PART II

Compilation Issues

4. Weights and Their Sources

A. Introduction

4.1 As an index number, the PPI is computed as an average of the price relatives of the many products for which prices are collected. The average is weighted to reflect the importance of each priced product in terms of its share of total output of the establishment.[1] Ideally a weight should be attached to each price collected. However, as noted in Chapter 5, this is not always feasible or cost-effective.[2] This chapter explores the statistical issues underlying the determination of weights. It outlines the objectives and criteria for determining weights, describes and evaluates the varying data sources that are traditionally used to generate the weights, and suggests some additional sources and methods for deriving the weights. Finally, it describes how the weighting might be accomplished in practice.

B. Role of Weights

4.2 The PPI is calculated from many prices collected from all types of establishments, covering the selected economic activities and products. The collected prices are first combined to compile indices for each individual product. For example, 10 prices for different types of transactions for a product may be collected from an establishment, and these prices are combined to produce the index for the product from that establishment. Weights are usually not available for these individual transactions, and the establishment's product index is thus computed as an unweighted average of the prices collected for the various transactions. Once this has been done, the establishment product indices are combined to produce the subgroup and group indices, and eventually the all product index (see Figure 4.1 in Section C.4). Because some products have greater production or sales than others, each product is given a weight to represent its importance in total output or sales during the reference (base) period for the weights. To arrive at the aggregate index figure, the price relatives of the individual products are multiplied by these weights to derive a weighted average aggregate index.

4.3 Thus, the weights are key elements in the construction of a PPI. They determine the impact that a particular price change will have on the overall index. For example, in some countries, a 5 percent rise in the price of milk products would have a much greater impact on the average rate of price change in the producer sector than a 5 percent increase in the price of tea products because the output value of milk is higher than that for tea. Without weights, relative price changes for all commodities in the PPI basket would be given equal importance in the calculation of the index above. Of course, if there is no dispersion of price changes, then weights would be unimportant.

4.4 Over time, establishment production levels shift in response to economic conditions. Some products and industries become more important while others become less important. Statistical offices periodically should update the weights in the PPI to reflect these changes in market structure. Best practice suggests that this be done at least once every five years. Details on how to introduce new weights into the PPI appear in Chapter 9, Section C.

[1] As noted in Chapter 5, when it is feasible to implement probability sample designs for selection of elementary products, the weight may also reflect the fraction of total output a sampled product represents among the totality of transactions in an economic activity or product class that are produced by businesses.

[2] Referencing the previous footnote, probability samples generally are not implemented except in large, advanced statistical systems. In the absence of such probability designs, elementary product weights may be judgmentally determined. In the interest of adopting transparent, reproducible procedures under judgmental sampling, elementary product weights generally are taken to be equal within an industry or product classification that is to be represented in the index. Equal weighting may also be implied by certain simple randomized sample designs, such as simple random sampling with replacement.

C. Appropriate Weights and Structure for PPIs

C.1 Value weights

4.5 As discussed in Chapter 14, the value aggregate from the national accounts framework that aligns with the basic price received by the producer of goods and services is the value of production. Thus, when estimating the PPI using the weighted average of long-term relatives formula (that is, the current price divided by the base-period price as in equation [4.2] of Section C.2), the best approach would be to have value of production weights at basic prices for all levels of index aggregation (from the elementary aggregate level of product/commodity within the establishment to the total output index by industry or product).

4.6 Since the PPI can also be used to measure the change in intermediate input prices, the value weights for the input index would be the cost of the input products to the producer. In the supply-and-use framework presented in Chapter 14, this value would be the cost of intermediate inputs valued at purchaser prices.

4.7 The use of values to weight long-term price relatives (that is, the current price divided by the base-period price) maintains the fixed-quantity relationship that existed in the base period. The value weight multiplied by the long-term price relative provides the estimate of what it would cost at today's prices to produce the quantity of product in the price reference period.

4.8 The value of production comprises the receipts from sales of all output by establishments and the change in value of inventories of finished goods on hand at the end of the period. If the value of production is unavailable or questionable because of concerns about the estimation of inventories, total sales (turnover) may be used. An analogous measure would be the value of shipments (that is, value of goods shipped at basic prices).

C.2 Quantity weights

4.9 In the traditional Laspeyres formula, base-period quantities can be used as weights to value base-period production volume at current-period prices. Consider the following:

$$(4.1)\quad I_L^{c,m} = \frac{\sum P_i^m q_i^0}{\sum p_i^0 q_i^0},$$

where $I_L^{c,m}$ is the Laspeyres price relative for subcategory "c" in month "m,"
P_i^m is the average price of product "i" in month "m,"
Q_i^0 is the quantity of product "i" purchased or sold in the base period "0," and
P_i^0 is the average price of product "i" in the base period "0."

The value in the numerator is often referred to as the current value of base-period production. It reflects what the cost would be at current prices to produce the quantity of output in the base period. This current value of base-period production is compared with the base-period value of production in the denominator to derive the long-term price relative.

4.10 The use of quantity weights is appropriate as long as the same specific product was produced as in the base period, that is, there is no qualitative difference between the current product produced and the base-period product. If the price-determining characteristics among the various transactions that are priced differ, then we have a dissimilarity, and the transactions with different characteristics should have separate weights.

4.11 Quantity weights are feasible only at the detailed product level. At higher levels of aggregation, such as at the product group level or industry level, a value aggregate is more appropriate for calculating the index because there are no unique, meaningful quantity levels available that apply to different products.[3] Thus, the index at the aggregate level would be the ratio of the sum of the base-period quantities valued at current prices to the sum of the base-period values, as in equation (4.1), but the values in the numerator are those summed from the calculation of values for each of the products at current prices. Alternatively, the simpler formula-

[3]This holds true unless one is willing to accept a notional or implicit quantity measure that is a representative aggregate of the different quality products being compiled. The problem with this approach is that the implicit quantity measure then must assume some type of average quality that should be comparable over time.

tion is to use a base-period value weighted average of price relatives such as

$$(4.2)\quad I_L^{c,m} = \sum \left(\frac{p_i^m}{p_i^0} \right) \frac{p_i^0 q_i^0}{\sum p_i^0 q_i^0} .$$

C.3 Net output weights

4.12 The output of one activity is often used as input to another activity within the same industrial grouping, as discussed in Chapters 2 and 17. The use of gross value weights for both activities would result in double-counting because the value of output in the first activity (for example, raw materials) is an input to the second (assembled goods). The value of output of the second activity, therefore, includes that of the first. If the two activities are aggregated to produce a group index, the importance of the first activity is counted twice in the group index. To eliminate this double-counting effect, net weights can be derived.

4.13 One of the principal uses of price indices is to analyze the price change faced by buyers of particular commodities. Such analysis may not be a problem at a detailed level because product price indices are particularly useful for this purpose. For example, a change in the index for "primary aluminum ingot, alloyed" is easily interpreted by a buyer of this product. However, the interpretation of price changes that involve various products or different industries may not be straightforward if they include the effects of overweighting. For example, if the basic price of aluminum increases, how should one interpret a metals products index that includes various types of aluminum at different stages of production? To interpret this aggregate index correctly, it is necessary to know how the various elements in the index have been combined, including specifically how these elements have been weighted together to form the higher-level index.

4.14 Using a weighting scheme based on net output weights eliminates double-counting when aggregating. However, before the net output weight can be defined, it is necessary to define the aggregation structure. It is the aggregation structure that determines which prices should be counted. Only then can the weight structure be identified and the value for each component determined. Thus, the process of constructing net output weights involves two steps:

(i) Define the aggregation of interest in such a way that it is possible to identify the portion of the products produced within the aggregation that is sold to buyers outside of the aggregation.

(ii) Assign weights to the products produced within the aggregation that reflect only the value of products sold to buyers outside of the aggregation. These weights are termed **net output weights** because they include only the value of output for products exiting the aggregation, that is, the net output.

4.15 When this type of weight structure is used, price movements of products are included only to the extent that the products are sold outside of the aggregation structure. Thus, each aggregate index can be viewed as a measure of price change for buyers of the final products from enterprises included in the aggregation structure.

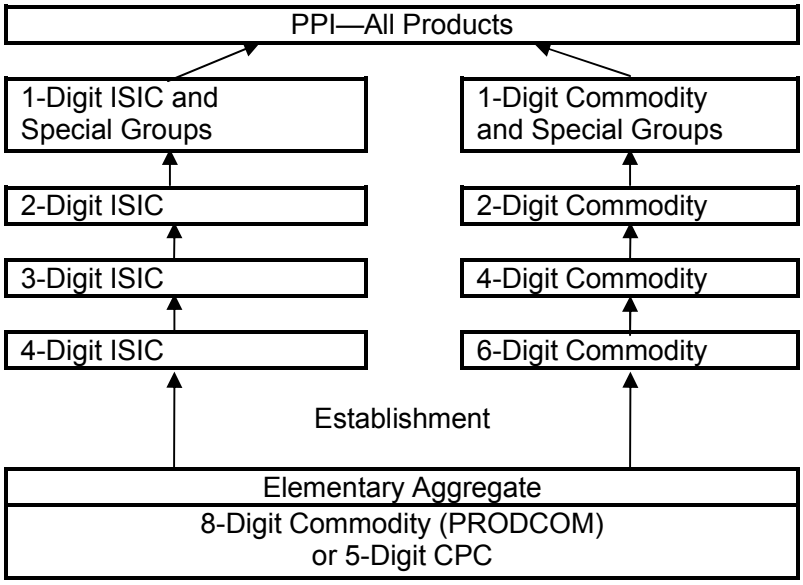
4.16 In many countries, net output weights are used to develop aggregate indices by the processing stage. In such aggregations, the weights used for products sold for final demand exclude the value of goods used as intermediate inputs. This approach avoids the problem of giving too much importance to price changes of intermediate goods as they wend their way through the production process.

C.4 Classification issues

4.17 For the purpose of applying the weights, products are grouped either because they have a common end-use or because they are considered substitutes for one another.[4] These families of products are joined at different levels to form a hierarchy in a classification system. Every product has a unique place in the classification used. Such criteria were used when the International Standard Industrial Classification of All Economic Activities and other classifications were established.

[4]Alternatively, some groupings are made for products that exhibit common price trends. Such groupings are important when a product index or group index is used to make imputations for missing products, which is discussed in Chapter 9.

Figure 4.1. Typical PPI Aggregation Structure



4.18 For the purposes of international comparison and internal consistency, the classification scheme of goods and services should be in line with the most recent version of the Central Product Classification, version 1.1 (CPC), or the Classification of Products by Activity. In terms of economic activities, the establishments should be classified using ISIC, Revision 3, the General Industrial Classification of Economic Activities within the European Communities (NACE), Revision 1, or a derivative of these industrial classification systems. From an individual country's perspective, it is also desirable that the classification used be consistent across all enterprise and production statistics (for example, establishment census and industrial survey statistics).

4.19 Each product selected for inclusion in the PPI is assigned a product code in accordance with the product classification system. Likewise, each establishment selected in the sample is assigned an industry code according to the industrial classification system. Subindices by product are computed from groupings of the selected transactions in accordance with the product classification system. The selected sample transactions can also be aggregated according to the industrial classification system to produce indices by economic activity. These subindices are further aggregated following the hierarchy of the classification systems to arrive at major groups or divisions and, finally, the all products index as shown in Figure 4.1. Because inputs may be overweighted in the aggregation to derive higher-level indices, the statistical agency may choose to use net output weights as discussed in the previous section.

4.20 This aggregation starts with the sample of specific product transactions selected within establishments. The transaction prices or price relatives are combined using the price index formula to arrive at the first level of index aggregation, which is referred to as the elementary aggregate or the elementary index. Weights are often not collected below the first level of aggregation. This aggregate is usually for specific types of products within the product classification. In our example, we use the eight-digit product code level. At all subsequent levels (establishment, six-digit product code, etc.), it is necessary to obtain a consistent set of aggregation weights. For example, the weights for the sample of establishments should cover the entire four-

digit industry even though not all establishments were selected. This means that the weights for the nonsampled establishments must be assigned to those selected. Also, the weights for the products selected within an establishment should include the entire weight for the sampled establishment. Once the weights have been established at these levels, it is a relatively straightforward procedure to aggregate by industry or product to higher-level aggregates.

C.5 Unimportant industries and products

4.21 Some industries and products will be of little importance in terms of their share of total production. For example, an industry that represents less than 0.1 percent of production within the industrial or service sectors could be excluded from the sample. In such cases, the output for the industry that is excluded should be distributed across those that were selected, or it should be assigned to a closely related industry. It may also be possible to make meaningful combinations of smaller industries producing related products that meet the criteria for minimum sizes. A similar procedure would also be applied to products that are insignificant. In either case, the weight for the nonsampled component needs to be included somewhere in the weighting structure.

4.22 A situation that will occur is having an important industry or product that falls below the size threshold chosen. In such an instance, if no meaningful combinations are apparent, the industry or product may have to be published on its own. This often is the case for growth sectors where industries and products are expected to become more important over time. The statistical office will want to include them because their contribution to economic activity will become significant before the next scheduled weight update.

C.6 Time period covered by weights (weight reference period)

4.23 The weight reference period is the time period—usually a whole year—to which the weights relate. The accuracy and reliability of a PPI are determined, in large part, by the weighting structure. For this reason, the choice of the period covered by the weights is crucial. The period chosen as weight reference period should be (i) reasonably normal/stable and (ii) not too distant from the price reference period.

4.24 The weight reference period and the price reference period used in the index formula should refer to the same period. When they differ, the weights should be updated for price changes between the weighting period and the price reference period. For example, if the weights refer to calendar year 2001 and the base price is for December 2001, the weights should be adjusted for the change in prices between the average price for the calendar year and the price in December. This is discussed further in Chapter 9.[5]

4.25 The weights may be chosen from multiple periods depending on the formula used to calculate the index. In Chapter 15, it is recommended that a symmetric index be used, which requires weights for the base period and the current period. In practice, weights are often not available for the current period on a sufficiently timely basis; base-period weights, therefore, are normally used. For example, the weights may represent (i) the value of output produced during the price reference period (Laspeyres index), (ii) the value of output produced during the current period (Paasche index), or (iii) a geometric average of the values in base and current period (Fisher or Törnqvist index). An index computed by using quantity or value weights for the current period can be produced only with a time lag, because it takes time to collect and process current production data. That is why most statistical offices adopt a Laspeyres-type index, which requires quantity or value weights for the price reference period only.

4.26 The weights that are used typically refer to a single calendar year. In some instances, a single year's data may not be adequate either because of unusual economic conditions or insufficient sample

[5] There is also an index reference period, which is the period when the price index is equal to 100. In many countries, the weight reference period, the price base period, and the index reference period are the same. More and more frequently, however, countries are introducing chained indices in which the weights are updated on an annual basis. In such cases, the three periods can be different. For example, weight reference period could be the previous year (2001), the base price period could be the previous December (December 2001), and the index reference period could be maintained as December 2000 = 100. This is discussed in more detail in Chapter 9.

sizes from survey data. An average of several years' data may provide the best weight reference period because it reduces the sampling and seasonal variance of the production or sales data for a given size of the annual sample.[6]

4.27 For seasonal products (as discussed in Chapter 22), it may be preferable to develop separate weights by month or by quarter to calculate indices at the elementary aggregate level. In addition to information for each period within the year, this approach may require additional data for the same period in a number of previous years.

D. Elementary Aggregate or Stratum-Level Weights

D.1 Coverage of weights

4.28 The calculation of the all industry or all products indices starts with the measurement of the relative price change for an elementary aggregate, which represents the first level at which price observations are combined to calculate an index. At this level, sometimes referred to as a stratum, weights are needed to calculate higher-level indices. This typically involves combining individual-product-level or establishment-level indices to derive product groups industry indices. The elementary aggregate index covers all prices collected for one product in one stratum. The stratification may be by product, industry, size of establishment, or some combination of these.

4.29 It is important that the weight for each elementary index represents the production value of *all* products produced within the stratum, not just the value for the selected sample of particular products at particular establishments chosen to represent this aggregate. (Chapter 5 deals with ensuring that the sample of elementary products is representative.)

4.30 Below the elementary index level, individual transaction prices may not have weights because the statistical office has not collected additional data on production or sales for sampled products within establishments.[7] If no weights are available, depending on the formula used (see Chapter 9), it is assumed that all the weights are equal (an average of price ratios approach) or the weights are proportional to their base-period prices (a ratio of average prices approach).[8] The former means that each price quote within the elementary aggregate is as important as any other price, that is, the shares of production value are equal. In the case of the ratio of average prices, the importance of each price quote depends on its price level in the base period and the fact that all the quantities produced are equal. This is appropriate if the production value in the base period is proportional to relative price levels in the base period. Thus, items with higher prices in the base period have more importance.

4.31 Once the price indices for the elementary aggregates are computed, the product/industry indices are obtained as weighted combinations of the indices for each elementary aggregate. Then the product indices are combined following the hierarchy of the classification, with appropriate weights applied along the way. For instance, assume the elementary aggregate is established at the eight-digit product code level (as in Figure 4.1). All transactions within this classification are used to estimate the eight-digit product index. Each eight-digit product index has an assigned weight, and the indices are aggregated to produce the six-digit product group level index. All six-digit product group indices are further aggregated using production weights at the six-digit level to obtain four-digit level indices and so on, until the all products index is obtained. In addition, the eight-digit product indices can be aggregated to derive industry-level indices, and industry indices can be aggregated according to the industrial classification structure to derive group- and division-level indices.

[6] During periods of high inflation, multiple-year weights should be calculated by averaging value shares rather than averaging actual value levels. Averaging the value levels will give more importance to the more recent years' data.

[7] The situation in the United States is somewhat different since compilers use probability sampling, where the weight within the elementary aggregate is determined by the inverse of the probability of being selected in the sample.

[8] The average of price ratios formula is $\frac{1}{n}\sum\frac{p_i^t}{p_i^0}$. The ratio of average prices is: $\frac{\sum p_i^t}{\sum p_i^0}$.

D.2 Sources for weights

4.32 The primary sources of weight information for the PPI are business- or establishment-based censuses, annual industrial surveys, and business registers.

D.2.1 Business or establishment censuses

4.33 The business census covers all establishments that have productive activity within the geographic borders of the country. These censuses may be conducted over several 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, and a census of services the year after that. In some instances, there may be a size cutoff to exclude small establishments. For example, some countries exclude establishments with fewer than five employees or with a low threshold of annual production. Alternatively, those countries might complete the census using a sample of small establishments only.

4.34 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 value 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.

D.2.2 Enterprise or industry surveys

4.35 These surveys differ from censuses primarily in three respects: (i) the coverage is limited to a sample of establishments rather than a full enumeration, (ii) the product detail is limited to higher aggregate levels such as groups, and (iii) the types of data requested are generally more limited. For example, product information in the census may be obtained at the eight-digit product code level with complete detail on product sales and inventories. In the industry, however, survey data are reported at the six-digit level and are requested only for sales. Also, data may be reported only at the enterprise level rather than broken down by establishment.

4.36 In these cases, the weights that are available will generally be for higher levels in the aggregation structure such as product group and industry, rather than detailed levels like product and establishment. The use of these weights for the PPI will depend on how the PPI aggregation structure has been established. If multitiered weights (for example, one set of weights for the industry level and above, 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.

D.2.3 National accounts

4.37 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 because of 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.

4.38 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 products by type of use is an excellent

source of net weight information for developing stage-of-processing indices.[9]

D.2.4 Business register

4.39 Most countries maintain a business register, which provides a list of firms that are involved in productive activities. Such registers usually contain information on location, economic activity, size (for example, employment, payroll, value of annual production, or turnover), contact persons, tax information, and so on. 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.

4.40 There are several shortcomings in 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 be outdated. The information on size of the firm also needs to be updated regularly. Much of the information may relate to the time when the firm was introduced into the register. Also, the business register may comprise a list of enterprises that is not completely suitable for the PPI, where the goal 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.

D.2.5 Other sources of weights

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

4.42 Another source for weight data is industry associations. Many associations conduct surveys of their membership that include detailed information on value of sales by product. Alternatively, in industries dominated by one or two large firms, the market shares for these firms can be a source of weight data.

4.43 In many countries, data on retail and wholesale turnover are produced regularly. Such data, if maintained at a detailed economic activity level, could serve as a source of weights for wholesale and retail economic activities. This 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.

4.44 Customs records are an alternative 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 frame for samples of export products.

4.45 *Statistical offices should make certain that data from any alternative sources conform to the definitions of the PPI.* Whichever weight concept the PPI uses (output, production, sales, or value of shipments), the data from these alternative sources should conform to that definition. For example, data on retail and wholesale turnover may be available at a detailed product level. One problem with these data is that they measure sales at purchasers' prices, which is inconsistent with other weights based on output at basic prices. The statistical office would need to adjust the sales information for taxes on products (for example, VAT) and separately invoiced transport charges. This adjustment derives turnover at basic prices; but, to derive an output measure, the statistical office would have to also make estimates of inventories for each product. In the example, if value of shipments were the weighting concept, only the adjustment used to derive basic prices is required.

4.46 Statistical offices also need to adjust primary source data for any known inconsistencies or errors. It often happens that reporting errors and in-

[9]Use of output data from the national accounts supply and use tables will provide weights that include nonmarket activities (see Chapter 14). Users must be aware of this fact if they intend to exclude nonmarket activity from the PPI.

consistencies are uncovered in censuses and surveys after final results are available. Statistical offices need to make sure appropriate adjustments are made to these source data when deriving PPI weights. For example, an establishment survey provides weights on total output by product, which includes the total value of inventories. The statistical office realizes that output values should only include the change inventories. It will be necessary to go back to the source data and adjust the final inventory figure to take out the value of inventories at the beginning of the period.[10]

E. Product and Transaction Weights

4.47 The selection of transactions to observe the price movements for each industry or product in the classification systems is a sampling issue discussed in detail in Chapter 5. The value weights at the industry or product level will generally be obtained from one of the sources discussed in the previous section. As soon as these results are available, one must determine what specific transactions in goods and services should constitute each elementary aggregate of the PPI. The data from industrial censuses are preferred because they provide a much larger coverage of goods and services than can possibly be observed in most surveys of enterprises. However, even the census will not contain details for each transaction that has transpired. For this reason, each elementary aggregate of the PPI must be represented by selected goods and services that are considered either important or representative of typical changes in relative prices for their class. The relative price changes of these particular goods and services are then monitored, and their average is subsequently used as a measure of relative price changes for that elementary aggregate.

E.1 Explicit and implicit weights

4.48 When the sample of representative transactions has been selected, a determination must be made about whether explicit weights can be derived. If probability sampling techniques are used, the inverse of the sampling fractions (or the sampling intervals)[11] are used as the weights.

4.49 In the case of judgmental samples, the weights for the selected industry and establishment should be adjusted to incorporate the weights of transactions not selected for the sample. Thus, the weight for small industries not selected should be allocated to those that were selected. For establishments, the same approach is used; the weight for the nonselected establishments within an industry must be allocated to those that were selected. Within the establishment, the total weight for the establishment can be distributed to the representative products in proportion to their share of sales. Finally, for each representative product, the weight for the product can also be distributed to each selected transaction in proportion to the selected transaction's sales. In this fashion, the weight for each establishment would be allocated to each price observation.

4.50 Alternatively, if certain products in an elementary aggregate are judged more important, higher weight may be assigned judgmentally or on the basis of secondary information from administrative or industry sources.

4.51 If no weights are available for the selected transactions, the formula used for averaging price observations will assign implicit weights to individual transactions. If the average of price ratios formula is used, as discussed in Section D.1, the implicit assumption is that relative price changes for each transaction within the elementary aggregate are equally important in terms of base-period quantities.[12] If the ratio of average prices is used, we assume that the importance of each observation is proportional to its base price.[13] The latter approach makes the strong assumption that production values are proportional to the base prices. In the ratio of average prices formula, transactions with

[10]This assumes that there has been no change in prices between the start and end period. If there has been such a change in prices, an inventory valuation adjustment must be made. See Bloem, Dippelsman, and Maehle (2001, pp. 60–63) or Shrestha and Fassler (2003) for techniques to make such an adjustment.

[11]For example, if total output for an industry is 10,000 and five establishments are to be selected, then the sampling fraction is 1 in 2,000, the sampling interval is 2,000, and the weight for each establishment selected is 2,000.

[12]This uses the first formula in footnote 8. For each transaction, the current price is divided by the base price, then the average of these price relatives is calculated.

[13]This uses the second formula in footnote 8. The current average price of the selected transactions divided by the base-period average price of the selected transactions yields the price relative.

higher prices receive more importance than those with lower prices. Often these differences in price levels occur because of the nature of the transaction specifications rather than real differences in the relative importance of transactions within the establishment.

4.52 Another alternative formula is the geometric average.[14] The geometric average of price relatives and the ratio of geometric average prices yield the same result. The use of this formula assumes that the weight of each observation is equal to its share of base-period production value (not its share of base-period quantities). Thus, as relative prices change, the assumption is made that there is an inverse relationship between the change in prices and the quantity produced consistent with a unitary elasticity of substitution so that a 1 percent rise in price results in a 1 percent decline in quantity produced. For the PPI, this inverse relationship between price and quantity may not be a valid assumption under some circumstances. See Chapter 20 for a detailed presentation of this issue.

E.2 Sources of product and transaction weights

E.2.1 Business censuses and surveys

4.53 As discussed previously, the censuses of business and the establishment census[15] are good sources for value of production or sales information to use as weights at the establishment and product level. Usually, such censuses would also contain information about products within establishment that is the most valuable source for obtaining weights by product classification within establishment. These censuses will not provide information by transaction because such information would place a heavy burden on reporting units.

4.54 Annual surveys by industry will often provide information at higher levels of aggregation, such as estimates of production by industry or key product lines within industrial activity. However, information at the establishment level is generally limited to the sampled establishments and will not contain full product detail within those establishments. Establishment and detailed product weights will be available from these surveys only to the extent that there is an overlap in the samples of establishments and products between the PPI survey and the industrial survey.

4.55 If a multitiered weighting system is used, such surveys would be a good source for updating weights at higher levels of the aggregation structure. They could also be used as a source for updating weights when producing annually chained indices (see Chapter 9).

E.2.2 Business registers

4.56 If the business register contains production or sales data, it forms a potential source of establishment weights. If the register is updated frequently, the weight information could be more current than census data. However, the business register is not likely to contain data on products produced within individual establishments. In addition, the weight information in the register may have differing reference periods for the establishments depending on procedures for updating information. If this is the case, the value weights will need to be adjusted for the differences in the weight reference period so that they are standardized across establishments.

E.2.3 Weights obtained from the probability sampling process

4.57 Sampling fractions or sampling intervals developed when the samples are drawn can be used as weights at the establishment and product level as appropriate. Individual weights at both the product level (if not available during the initial sampling phase) and the transaction level can be obtained through a sample disaggregation process using probability sampling techniques at the establishment level.

4.58 Disaggregation within establishment is accomplished by working with a knowledgeable respondent to determine probabilities of selection from production or sales data available at the estab-

[14] The geometric average formula is $\prod_{i=1}^{n}\left(\frac{p_i^t}{p_i^0}\right)^{1/n}$.

[15] The censuses of business usually are conducted by economic activities such as agriculture, mining and manufacturing, trade, services, and so on. These are usually collected in a cyclical fashion with one or two censuses per year over a five- or seven-year period. The establishment census covers all establishments at one time, regardless of their economic activity. Thus, the establishment census has broader economic coverage than the individual economic censuses.

lishment level as described in Chapter 5. By applying this technique at various levels, products and transactions are selected and the sampling factors ultimately determine the weights for the products and transactions.

E.2.4 Internal product and transaction weights obtained from establishments

4.59 When judgmental selection of products is used, the weights for each product are adjusted proportionally upward to represent all products within the establishment or product classification as discussed in Section E.1 of this chapter. Similarly, when judgmental selection of transactions is used, the weights for each transaction within the selected product can be adjusted proportionally upward to represent all transactions for the product.

E.2.5 Other sources

4.60 Data from administrative and regulatory sources can also serve as a source of weights if there is reporting of product and transaction information in sufficient detail. In the last few years, some countries have started to use electronic databases maintained by enterprises, marketing firms, and trade organizations to derive weights at the establishment, product, and transaction level. The databases consist of electronic data records that are maintained by or collected from producer enterprises. These data sets include information on the quantity sold, inventories, and the corresponding values for each. They also include the individual transactions, their prices, and the specifications for the transaction. This information can be used to derive PPI weights at the product and transaction level more frequently than otherwise would be possible. However, one should bear in mind the limitations of this source of information: the data usually represent only large producers. This may be adequate in highly concentrated industries, but it is less useful where small enterprises are prevalent.

4.61 Additional data may be available from tax revenue sources. Many countries have value-added or gross sales tax schemes that provide detailed information on sales revenue for a variety of enterprises and economic activities. Electronic scanner data on sales collected at the point of purchase are also available and used by a number of countries to derive weight information for detailed classifications.

E.2.6 New revolutionary products

4.62 As discussed in Chapter 8, new products should be introduced into the PPI as soon as possible to avoid potential bias in the index. The chapter discusses two types of new products: *evolutionary* products that represent continuous improvement over existing products and *revolutionary* products that represent a break from products previously available and are a new genre. Traditional surveys and sources of weight information typically will not provide the statistical office with any usable data. Some examples of revolutionary products include video recording devices and mobile phones.

4.63 If the new product falls within the existing classification structure, it can be introduced into the PPI calculation system by adding the product within an existing class. The weight for the product class remains the same, but the weights for the individual products will have to be recalculated. Since information on output of the revolutionary product will not be available from existing establishment surveys, the statistical office will have to seek production data from other sources. If only a few enterprises are involved in the production or distribution of the product, the statistical office can do a special survey to collect the value of output data directly from the enterprises. Other alternatives are to contact a trade association that represents the industry and product or, if it is subject to regulation, to contact the regulatory authority. The statistical office will also have to determine the specific sample transactions to price on an ongoing basis.

4.64 Once a production value is obtained, the temporal value of the new product weights is aligned with that of the other products in the class. For example, if the new weights refer to calendar 2002, but the weights for the other products in the class are for 2000, the new product weights should be deflated to a price level reflecting calendar year 2000 prices. The statistical office can use the product class PPI to deflate the 2002 production value. For example, a new mobile phone system is introduced in addition to the traditional land line system. The enterprise offering the new system can provide data on total revenue received during 2002, its first year of operations. The other weight information in the telephone class has a weight reference period of 2000. Mobile phone data for 2002 reflect average

prices in 2002 and can be adjusted to 2000 price levels by dividing the 2002 value by the price change in the telephone class between 2000 and 2002.[16] The new weights for land line and mobile are then used to calculate the aggregate index for telephone services using the old elementary index for land line phones and the new index for mobile phones.

E.2.7 Household enterprises

4.65 Household enterprises engaged in economic activity should be included in the PPI. Often statistical offices will exclude establishments below a certain size, for example, those that have fewer than 10 employees. Such size cutoffs are made because of the lack of good source data for weights and the relative unimportance of such establishments in most industries. This will exclude most unincorporated household enterprises; but, in many industries, this type of establishment dominates. For example, small establishments dominate many home craft industries and agriculture. It is important that they be included in the PPI for these industries.

4.66 In many countries, statistical offices can identify these establishments as part of their establishment censuses where data on value of production or turnover can be obtained. Such censuses can be used to develop sampling frames for industries with significant concentrations of small establishments. In other countries, government tax authorities maintain records on such establishments for administrative purposes. As mentioned earlier, the tax records may not have adequate information to derive PPI weights, but they can serve as a sampling frame for identifying units. The statistical office may have to derive the value of the weights for these establishments as part of a separate survey. For example, a random sample of small establishments can be drawn from the tax files to collect information on production, products, and prices. This information would then be used to estimate production weights for the establishments and products in the PPI price survey.

[16]The index compilers would have to make a similar adjustment to the 2002 prices for the selected mobile phone transactions to estimate base prices for calendar 2000. Current prices would then be compared with the estimated base-period prices to derive the elementary index for mobile phones on a 2000 reference period.

F. Practical Steps for Selecting and Determining Weights

4.67 The process for determining weights in the PPI structure can be viewed in a variety of ways. The following represents an overview of the steps required so that readers of this *Manual* have a sense of the milestones involved in developing a full set of PPI weights.

F.1 Determine sources for weights of economic activities and products that are in scope

4.68 The initial scope of the PPI is established in terms of the economic sectors (manufacturing, mining, construction, agriculture, transport, etc.) and the products that are produced in those sectors. The sources of the weights for each sector must be determined from those sources discussed in Section D. A review of these sources may indicate that additional surveys or censuses are needed.

F.2 Determine weights for sampled industries and products

4.69 Samples with a cutoff are used to determine the activities and products that will be included in the index for each sector. Such samples exclude activities and products that fall below a certain threshold. For the selected industries and products, weights must be established from those sources presented in Section D.

F.3 Determine weights for sampled establishments

4.70 The establishment weight could be taken directly from information provided in the sources discussed in Section E.2. However, if probability or judgmental sampling techniques must be used, the weight for each sampled establishment must be derived. If establishments are selected using probability techniques, the weights will be derived from the sampling probabilities. If judgmental techniques are used, the weights for selected establishments must be adjusted upward to include the weight for the remaining establishments.

F.4 Determine weights for sampled products

4.71 The product weight could be taken directly from information provided in the business census if that information is available as discussed in Section E.2. However, if probability or judgmental sampling techniques must be used, the weight for each sampled product must be derived. If products are selected using probability techniques, the weights will be derived from the sampling probabilities. If judgmental techniques are used, the initial weights for selected products must be adjusted upward to include the weight for the remaining products. When all products are selected within an establishment, the final product weight can be determined by distributing the establishment weight to each selected product using the product's relative importance among all selected products.

F.5 Determine weights for the sample of transactions

4.72 If no weights are used, the index formula for combining the transaction prices will determine the implicit weights as discussed in Section E.1. If transaction weights are used, they will be determined as discussed in Section E.2.

F.6 Adjust weights on the basis of sample yield

4.73 After all the establishments have been brought into the sample, the weights must be further adjusted for sample losses during the recruitment phase. Establishments that refused to participate must have their weight allocated to those that did respond, and other adjustments, as discussed in Chapter 5, may be necessary for establishments that were out of scope or out of business. When establishment weights are adjusted, the product and transaction weights may also have to be adjusted if they represent actual values rather than proportions.

F.7 Update weights for price changes due to differences between weight reference date and price base period

4.74 When the weights are introduced into the monthly or quarterly processing system, they must represent the same time period as the price base period used in the index calculation. If the weight reference period and the price base period differ, the weight should be adjusted for price changes (see Chapter 9).

F.8 Adjust weights as sample augmentation occurs

4.75 Sample attrition results in a continual decrease in the number of products and establishments in the sample. In addition, new products are produced by enterprises in response to customer demand. The PPI should have a cycle that augments the sample; doing so maintains the size and representation of the sample as discussed in Chapter 5. When new samples of establishments and products are introduced, the weights for establishments and products within establishments will need to be adjusted.

4.76 Portions of Chapter 10 deal in greater detail with determining the weights for some specific economic activities such as insurance, financial services, and retail trade.