (for example, national accounts growth estimates). All key stakeholders need to be consulted early in the index design stage to ascertain what their needs are (that is, what are the questions they are aiming to answer and, hence, the characteristics of the required statistics).

1.288 After the objectives have been determined, informed decisions need to be made about the economic scope of the index—that is, what is the domain of price transactions that the index is aiming to measure.

1.289 As discussed earlier and in Chapter 2, it is necessary to determine whether the index is to be demand based (an input index) or supply based (an output index).

1.290 Assuming it is to be supply based (which is the most common form of PPI compiled by national statistical agencies), an important consideration in defining the scope of the index is whether it should be a *net* or *gross* output index (see Chapter 2). The scope of a gross output index is broader than that of a net output index in that it also includes *intrasectoral* transactions. That is, taking as an example a manufacturing sector output index, transactions between different manufacturers would be in scope (for example, sales of refined sugar for the production of soft drinks), not just sales outside the manufacturing sector.

1.291 A further consideration is whether the scope of the index should be confined to domestic transactions only, or be broadened to include transactions with the rest of the world (exports).

1.292 Having decided on the objectives and scope of the new PPI, it is then necessary to formulate the detailed conceptual basis of the measure, again in consultation with users as necessary. Conceptual characteristics to be determined include the point of pricing, the valuation basis, coverage, and classification structure.

1.293 Decisions on the point of pricing and on the valuation basis of the index largely fall into place once the objectives and scope have been determined. As a rule of thumb, for an output (supply-based) index, the pricing point is ex-producer (for example, ex-factory, ex-farm, ex-service provider) with a valuation basis of "basic prices" (that is, reflecting the amount received by the producer exclusive of any taxes on products and transport and trade margins). On the other hand, for an input (demand-based) index, the pricing point is "delivered into store" with a valuation basis of "purchasers' prices" (that is, reflecting the amount paid by the purchaser inclusive of any taxes on products and transport and trade margins).

Step 2. Deciding on the index coverage and classification structure

1.294 The issue of the actual coverage of the domain of transactions defined by the economic scope of a PPI can be viewed from several perspectives.

1.295 Choices need to be made as to whether *nonmarket* transactions should be included or excluded. The decision will be based on a consideration of the primary objective of the index and on practical pricing considerations such as the following.

1.296 For example, for an index that aims to reflect changes in actual market transaction prices, prices of notional transactions such as changes in stocks and imputed dwelling rents have no place (in contrast to the national accounts, where conventions provide for the valuation of certain nontraded goods and services so that economic activity is not omitted). Further, it can be argued that for a price index designed primarily for the purpose of analyzing inflation, prices of commodities that are not determined on the basis of buyers and sellers interacting (that is, as a result of supply and demand forces) should be excluded because they do not provide signals of market-driven inflation. Examples include the nominal prices sometimes charged by providers of general government services (for example, health and education) and prices that are heavily subsidized through government funding or regulated by government policy.

1.297 Similarly, practical decisions need to be made about whether efforts should be expended on trying to capture price changes of goods and services transacted in the nonobserved ("hidden") economy. Issues such as the relative size of the nonobserved economy and its accessibility for price measurement need to be considered.

1.298 Other coverage issues include the treatment of intracompany transfer prices and capital formation on own account. A decision needs to be made whether these flows are to be included or excluded. If they are to be included, an assessment needs to be made about whether the book entry valuations recorded in the company accounting records are realistic in terms of being contemporary market-based estimates, or are merely notional estimates. If the latter, the preferred approach would be to assign the weight associated with these transfers to the prices obtained from businesses engaging in arms-length trading.

1.299 There are also issues of geographic coverage, in particular whether international transactions should be priced for the PPI. That is, should prices for direct imports and foreign purchases of residents used as inputs to productive activity be priced for an input index, and, on the other hand, should prices for exports and domestic purchases of nonresidents be included in an output index.

1.300 An output index can be constructed under alternative classification structures. The most common constructs are based on industry, commodity, or stage of processing. International industry classifications (for example, ISIC) and commodity classifications (for example, CPC) are available for use in index construction to ensure adherence to accepted statistical standards and facilitate international comparisons. Many countries or regions have developed local adaptations of these classifications that still conform to the underlying principles.

1.301 Formal classifications are hierarchical in nature. For example, ISIC covers the entire economic activity of an economy and provides for the progressive aggregation of data from a fine level of detail (for example, soft drink manufacturing), through successively broader levels of aggregation (for example, manufacturing of beverages; food, beverage, and tobacco manufacturing; total manufacturing). In designing an index classification structure, it is important to consider issues such as:

- *Publication goals*. In particular, the level of detail to be released, whether the indices will be national only or include regional series, and the needs of internal users;
- *Potential bias in the index due to product replacement and new goods*. There are opportunities to minimize such bias through grouping products that are close substitutes.

1.302 Having determined the index classification structure, the weighting pattern needs to be derived and issues of sample design and price collection addressed.

Step 3. Deriving the weighting pattern

1.303 A price index can be considered as being built up from samples of prices of individual (or price relatives) that are progressively weighted together through successive levels of aggregation within a classification framework.

1.304 In considering the development of an index weighting pattern, two different categories of indices need to be considered: lower-level indices (sometimes referred to as elementary aggregates) and upper-level indices.

1.305 The lower-level indices are built up by combining together the individual prices using one of a range of available price index formulas. At this initial level of aggregation, the internal weighting can be either *explicit* or *implicit*. If *explicit* weights are used, then, as part of the price collection activity, it is necessary to obtain relevant value data (for example, product sales). This is discussed further under Step 5 below. On the other hand, if *implicit* weights are used, then the design features of the sampling techniques employed to select the product specifications for pricing need to result in the prices being "self-weighted." Such a result would be achieved, for example, by using probability sampling proportional to size.

1.306 Upper-level indices are formed through weighting together lower-level indices through progressive levels of aggregation defined by the classification structure, usually employing weights that are fixed for a period (say one, three, or five years) between index reweighting.

1.307 The selection of the level in the index hierarchy at which the structure and weights are fixed for a period is particularly important. The main advantage of setting the level relatively high (for example, at the four-digit industry or product group level) is that the price statistician then has greater discretion to update the lower-level price samples (at the establishment and product level), their structure, and their internal weighting on a needs basis as market activity changes. New products and establishments can be introduced easily into the samples, and the weights at the lower level reestablished on the basis of more recent market conditions. That is, there is greater opportunity to keep the index representative through an ongoing program of sample review (see Step 9).

1.308 On the other hand, if the level is set relatively low in the index structure, there is less freedom to maintain the representativeness of the index on an ongoing basis, and there will be a greater dependence on the periodic index review and reweighting process (see Step 10). In such circumstances, the argument for frequent reweighting becomes stronger.

1.309 Assume a manufacturing output index is to be developed with the broad index structure based on ISIC. In order to derive the upper-level weighting pattern, a data source is required; potential sources include industry surveys, economic censuses, input-output tables, and international trade statistics.

1.310 The relevant values need to be assigned to each of the industry groupings, taking a top-down approach. It may be appropriate to assign the values associated with industry output that is not going to be directly priced in the index (either because it is too small, or because of practical pricing difficulties) to a related industry in order to maintain the correct broad weighting relativities. The assumption underlying this practice is that the price movements of the unpriced products are more likely to be similar to those of related products than to those of the aggregate of all the products priced in the index.

1.311 Weights aim to be representative of the pattern of transactions expected to prevail during the period for which they are used in the index construction (perhaps one year, or five years, depending on the frequency of reweighting). It may therefore be necessary to adjust some of the values to *normalize* them and overcome any irregularities in the data for the particular period from which it is being sourced (for example, as a result of a one-off increase in production of a product in response to a temporary increase in demand). Alternatively, the weights may be *smoothed* by basing them on data from a run of years (say, three years). Other adjustments may be needed to overcome problems of seasonality that are discussed in Chapter 22.

1.312 If the time reference (base price) period of the index is different from the period from which the value weights are derived, then it is important to *revalue* the weights to the prices of the time reference (base price) period using relevant price indices in order to ensure that the weights are effectively based on the underlying quantities or volumes.

1.313 Having assigned weights to the upper-level index structure that are to be fixed for a period of one or more years, the next step is to consider the lower-level index construct and the sample design.

1.314 If explicit lower-level weighting of price samples is to be incorporated, then output or sales data[15] will need to be obtained directly from manufacturing businesses during the process of establishing price collections (Step 5).

Step 4. Designing the sample

1.315 Take the example of soft drink manufacturing used in Step 2, and assume that this is an *index regimen item* with a fixed weight of, say $100 million, within the upper-level index structure. It is now necessary to choose techniques for selecting samples of businesses (statistical units) to provide transaction prices of a selection of representative products on an ongoing basis. The prices, or price movements, collected from different businesses will be aggregated to form indices.

1.316 To select a sample of businesses (for example, manufacturers of soft drinks), the first step is to identify the sample frame (that is, the population of units from which to select). Possible frame sources include registers of businesses maintained by national statistical agencies, commercially maintained lists (for example, as used for marketing mail outs), company registers, taxation records, telephone directory "yellow pages," etc., or some combination of such sources.

1.317 Either probability (scientific) sampling or nonprobability (judgmental) sampling techniques can be used,[16] and the choice may be based largely on practical considerations such as resources to develop sampling frames, data sources like a business register to develop sampling frames, and data collection resources to undertake the required intensive efforts to recruitment establishments. Some agencies use a combination of techniques, for example scientific sampling to select the businesses and judgment sampling to select the product specifications for pricing.

1.318 In deciding how to select the sample of businesses, the degree of industry concentration is a relevant consideration. For example, in a highly concentrated industry dominated by, say, three businesses producing over 90 percent of the output, it may be acceptable to aim for high, rather than complete, coverage and to select only the three largest businesses.

1.319 However, as the degree of concentration decreases, the greater is the need for the sample to include a selection of smaller businesses. If, for example, the three largest businesses account for less than 70 percent of the industry output, with the remaining 30 percent being produced by a large number of small businesses, it may not be possible to

[15]As discussed in Chapter 4 and subsequently in Chapter 5, output data, which include sales and inventories, are preferred for weights. However, in many countries data on output are not readily available, and for practical reasons a close proxy such as sales or value of shipments may be used.

[16]Judgmental sampling should be avoided, if possible. Often cutoff sampling, as discussed in Chapter 4, Section D, can be used in place of judgmental sampling.

achieve adequate representation of price movements by relying only on prices reported by the three largest businesses. That is, it may not be reasonable to assume that the pricing behavior of the small businesses mirrors that of the large ones, because, for example, they may target separate niche markets and direct their pricing strategies accordingly. Therefore, it would be prudent to select a sample of the small businesses to represent the markets they serve.

1.320 The less concentrated is the industry structure, the stronger is the case for using probability sampling techniques. Experience has shown that, although many manufacturing and mining industries may be dominated by a few large businesses, many service industries have a very large number of small businesses, and, if there are any large businesses, they produce a relatively small proportion of the output. An added advantage of probability sampling techniques is that they enable sampling errors to be calculated, which provide some guide to the accuracy of the resultant indices.

1.321 Procedures need to be implemented to ensure that samples of businesses remain representative through, for example, regularly augmenting the sample by enrolling a selection of new businesses as they enter the market. Also, a sample rotation policy needs to be considered in order to spread the business reporting load.

1.322 Once the sample of businesses has been selected, they need to be contacted to agree on a sample of representative product specifications for ongoing price reporting. This is discussed further under Step 5.

Step 5. Collecting and editing the prices

1.323 The main source of ongoing price data is usually a sample of businesses. The sample can relate to either buyers or sellers, or a combination of both. The choice will be influenced by the pricing point of the index (input or output) and practical considerations such as the relative degree of concentration of buyers, and of sellers, and the implications for sample sizes and costs.

1.324 The statistical units to be sampled may be head offices reporting national data, establishments reporting regional data, or a mixture. Decisions on the units to be surveyed may be based largely on pragmatic grounds such as efficiency of collection, location of relevant business records, etc.

1.325 The aim of the price collection is to enable the calculation of reliable indicators of period-to-period—say, monthly—price change. As such, choices need to be made as to the type and frequency of pricing. For example, point-in-time prices may be the easiest to collect and process (for example, transaction prices prevailing on a particular day, say the 15th of the month) and commonly prove to be reliable indicators. For workload management, it may be decided to spread pricing over the reference period with, say, three or four pricing points and different commodities priced on different days.

1.326 For commodities with volatile prices, it may be necessary to price them on several different days of the month and calculate time-weighted averages; alternatively, businesses can be asked to provide weighted average monthly prices (usually derived by dividing the monthly value of product sales by the quantity sold). This approach should be avoided because it is susceptible to the unit value "mix" problem, where products of different qualities are included.

1.327 The most appropriate pricing methodology to use is *specification pricing*, under which a manageable sample of precisely specified products is selected, in consultation with each reporting business, for repeat pricing. In specifying the products, it is particularly important that they are fully defined in terms of all the characteristics that influence their transaction prices. As such, all the relevant technical characteristics need to be described (for example, make, model, features) along with the unit of sale, type of packaging, conditions of sale (for example, delivered, payment within 30 days), etc. This technique is known as *pricing to constant quality*. When the quality or specifications change over time, adjustments must be made to the reported prices (see Step 7).

1.328 Another important consideration in establishing and maintaining price collections is to ensure that the prices reported are *actual market transaction prices*. That is, they must reflect the net prices received (or paid) inclusive of all discounts applied to the transactions whether they be volume discounts, settlement discounts, or competitive price-cutting discounts, which are likely to fluctuate with market conditions. Any rebates also need to be considered. The collection of nominal list prices, or book prices, is not reflective of actual transactions, is unlikely to yield reliable price indices, and may result in quite misleading results because fluctuations in market prices are not captured.

1.329 The principles underlying the selection of the sample of product specifications from a particular business, whether using probability or nonprobability sampling, are similar. That is, the outputs of the business and the markets are stratified into categories with similar price-determining characteristics. For example, in selecting a sample of specific motor vehicles in consultation with the manufacturer, the first dimension may be the broad category of vehicle (for example, four-wheel-drive recreational vehicles, luxury cars, family cars, and small commuter cars). These categories will reflect different pricing levels as well as different pricing strategies and market conditions. A further dimension may be to cross-classify by the type of market (for example, sales to distributors, fleet sales, and exports).

1.330 Then, from each of the major cells of the matrix of vehicle category by market, a sample of representative vehicles can be selected, with each one representing a broader range of vehicles.

1.331 If *explicit* internal weights are to be used in the construction of the lower-level indices (for example, for motor vehicles), then the relevant sales data for (i) the individual vehicles in the sample, (ii) the wider range of vehicles being represented (that is, as defined in the matrix of vehicle category by market), and (iii) all vehicles should be collected from the business for a recent period. This will enable internal weights to be calculated for combining the prices of individual product specifications and the prices of different producers.

1.332 Ideally, initialization of a collection from a business will be undertaken through a personal visit. However, this is an expensive exercise, and budgetary considerations may necessitate compromise. Alternative, though less effective, approaches to initialization include the use of telephone, Internet, fax, and mail, or some combination of approaches. At a minimum, the larger businesses and those producing complex (for example, high-tech) products and operating in changing markets should be visited.

1.333 In cases where the products are unique and not reproduced over time—for example, construction industry output and many of the customized business services—specification pricing is not feasible, and alternative pricing techniques must be used, often involving compromise. Possibilities include model pricing, collecting unit values for reasonably homogeneous components of a good or service, input pricing, and the collection of charge-out rates (for example, for a legal service).

1.334 Most national statistical agencies use mail questionnaires to collect their producer prices. Collection procedures include the design of tailored forms incorporating the particular product specifications for each sampled business and collection control to facilitate the dispatch, mark-in, and follow-up with the reporting businesses.

1.335 It is important that rigorous *input editing* techniques are employed, and that any price observations that do not appear credible are queried (usually by telephone) and either confirmed with an acceptable reason or amended. Input editing involves analyzing the prices reported by an individual business and querying large changes (editing tolerances may be built into processing systems) or inconsistent changes across product lines. An important objective of the editing process is to ensure that actual transaction prices are reported, inclusive of all discounts, and to detect any changes in the specifications.

1.336 If the price of a product has not changed for, say, six months, it may be appropriate to contact the business to make sure the prices reported are not being automatically repeated.

1.337 *Output editing*, which is often an integral part of calculating the lower-level indices (see Step 6), involves comparing the price levels, and price movements, of similar products between different businesses and discretely querying any outliers.

1.338 In undertaking these editing processes, reference to other supporting price information is often valuable. Examples include international commodity prices (for example, London Metal Exchange), exchange rates, press and wire service reports, and general market intelligence obtained during the sample maintenance activities described under Step 9.

1.339 Alternatives to the traditional mail questionnaire include telephone, e-mail, Internet, telephone data entry, fax, and the use of electronic data transfer from company databases. Several national statistical agencies have had experience with at least some of these alternatives. Important factors to be considered are data security, the convenience of reporting for the business, cost, and effectiveness.

Step 6. Adjusting for changes in quality

1.340 The technique of *specification pricing* was outlined under Step 5. The objective is to *price to constant quality* in order to produce an index showing *pure price change*. This is the most common technique employed by national statistical agencies in compiling PPIs.

1.341 To the extent that pricing is *not* to constant quality, then, over time, the recorded prices can incorporate a nonprice element. For example, if a product improves in quality and its recorded price does not change, there is an effective price *fall* because an increased volume of product is being sold for the same price. Conversely, if the quality of a product declines without a recorded price change, there is an effective price *rise*. In such circumstances, the recorded price of the new product of changed quality needs to be adjusted so that it is directly comparable with that of the old product in the previous period.

1.342 Failure to make such adjustments can result in biased price indices and consequently biased constant price, or volume, national accounts estimates.

1.343 It is possible to identify fairly readily the main price-determining characteristics of many goods (for example, a washing machine) that are mass produced to fixed technical specifications and can be readily described in terms of brand names, model codes, etc. However, specification pricing cannot be used for customized goods such as the output of the construction industry. Nor can it be used for much of the output of business service industries (such as computing, accounting, and legal services) because it is unique in nature (each transaction is commonly tailored to the needs of an individual client). Further, it is far more difficult to identify all the price-determining characteristics of many services because some are intangible.

1.344 In such cases, other approaches to pricing to constant quality must be employed—for example, model pricing—using narrowly defined unit values or collecting charge-out rates (see Step 5).

1.345 Even in areas that do lend themselves to specification pricing, problems arise when there are *changes* to the specifications, and hence the quality, of the products over time. Examples of possible product changes would include:

- Presenting it in new packaging;
- Selling it in different size lots (for example, 1 kg packets of sugar replaced with 1.2 kg packets); and
- Replacing it with a product with different technical and design characteristics (for example, a new model of motor vehicle).

1.346 The first step, in consultation with the provider, is to fully identify the changes and assess whether they are, in fact, quality changes.

1.347 The first example above (new packaging) may be deemed to be cosmetic only; alternatively, it could be assessed as being substantive if, for example, it led to a reduction in the damage to the contents. In the latter case, a value would need to be placed on the improvement on the basis of some estimate of the value of reduced damage.

1.348 The second example (change in size lot) would be likely to involve an office adjustment based on matching the new and old prices per a common unit of measure (for example, price per kilogram).

1.349 The third example (new model of motor vehicle) is the most complex. Possible techniques include using an assessment of the difference in the production costs of the old and new models to adjust the price of the new model. Alternatively, the different product characteristics can be identified and a value placed on them. The valuation can be based on consultation with the producer or, if the new model has features that were available as options on the old one, market prices will exist for those options. In cases where the old and new model are sold (in reasonable volume) in parallel, the difference in the overlapping transaction prices may be taken as a guide to the value of the quality difference.

1.350 Increasingly, national statistical agencies are researching and selectively implementing hedonic regression techniques as a means of placing a market value on different characteristics of a product—for example, the value of an additional unit of RAM on a personal computer. When the characteristics of a particular product change, these techniques enable its price to be adjusted to make it directly comparable with that of the old model. Unfortunately, hedonic techniques tend to be very costly, involving extensive research and analysis and the collection of large volumes of data.

Step 7. Calculating the index

1.351 Under Step 3, the two categories of indices were described: lower-level and upper-level indices. Having established the structure and weighting pattern of the index, constructed a processing system, and established the regular price collection, the first step in the routine production cycle is to aggregate the input-edited prices to form the lower-level indices. There is a range of micro-level index formulas available for use, each being based on different assumptions about the relative behavior of prices and quantities in the economy (see Chapters 15 and 17).

1.352 The initially compiled lower-level indices should be scrutinized for credibility in terms of the latest period movement, the annual movement, and the long-term trend. *Output editing*, involving comparisons of price levels and movements between different businesses, is an integral part of the credibility checking. Reference to the type of supporting information described under Step 5 will be valuable for this analysis.

1.353 Despite the most rigorous collection processes, there are often missing prices that need to be imputed. Prices may be missing either because the provider failed to report on time or because there were no transactions in that product specification in the relevant period. Imputation techniques include applying the price movements of like products to the previous period price observations. The like products may either be reported by the same business or by other businesses. Another approach is to simply repeat the previous period prices, but this approach should be used only if there is reasonable certainty that the prices have not changed.

1.354 Once the price statistician is satisfied with the lower-level index series, the series should be aggregated to form the hierarchy of upper-level indices, including the total measure. This aggregation is undertaken using the classification structure and weighting pattern, determined in Step 2, and an appropriate index formula.

1.355 Extensive studies have concluded that the theoretically optimal formulas for this purpose satisfy a range of tests and economic conditions, and as a class are known as superlative formulas (Chapter 15). A basic characteristic of such formulas is that they employ weights based on volume data from both the current period and the period of index comparison. In practice, since the volume data from the current period are not available at the time of index construction, the use of a superlative formula would necessitate the estimation of the current-period volume data in order for timely indices to be produced. When the current-period volumes subsequently became available, the index numbers would need to be recompiled using the actual volumes, and the earlier index numbers revised. This ongoing cycle of recompilation and revision of published index numbers would create a high degree of uncertainty among users (as explained under Step 8) and is highly undesirable. Therefore, most national statistical agencies compromise and use a base-weighted formula such as Laspeyres.

1.356 The upper-level indices are aggregated across industries, commodities, and/or stages at the national or regional level, as defined in Steps 2 and 3, to produce the aggregates required for publication (Step 8).

1.357 Finally, annual average index numbers and the suite of publication and analytical tables should be produced and the commentary on main features prepared for publication (see Step 8). It is prudent to apply broad credibility checks to the aggregates before release. Are the results sensible in the context of the prevailing economic conditions? Can they be explained?

Step 8. Disseminating the indices

1.358 During the initial user consultation phase described under Step 1, and the formulation of the index classification structure under Step 2, broad publication goals will have been formulated. It is now time to refine and implement these goals, probably involving further user contact.

1.359 As well as releasing time series of index numbers for a range of industries or commodities or stages, and aggregate measures (for example, all groups), user analysis can be enhanced by the release of time series of percentage changes, as well as tables presenting the contribution that individual components have made to aggregate index point changes. This latter presentation is particularly important to help gain an understanding of the sources of inflationary pressure.

1.360 Different tabular views of the data can be provided. For example, classification by:

- Source—imported or domestically produced;
- Economic destination—consumer or capital goods; and
- Industry and/or commodity.

1.361 Some form of analysis of the main movements and, ideally, the causes of those movements, should be provided. These will be based on the percentage change and point-contribution tables described above.

1.362 In addition to the summary tables, analytical tables, and detailed tables, explanatory notes should outline the conceptual basis of the index, including the objectives, scope, coverage, pricing basis, sampling techniques, and data sources. The weighting patterns should also be published. Any caveats or limitations on the data should be included to caution users.

1.363 As well as release in hard copy form, electronic delivery and access through the Internet website of the national statistical agency should form part of the overall dissemination strategy.

1.364 In terms of timeliness of release, there will be a trade-off between accuracy and timeliness. In general, the faster the release, the lower the accuracy of the data, and hence its reliability, as the need for revisions increases. Price index users—whether they be public policy economists, market analysts, or business people adjusting contract payments—place a high value on certainty (that is, the nonrevisability of price indices). Accordingly, some compromise in the timeliness of release will probably need to be made in order to achieve a high degree of certainty and user confidence.

1.365 Policies need to be developed in relation to:

- Security of data through the uses of a strict embargo policy;
- Publication selling prices and electronic access charges based on relevant principle—for example, commercial rates, cost recovery, or rationing of demand; and
- Community access to public interest information—for example, through free provision to public libraries.

1.366 Ongoing consultation with users should be maintained to ensure that the indices, and the way they are presented, remain relevant. The establishment of a formal user group, or advisory group, should be considered.

Step 9. Maintaining samples of businesses and product specifications

1.367 As noted in Section R above, some of the necessary prerequisites for the production of an accurate price index are to incorporate prices that, over time, relate to:

- Product specifications that are representative indicators of price change;
- Constant quality products with fixed specifications; and
- Actual market transactions inclusive of all discounts, rebates, etc.

1.368 Step 5 above expands on these principles and outlines the methodology for selecting the sample of product specifications from a business at initialization, preferably by personal visit.

1.369 Given the dynamics of many marketplaces in terms of changing product lines and marketing strategies, it is important that procedures are put in place to ensure that the product samples remain representative and have fixed specifications, and that the prices reported incorporate all discounting.

1.370 Further, if explicit internal weighting is used in the lower-level index aggregation, these weights need to be monitored and updated as necessary, on a component-by-component basis.

1.371 Ideally, a rolling program of regular interviews of the sampled businesses would be established to undertake these reviews on a fairly frequent basis. Costs may prohibit regular visits to all of the businesses, so it may be necessary to prioritize them according to factors such as their weight in the index, the extent of technical change in the industry, and the volatility of the markets. A program may be devised such that the high-priority businesses are visited frequently and the lower-priority ones visited less frequently and/or contacted by telephone. Many national statistical agencies have such structured programs in place.

1.372 In addition to these structured proactive reviews, resources should be made available to enable a quick reaction to changed circumstances in relation to a particular commodity or industry and to undertake specific reviews on a needs basis. For example, competitive pressures resulting from deregulation of a particular industry may quickly, and radically, transform the product lines and methods of transacting and produce substantial market volatility. Examples in recent years include the deregulation of the electricity supply, telecommunications, and transport industries in many countries.

1.373 The samples of businesses also need to be reviewed, either through a formal probability-based sampling process incorporating a rotation policy, or some more subjective approach that includes initialization of price collections with substantial new businesses as they enter the market.

Step 10. Reviewing and reweighting the index

1.374 Other necessary prerequisites for the production of an accurate and reliable price index that were listed in Section R above are that:

- The weights need to be representative of the relevant pattern of transactions over the period for which they are used for index aggregation; and
- The aggregation formulas used must be appropriate to the needs of the particular index and not yield significant bias or drift.

1.375 Studies have concluded that, in practice, price indices are often not highly sensitive to small errors in weighting patterns. However, the greater is the variation in price behavior across different commodities, the more important are the weights in the production of an accurate measure of aggregate price change.

1.376 Assuming that a rolling sample review program is in place for the maintenance of price samples and the lower-level internal weights (see Step 9), then the question of the frequency of reweighting of the upper-level indices (which were established under Step 3) needs to be considered. Alternatively, if no such sample review program is in place, a strategy needs to be put in place for the periodic reweighting of the entire index (lower and upper levels) along with a complete review of the product samples.

1.377 Practices in this regard vary among national statistical agencies. Some agencies update the upper-level weights on an annual basis and link the resultant indices at the overlap period such that there is no break in continuity of the series. That is, if the link was at June 2000, then the “old” weights would be used to calculate the index movements between May and June, and the new weights used to calculate the index movements between June and July (and subsequent months), with the July movements “linked” onto the June level. This process is termed annual *chaining* or *chain linking*.

1.378 The more common national practice is to reweight and chain on a less frequent basis, perhaps once every three or five years. Considerations in making decisions on the frequency of reweighting include:

- Changes over time in the pattern of transactions covered by the index:

  (i) The greater the volatility in the transaction patterns, the greater the need for frequent reweighting to maintain the representativeness of the weights. If the trading patterns are highly volatile, it may be desirable to “normalize” or smooth them by using data from a run of years in order to mitigate against *chain-linking bias* or *drift*,
  (ii) If the trading patterns are relatively stable and tend to shift on a trend basis, very frequent reweighting is of little benefit, and it may be assessed that reweighting every three, five, or more years is adequate;

- The availability of reliable and timely weighting data sources; and
- Resource constraints.

1.379 If reweighting is done on an infrequent basis using data from a single year, it is important that a *normal* year is selected in terms of providing weights that can be expected to be representative of the period (say, five years) for which they are used in the index. Again, the use of data from a run of years may be prudent.

1.380 In addition to developing a reweighting strategy, it is desirable to undertake thorough periodic (say, every five or ten years) reviews of the PPIs to ensure that the conceptual basis is still relevant to the needs of users.

Summary

1.381 Early consultation with users and decisions on the scope and conceptual basis of a PPI are fundamental to the production of a *relevant* index. In order for the index to be *accurate*, it must be constructed using indicative transaction prices (measured to constant quality) and representative weights.

1.382 The issue of reporting burden is an important consideration in seeking the cooperation of businesses and, along with resource constraints facing national statistical agencies, heavily influences decisions on sampling strategies and other methodological matters. Ensuring the security of often commercially sensitive price data is another essential prerequisite to building good business relationships.

1.383 A dissemination strategy that meets the needs of the wide variety of users must be developed, and ongoing consultation maintained, to ensure to insure that users' requirements continue to be met.

1.384 It is important to appreciate that a price index seeks to provide contemporary information in relation to dynamic markets. As such, it is not sufficient to develop a new index framework, establish the collection of the price samples, and simply aggregate them over time. Mechanisms need to be put in place to ensure the ongoing integrity and representativeness of the measure. That is, the price samples and weights need to be systematically reviewed and updated periodically.