6. Sampling Issues in Price Collection

A. Introduction

6.1 Price surveys normally collect information on products—their prices and price-determining characteristics—from establishments. For foreign trade in goods, however, administrative sources (customs records) usually are available from which compilers can calculate unit values. The central question of external trade price indices thus arises: is a unit value a price and a valid data element for constructing an export or import price index? Can we, therefore, use it rather than spending the effort and cost to collect prices from establishments? In Chapter 2 concerns were expressed about the widespread use of unit values as surrogated for price indices, even at the most detailed level of Harmonized System classes of goods, and even if subdivided by source (import) or destination (export) country. A strategy was outlined in Chapter 2 for countries whose trade price indices rely on unit values for moving from such a system to a hybrid system, that includes actual prices surveyed from establishments, and subsequently, resources permitting, to one in which most of the commodities included has prices based on survey information. A price must be associated with a given and complete description of the product encompassing the product and transaction characteristics that affect the exchange value or price. Unit values more often than not are averages across a variety of such descriptions, and thus, they are subject to composition effects. Unit values will change not only because of change in the price of any given product description within the HS class, but also because an HS class contains a different assortment of priced product descriptions from month to month, quarter to quarter, and year to year.

6.2 More often than not, as outlined in Chapter 2, a given HS class, even crossed with country of origin (import) or destination (export), does not define a homogeneous class of transactions. As a product description, it is insufficient. Unit value indices are used by many countries and a move to price indices has resource consequences. The preferable, though resource-intensive, approach is a one-off switch to an index based on establishment-based price surveys. This may be prompted by a country joining a customs/monetary union. While the main problem with simply introducing a new program is the resource cost, if a PPI program is already established, there will be natural synergies between the export and import price indices and the PPI. If resource constraints preclude this, one possibility is to identify whether there are particular products less prone to unit value bias and utilize unit value indices only for these sub aggregates in a hybrid overall index. The compilation techniques for such hybrid indices was outlined in Chapter 2, Sections E.1 and E.2 and an example is provided in table 6.2 of this Chapter. The use of hybrid indices is a strategic option outlined in Chapter 2 Section E. *It is stressed that this is a strategy for statistical offices with limited resources.* The intention of this gradualist approach is that it be a staged progression for an eventual move to a system in which the primary data source would be survey based prices of well specified representative items. A gradualist approach has major resource benefits. There will be some "low-hanging fruit," establishments responsible for relatively high proportions

of exports and imports, some of which may be owned by the state and may have some reporting obligation. Likely examples of such commodity groups include natural gas, petroleum, electricity, and airlines. There will also be industries in which unit values indices are *prima facie* inadequate measures of price changes, largely because of the churn in highly differentiated products, or the custom-made nature of the products, such as shipbuilding and oil platforms. Further, there may be industries which account for a substantial proportion of trade and the pay off of reliable data far outweighs the survey costs, for example, the use of surveys of fish-processing plants for major exporters of fish products and of agricultural marketing cooperatives for exports of primary products.

6.3 The gradualist approach requires as a first step a rigorous evaluation of each commodity group of the relative pay-off and cost of abandoning unit value indices. A good starting point would be a listing by commodity group that includes the weight, the perceived reliability of the unit value series, the likely source and reliability of alternative series, and a grade for the relative cost of obtaining such data. The initial aim would be to identify important commodity groups whose current series are deemed unreliable for which there are readily available alternative sources. The gradualist approach is outlined in Chapter 2, Section G. The evaluation implies that we also look to see if there commodity classes which are *prima facie* suitable for unit value indices, at least in the short run. Section C of this Chapter outlines some simple distributions and summary measures to help identify potentially homogeneous classes for which unit value indices may be used as a starting point.

6.4 Empirically, we say a product description is sufficient if the variability of the prices of transactions classified under that description is low at any given time. However, the empirical justification is not sufficient. The dispersion must be expected to continue to be low and this is likely to be the case if the customs class is a single product, not differentiated. XMPI compilers cannot be expected to be experts in the details of every commodity market. If the variance in prices is particularly low for a detailed HS class, good practice would be for the compiler to contact the importing/exporting establishment to establish whether commodities are available in differentiated forms not only with regard to their characteristics, but also the terms of sale.

6.5 When the detailed customs class is not a single product, it is necessary to identify the products within the class and track their prices. The standard approach in this case, which is no different than for other price indices such as the Consumer and Producer Price Indices, is to design a survey of establishments engaged in foreign trade, select a representative sample of fully described products from them, and follow the prices of this sample of products to construct export and import price indices. The focus of this chapter is with obtaining price information from surveys of importing and exporting establishments.

6.6 In an ideal world, it would always be possible to use statistically sound sampling techniques to produce price indices with a high degree of accuracy and within given resource constraints. There are, however, factors mitigating against achieving an ideal, statistically efficient sample within a given sampling budget: (i) accurate estimates of population variances, required for allocation of sample units to strata, are rarely available; (ii) sampling frames are always deficient to some extent, missing some key information, such as births of new establishments, or desired stratification variables; and (iii) response rates are

unpredictable and may prove to be deficient, which affects the accuracy of the price index levels and measured price changes.

6.7 The aim of the sampling statistician is, therefore, to make the best use of what is available and to apply the principles of sampling theory in a common sense and practical way. Arguably the most important steps in sampling are to establish and understand fully what the survey is trying to estimate, the limitations of the sampling frame, and the environment in which the survey will be conducted, that is, likely response rates, data quality, and levels of resources.

6.8 There is a direct relationship among the uses of the XPI and MPI, the scope of the external trade price survey coverage, and the requirements for sampling frames. Two of the major uses of the XPI and MPI are as general indicators for inflation in a country's international trade and deflators for goods and services trade in the national accounts. The broader the trade coverage of the indices in terms of economic activities, particularly services, the more useful they are in inflation analysis and compiling constant price GDP measures. But broad coverage requires the ability to develop sampling frames for the full range of economic activities engaged in international trade, including both goods-producing and service-producing activities. These sampling frames also must be kept up-to-date by recording both the births and deaths of enterprises in each sector.

6.9 Once coverage and uses have been established, a sample design can be drawn up, with decisions made about stratification, sample size, and allocation. Random sampling techniques may be employed in countries where large amounts of data are available and reasonable estimates of variance can be made. In many country situations, only limited details of sampling parameters are available, and the statistician may have to fall back on procedures that use expert knowledge at many stages in the selection process. To the extent possible, acceptable, practicable sampling procedures should be used. Judgmental approaches should be used only as a last resort.

6.10 As with most panel samples collected through time, price surveys suffer from problems associated with a changing population. Any sample of establishments and products will become increasingly unrepresentative over time, and is likely to be depleted as establishments cease the sale or the production of selected products or cease operations altogether. Some form of panel rotation or supplementation for the samples is advised to minimize any bias caused by sample attrition, noncoverage of new products, new establishments, and new production technologies.

B. Starting Position

6.11 The data to be collected must be identified and understood. Focusing on goods first, if there are customs records available on a timely basis, the problem is to determine which of the detailed customs strata or classes contain a single product. For these classes, use of already collected administrative information is more efficient and much less expensive than

developing establishment survey sources from which prices are directly collected.[1] However, such cases, as outlined in Chapter 2, are likely to be the exception rather than the rule, but may be a necessary undertaking because of resource requirments. Regardless of the source of information, whether administrative or survey, the source must be assessed for its compliance with the statistical concepts set out in Chapter 3.

B.1 Collecting transaction prices

6.12 It can be difficult to define and collect prices for many goods and services. Often the quoted list or book price does not represent the price received or paid by the establishment. Ideally, we want to collect actual prices received for a representative sample of establishment transactions. For goods, this can be achieved quite regularly. This is also the case with most services. However, for some services—for example, banking and insurance services—the volume and price of financial intermediation are not clear-cut and the price may have to be derived from transaction information. (Additional information on prices for these services is provided in Chapter 11.) In addition, the prices collected should be based on transactions rather than estimates, except as a last resort.

B.2 Basic price valuation

6.13 According to the *1993 SNA*, the supply of goods and services should be valued at basic prices, and, as a supply component, so should Import Price Indices. As we noted in Chapters 3 and 4, f.o.b. prices are effectively basic prices for imports, but are often difficult to obtain, so for practical reasons most countries use c.i.f. prices. When MPIs use purchasers' prices—c.i.f. prices plus tariffs on imports—rather than basic prices, we would need to use them to deflate a comparably valued import aggregate rather than imports supplied, f.o.b. In fact, the orientation of such an index would be different from the *1993 SNA* view of imports as a supply component. Instead, we would view them as an intermediate use of domestic production and as a final use in capital formation, and to a lesser extent, in their use as final consumption.

6.14 The *1993 SNA* recommends purchasers' prices for the uses of goods and services. As the *1993 SNA* views exports as a final use of domestically produced goods and services by nonresidents, exports should be valued at purchasers' prices. They thus include all relevant taxes on products levied by the government of the source country, and all transport, insurance, and distribution charges to the point of international shipment. This is the f.o.b. price.

B.3 Date of recording transactions

6.15 In line with the valuation of international trade in the *1993 SNA*, accrual accounting rules should be followed as far as possible, so that in the XPI and MPI, prices are recorded at the time of shipment (goods) or delivery (services). Although country practices often differ—

[1] An exception would be when establishment price data are already collected for a PPI and the output (input) is mainly for exports (imports), or if the establishment is willing to extend the survey to domestic and foreign markets.

for example, prices may be recorded at the time of purchase or order—the preferred timing generally is at the time of shipment for goods and of delivery for services. Prices could be an average of several observations during the month or the price on a particular day of the month; both approaches are used and are acceptable.

B.4 Structured product descriptions

6.16 The price-determining characteristics of each product or variety should be identified so that transaction specifications are complete; that is, that there is little variance in the price of a product with that description at any given point in time. For example, the price per liter of paint will depend on the number of cans to be shipped, type and quality of paint, terms of payment (net 30 days), type of customer, and any special discounts that may apply. The **product specification** or description in XPIs and MPIs, as for other price indices, should be systematic and permit statistical analysis of how product and transaction characteristics contribute to the transaction price at any given point in time. International good practice in product description is exemplified in the **Structured Product Descriptions** of the International Comparisons Program, which are similar to the "checklists" used by certain national statistical offices.

C. Goods: Testing customs elementary aggregates for multiple elementary items

6.17 The first phase in setting up a hybrid compilation system for export and import goods price indices is to identify the elementary items whose prices will be tracked by the index. Resource constraints may dictate that unit value indices have to play a major role in XMPI measurement and the logic of this process is to begin with evaluating the data already available from customs sources to form hybrid indices. The objective is to test whether each good's elementary aggregate defined by the detailed customs commodity code and destination or source country comprises a single elementary item. If so, because the unit value can be more used as a surrogate price, a price relative can be formed from them directly, and no further collections are required, assuming customs timing errors are not too severe relative to the change of ownership principle. The requirement that a single elementary item comprises a code, for example, *Harmonized Code 6402991815: tennis shoes, basketball shoes, gym shoes, training shoes and the like*, contains trade in an item of a single quality under the same terms of sale over time, and will continue to do so, is a restrictive one. If the good's elementary aggregate, defined by the detailed customs commodity code and destination or source country, cannot be deemed to comprise a single elementary item, additional surveys will be required to identify the underlying elementary items within those commodity codes. Yet there may be cases where unit values have to be used in hybrid indices that cover more than one item in a class. To evaluate the fitness of customs unit values as the basis for elementary aggregate price indices, we consider two suites of testing protocols below—though see also Chapter 10 Section D for an account of

outlier detection routines.[2] These are illustrative and there are many alternative approaches that may be more suitable for a particular country's needs. In adopting a set of tests a country should apply them to past data not only to decide on the appropriate methods, but also the appropriate (cut-off) parameters to adopt for a particular method. Contacts with importing and exporting establishments will be useful to decide on what constitutes outliers or alternative groups.

6.18 However, the use of unit values in such circumstances is born out of limited resources to do otherwise and should not be relied upon. Thus the rationale behind the strategy in Chapter 2, Section G of such hybrid indices being part of a staged progression to XMPIs, primarily based on survey data.

C.1 Price dispersion test

6.19 Our definition of an elementary item is based fundamentally on the price dispersion of all transactions falling within the group defined by the item. We can consider a given domain of export or import transactions defined by a particular set of commodity and transaction characteristics to be an elementary item if there is very little price dispersion within the domain at any given point in time. An elementary aggregate defined by a customs commodity class crossed with destination/source country may satisfy this condition. In this case, there is one elementary item in the elementary aggregate and the unit value that can be derived from customs information may be considered a reasonable estimator of the desired unit value estimate for the elementary item.

6.20 We would test this for a given month by constructing a unit value for every transaction (customs document) in the domain in that month and measuring the statistical variance of the resulting collection of unit values. A sufficiently low variance measurement would allow us to use the customs unit value of the elementary aggregate as the basis for an index. There are caveats to the accuracy of this empirical test, however.

- First, as noted in the previous section, the quantity measure available on customs documents is a shipping quantity rather than a transaction quantity. The shipping quantity and the transaction quantity must be in a fixed (for all time) proportion to one another across all transactions in the elementary aggregate for the customs (shipping) unit values to be accurate estimators for the desired transactions unit values. They can and often do differ because of recording errors in the shipping quantities on the customs form, as already noted.

- Second, the date of each customs document must be within the same month as the date of the change of ownership of the goods in question (i.e., the date the transaction accrues).

[2] Systems for outlier detection tailored to the needs of unit value indices are outlined in Technical Annex B, pages 190–198 of the World Trade Organization, International Trade Centre, and UNCTAD, 2007, *World Tariff Profiles 2006* (Geneva: WTO).

6.21 If the first condition is seriously violated, the shipping unit values are inaccurate estimators for the desired transaction unit values. If the second condition is violated, the unit values will be classified in the wrong month even if they are correctly calculated. Both of these types of error would tend to cause us to conclude that a customs elementary aggregate is not coincident with a single elementary item, even if it is true.

6.22 Consider the following price dispersion testing suite to ascertain whether a customs elementary aggregate contains one or multiple elementary items:

- Compute trade values, total shipping quantities, and unit values for each customs document classified into a set of **cells** of data defined by the

  Month or quarter of observation,

  Detailed (6- to10-digit HS) **commodity code**, and

  Destination (exports) or source (imports) **country**.

- Compute and examine the mean and standard deviation of the unit values in each cell.

- For those cells whose unit value coefficient of variation is **greater than 0.5**, **test for multiple elementary items** within the cell.

  Construct a histogram of the unit values within the cell.

  Examine the unit values in the cell for a few extreme upper or lower outlier unit values, for example three or more standard errors from the mean unit value, whose removal might significantly reduce the coefficient of variation. If there is evidence of outliers, examine the customs document of the outlier for errors (e.g., in order of magnitude) in the value or quantity recorded, or consult the text description for indications of different product or transaction characteristics explaining the unusual unit value from the document. Examine several quarters of data to verify the persistence of any observed patterns.

- Examine the histogram for evidence of separate clusters of unit values. If two or more clusters are visually identified, examine the customs forms in the identified clusters for comments indicating differences in product or transaction characteristics explaining the difference in the unit values represented between the clusters. Examine several quarters of data to verify the persistence of any observed patterns.

  If outliers or clustering identified by differences in product or transaction characteristics are not present or occur only once in several quarters, trim the outliers and smallest clusters from the cell and examine the time series of the resultant estimates of unit values for the cell for erratic temporal behavior inconsistent with any available anecdotal evidence about the prices in the product class and concurrent exchange rate movements.

If outliers or clustering identified by differences in product or transaction characteristics are persistent (e.g., more than one in five quarters),

Recommend adding a coded data item identifying the clusters of items on the customs forms for the product class to permit sub-classifications of customs forms within the existing cell; and/or

Consider collecting a survey of exporters/importers (discussed below) to measure and test for statistically significant price determining characteristics within the product cell.

- If there is no conclusive evidence of outliers and clustering but there is an erratic time series of cell unit values at the desired frequency of the export or import price index (e.g., quarterly or monthly),

  Consider collecting a survey of exporter/importers (discussed below) to measure and test for statistically significant price determining characteristics within the product cell. (Some commodities do have erratic price variations, which the UV can truly reflect. This should also be taken into account.)

- If the cell coefficient of variation in customs form unit values is consistently less than or equal to 0.5; or there is no conclusive evidence of outliers and clustering and a reasonable time series of unit values for outlier-adjusted cells at the desired frequency of the export or import price index,

  Consider the commodity/country cell as an elementary item and the cell unit values as the average prices of elementary items.

6.23 In the presence of some price dispersion in the elementary aggregate, a further test would require additional information on the price-determining characteristics of the transactions in the customs elementary aggregate, if available. If there is no variation in any of these characteristics, we would tend to accept that there is a single elementary item in the aggregate. If, however, there is significant variation in the price determining characteristics, we would conclude that the customs aggregate contains more than one elementary aggregate.

6.24 Detection of variation in price determining characteristics usually would proceed judgmentally on the basis of any text notes included on the customs documents in the domain of the elementary aggregate. Multiple elementary items also might be suspected if there are dissimilar clusters of shipment unit values within the domain. Credible evidence of multiple elementary items within a customs elementary aggregate would be the basis for augmenting customs data with survey information in order to identify those elementary items by measuring the associated product and transaction characteristics common to the elementary aggregate.

C.2 Quantity proportionality test

6.25 We note for completeness that if the quantities transacted of elementary items within an elementary aggregate are highly correlated from period to period—that is, product

quantities remain in roughly fixed proportions—then unit values across elementary items can be used to track price change in the elementary aggregate *even if it fails the price dispersion suite*. Fixed quantity proportions over time at high levels of product detail is, however, an untenable assumption for most price indices, and this is certainly true of export and import price indices. Table 6.1 illustrates the pitfalls of using unit values to produce a price index for a metals elementary aggregate in a case where there has been no change in prices, but there has been a shift in relative quantities. The challenge to constructing a test of proportional quantities within a customs product-destination/source cell is that it often is not possible to obtain repeated observations of the value and quantity shipped by a given shipper of a given specific product, as this would require a laborious process of going through customs documents from month to month looking for possibilities of matches. In some months for some types of products, there will be few if any matches to be found even if the effort is spent to find them. This tends to bolster the view that if the price dispersion suite fails, the default methodology is to survey importers/exporters in the cell.

6.26 Nevertheless, an approach to testing for proportional shipping quantities would be to examine the average shipping quantity across quantiles (say, quintiles or deciles) of the values of shipments, and compare relatives of the average quantity shipped between months for each quantile. If the quantity relatives are all the same across quantiles or tightly clustered, some support would be lent to the proposition that relative quantities have not changed.

D. Goods and services: Surveying enterprises to identify elementary items

6.27 When a customs elementary aggregate is deemed to contain multiple differentiated elementary items, or in order to survey goods and services transactions beyond the scope of customs sources, it is necessary for good measurement to design surveys of enterprises to obtain elementary item prices. In part because statistical surveys can be designed for export and import price indices whereas customs administrative files are designed principally for tax collection, surveys can capture information on the characteristics of goods and services to the level of specificity needed for a complete product description. The kinds of characteristics of products and transactions on which survey information is to be recorded for each product type are determined as a result of review of trade association literature, comment fields on customs forms, various press sources, and previous survey experience with the elementary aggregate, if available.

6.28 These aggregates generally will be defined for goods by the same detailed HS codes used by the customs administration, crossed with destination (exports) or source (imports) country. The particular goods aggregates subject to survey will be those that testing has indicated contain multiple elementary items, or for which no customs data are available.

Table 6.1. Unit Values and Product Mix

This example uses import data for three commodities considered to be three specifications of the same item where the nature of the sales contract (spot versus long term) affects the price.

Commodity		Period *t*-1			Period *t*		
2616.10.00.10	SILVER ORES AND CONCENTRATES: COPPER CONTENT						
Source:	Orlandia	Value (V)	Quantity (q)	Unit value (p)	Value (V)	Quantity (q)	Unit value (p)
Shipping volume: metric ton		(1)	(2)	(3) = (1)/(2)	(1)	(2)	(3) = (1)/(2)
		Observed consignment customs value and shipping volume					
A –	Acme Ores, spot market transaction, Orlandia metals exchange	300	6	50	600	12	50
B –	Acme Ores, contract with Orlandia Metals Corporation	450	10	45	450	10	45
C –	Metals, Inc., contract with Orlandia Mining Company	1200	30	40	1200	30	40
	Total	1950	46	135	2250	49	135
	Weighted average unit value (A, B, C)			42.39			45.92

Illustrating the discussion in the text, we have been able to match only three unit value calculations in this customs aggregate between periods *t*-1 and *t*. The price dispersion test for these three would be inconclusive, though possible elementary items would be [A B], which combines on the basis of the shipper (here, Acme Ores) or [B C], which combines on the basis of the type of transaction (here, long-term contract). A quantity proportionality test applied to this example suggests that quantities are not proportional for any combination except [B C]. Hence there are at least two elementary items in this matched subset of customs elementary aggregate 2616100010, namely A and [B C]. Accordingly, it will be necessary to survey transactors in silver ores with copper content in order to accurately measure price change.

To demonstrate the importance of identifying and tracking elementary items in customs aggregates observe that in this example prices remain unchanged, but quantities (product mix) change. This leads to change in the unit value of the customs subset from

$$\mathbf{42.39} = ([\sum_{j=A}^{C} p_j^{t-1} q_j^{t-1}] / 3) \text{ to } \mathbf{45.92} = ([\sum_{j=A}^{C} p_j^{t} q_j^{t}] / 3).$$

In other words, the unit-value index based on data for specific transactions A, B, and C equals **108.33** (or 45.92/42.39). On the other hand, because the price relative of each of the individual goods is equal to **one** between *t*-1 and *t* in our example, any price index formula applied to the price, quantity, and value data for varieties A, B, and C would yield the value **100.00** for the commodity group comprising them. The unit value thus has a **composition error** of 108.33 – 100.00 = **8.3 percent**.

6.29 For services, a starting set of definitions for elementary aggregates would be the five digit codes under the United Nations Central Product Classification, version 1.1, divisions 59.

6.30 In designing a sample of exporters or importers, the first step is to assemble a comprehensive list or frame of resident trader establishments in the various goods and services to be surveyed. For goods, establishment sample frames normally are set up using the customs source, as the names of the seller and buyer should be on each customs form. The seller's name would be captured from export declarations to form an establishment sample frame for exports, and the buyer's name would be captured from tariff filings to form the frame for imports.

6.31 Lists of establishments engaged in international trade in services also may be set up from administrative sources, for example exporters of services might be assembled from individual and business income tax filings reporting a foreign source of, respectively, earned income and sales, after making a comparison with foreign sales of goods from examination of shippers export declarations collected by customs. The most efficient source would most probably be available to balance of payments compilers which in many cases come from commercial banks reports of international payments. XMPI compilers should examine these records, and ensure good co-operation with BOP compilers. A good if less focused alternative would be implementing and/or expanding general establishment survey collections for service activities, such as for the PPI. These collections are based on survey frames compiled from lists of resident establishments and enterprises extracted from business registers, or from frames assembled from tax filings. Elementary items comprising export transactions can be determined and collected from the same sample of establishments used for these general collections, but large increases in sample size may be required to capture an adequate sample of products and transactions in external trade.

6.32 However, for transportation and transportation insurance activities related to imports, if the establishment's information is recorded on properly completed customs documents then customs sources can be used to assemble the survey frame of establishments.

6.33 Other factors to consider in sourcing data for trade price indices from direct price surveys are:

- The elementary items should be periodically reselected within the elementary aggregates, at least every five years, to keep the sample representative of current trade flows—usually a costly process for both the statistical agencies and the respondents;

- For customs elementary aggregates containing multiple elementary items, the coverage of commodities and traders from samples may be volatile from month to month or quarter to quarter if infrequent (or casual) traders account for an important share of exports or imports; and

- Through a properly designed questionnaire and good response from sampled establishments, the timing of prices collected through direct surveys can be made to closely approximate the change of ownership accrual principle required by the *BPM5* and *1993 SNA*.

6.34 Table 6.2 provides a stylized example of how both customs and survey sources of price data could be incorporated into an import price index. In this example customs classes

- 2616100010 SILVER ORES AND CONCENTRATES: COPPER CONTENT;
- 2701200000 BRIQUETTES, OVOIDS AND SIMILAR SOLID FUELS FROM COAL, and
- 2707100000 BENZENE, WEIGHT OF AROMATIC CONSTITUENTS GREATER THAN NONAROMATIC

are considered to comprise only one elementary item, while

- 8419899085 INDUSTRIAL MACHINERY, PLANT OR EQUIPMENT FOR THE TREATMENT OF MATERIALS, INVOLVING A CHANGE IN TEMPERATURE

comprises multiple elementary items. Hence, a sample of elementary items is developed for industrial machinery, comprising

- Cracking tower, heavy crude oil feedstock; mfr: Petro-Equipment Corporation
- Paint drying booth; 4m ×10m; mfr: Radiant Products, Pte.

E. Common Problems in Price Survey Sampling

6.35 There may be many reasons why price surveys are thought to be unrepresentative and thus liable to lead to inaccurate results. All national price surveys suffer from problems to some extent. The following are some examples:

- Samples are selected purposively rather than using probability sampling methods, increasing the chances of bias. For example, establishments may be selected for their convenient geographical location or because they are known to be good respondents;
- Without probability selection methods, estimates of statistical accuracy cannot be made (but without some initial estimate of variance, a randomly selected sample cannot be optimized—that is, lowest variance given cost constraints—either. This is a difficult problem that is dealt with later);
- The sample size for an industry or commodity may have become outdated if the industry or commodity has grown or contracted since the base period (period when sample was selected);

- New products may not be identified or included in the survey. This problem may be relieved to some extent by rotating the sample of establishments;

Table 6.2. Using Price Surveys and Customs Unit Values in the Same "Hybrid" Index

Commodity/source (destination)	Period *t*-1			Period *t*		
Trade index item group – import example	Value (V)	Quantity (Q)	Unit value or price (P)	Value (V)	Quantity (Q)	Unit value or price (P)
	(1)	(2)	(3) = (1)/(2)	(4)	(5)	(6) = (4)/(5)
1 – 2616.10.00.10 SILVER ORES AND CONCENTRATES: COPPER CONTENT; Source: Orlandia; Shipping volume: metric tons	300	6	50	1200	20	60
2 – 2701.20.00.00 BRIQUETTES, OVOIDS AND SIMILAR SOLID FUELS FROM COAL; Source: Mineland; Shipping volume: metric tons	450	10	45	450	10	45
3 – 2707.10.00.00 BENZENE, WEIGHT OF AROMATIC CONSTITUENTS GREATER THAN NONAROMATIC; Source: Lubovia; Shipping volume: metric tons	1200	30	40	600	22	27
Total value of unit value items	1950			2250		
4 – 8419.89.90.85 INDUSTRIAL MACHINERY, PLANT OR EQUIPMENT FOR THE TREATMENT OF MATERIALS, INVOLVING A CHANGE IN TEMPERATURE; Source: North Machindia.	***600***			***700***		
The following specific transactions collected from a survey sample:						
A – Cracking tower, heavy crude oil feedstock; mfr: Petro-Equipment Corporation	***500***		***500***	***500***		***500***
B – Paint drying booth; 4m×10m; mfr: Radiant Products, Pte.	***100***		***10***	***200***		***30***
Total value of hybrid items (unit value and directly priced)	**2550**			**2950**		

Unit value price index for items 1-3, and also aggregate unit value price index if index for 1-3 is imputed to item 4

Laspeyres: $[300/1950 \times 60/50 + 450/1950 \times 45/45 + 1200/1950 \times 27/40] \times 100 =$ 83.08

Paasche: $([1200/2250 \times (60/50)^{-1} + 450/2250 \times (45/45)^{-1} + 600/2250 \times (27/40)^{-1}])^{-1} \times 100 =$ 96.20

Fisher Ideal: $[(83.08 \times 96.20)^{1/2}] =$ 89.40

Elementary aggregate price index for item 4

Laspeyres: $[500/600 \times 500/500 + 100/600 \times 30/10] \times 100 =$ ***133.33***

Paasche: $((500/700) \times (500/500)^{-1} + (200/700) \times (30/10)^{-1})^{-1} \times 100 =$ ***123.53***

Aggregate hybrid external trade price index

Laspeyres: $1950/2550 \times 83.08 + 600/2550 \times \mathbf{133.33} =$ **94.90**

Paasche: $(2250/2950 \times (96.20)^{-1} + 700/2950 \times (\mathbf{123.53})^{-1})^{-1} =$ **101.53**

Fisher Ideal: $[(94.90 \times 101.53)^{1/2}] =$ **98.16**

This example presumes that the elementary aggregates silver ores with copper content, coal briquettes, and aromatic benzene each comprise a single elementary item and consequently that period to period relatives of their unit values are proper estimators for their price indices.

- The sampling frame may be out of date or may not include certain groups of the target population. For example, a common problem in the PPI is that information on small producers is unreliable because this group often is volatile, and difficult for administrative authorities to track, resulting in the weight for small producers may being wrong (typically they are underrepresented; and

- Surveys may be voluntary, increasing the chance of non-response bias that results when those who do not respond have different price experiences than those who do respond.

6.36 A decision should be made about the **level of accuracy** required:

- *Ideally, a maximum acceptable sampling error should be identified for each published index.*

Sampling error can be assessed only, however, if probability sampling techniques have been used. This often means starting with some estimates of variance for the component index to determine initial sample sizes. Then, once samples have been collected and variances calculated, the sample can be optimized based on the new variance information. However, the calculation of variances and sampling errors is very difficult to accomplish (Leaver, Johnstone, and Archer, 1991; Leaver and Swanson, 1992; Cope and Freeman. 1998; and Morris and Birch, 2001).[3]

- *In practice, there is a trade-off between cost and accuracy.*

A high level of accuracy that would be desirable requires larger sample sizes that may not be affordable. In such cases, costs often determine the sample sizes, and the level of accuracy may suffer somewhat. Many countries find that they can contain costs and maintain acceptable levels of accuracy by using cutoff sampling (discussed later).

6.37 Once the coverage is decided, the **population** to be sampled should be identified and the sampling frame reviewed to determine whether the existing frame needs to be supplemented.

- *Does the frame contain all of the units in the target population? Does it cover all of the industries that are in the scope and all of the establishments in the targeted industries? Will separate frames have to be developed for each industry, group or division?*

Customs documents contain useful information on the enterprises trading. Most business registers have a cutoff (threshold) below a certain size (number of employees or value of sales), and probably some industries that are less well covered, for example, construction and

[3]The United States has estimates of variance for its CPI and the United Kingdom has estimates of variance for its PPI. In both cases, the sample design was set up first without information on variances. The resulting variances are greater than if they had been known in advance. Once these first variances have been calculated, they can then be used to improve the efficiency of the sample design by reallocation of sample strata and the number of price observations in each.

retail trade. Also, there is a need to identify establishments separately from parent enterprises.

- *How are units defined in the frame? There are probably borderline units where it is uncertain if they belong in the population.*

Although the PPI survey could in principle serve as the "spine" for a supplemental survey of the prices for internationally traded goods and services, it may not have a sharp enough focus on international trade, and will typically cover only exports. For example, ancillary or auxiliary units of an enterprise may be out of scope, and the PPI sample might have to be uneconomically large to detect foreign trade when this trade covers a rather small part of total output. Thus, compilers need to develop a separate establishment sample frame for establishments engaged in international trade to select the sample of establishments that will report the prices of internationally traded products. Customs documents have the required information for this purpose.

- *Are units mutually exclusive?*

There could be double counting, which occurs when an establishment could be included both in its own right and as part of its parent enterprise.

- *Is there information available to allow stratification?*

We need certain data elements that will serve as stratification variables–for example, sales or purchases from abroad by type of good or service, industrial classification, production or sales, number of employees, and location of establishment, in order to select the sample. Again customs documents have information on the value of trade, HS classification, and location.

- *Is there information available to allow weighting for probability proportionate to size (PPS) selection?*

We will need measures of size, such as output, total sales, and value of shipments. Customs documents should be helpful in this respect. If such measures of value are not available, employment may have to be used as a proxy.

6.38 The level of available **resources** should be decided:

- *This will be a constraint on sample sizes.*

It is generally more expensive to increase the number of establishments sampled, as opposed to increasing the number of prices collected from each establishment. On the other hand, simply increasing the second may add little to accuracy, when intra-establishment (within an establishment) variance is low compared to inter-establishment (between establishments) variance.

- *And this may dictate the methods of measurement.*

The methods of measurement would include, for example, personal visits, telephone collection, or postal or electronic questionnaires.

6.39 **Legislative issues** may affect the sample design.

- *Will the survey be voluntary or statutory?*

This will affect response rates, which, in turn, have implications for accuracy and sample sizes. Statutory surveys will have higher response rates, although they may result in lower data quality.

- *Are there rules concerning confidentiality?*

This may impose a lower limit on sample sizes–for example, a minimum of four units per stratum may be required. In many small countries, this could be problematic when there are only one or two major exporters for certain goods and services. In such instances, the national statistical office may have to request approval from the companies to publish detailed product-level price indices.

F. Sample Design

6.40 Given information about what the XMPI survey is intended to achieve, the format of the inputs and outputs, desired level of accuracy, and available resources, the process of designing the sample can begin.[4] Again, decisions need to be made, but the main objective of the design process is clear—to maximize efficiency—that is to minimize sampling and nonsampling errors, and to minimize costs.

6.41 Decisions will need to be made about:

- Sampling techniques (probability vs. nonprobability),
- Sampling frames,
- Sample structures and stratification,
- Sample allocation between strata, and
- Methods for reducing non-sampling errors.

[4]There are many textbooks that can be consulted on the theory and application of sampling. One text used quite often is Cochran (1977), available worldwide.

F.1 Sampling techniques

F.1.1 Probability vs. nonprobability sampling

6.42 The statistician, confronted with any measurement problem, must initially consider the possibility of installing a rigorous probability sample. In the context of XMPIs, probability sampling means the selection of a sample panel of exporters, importers and products (transactions) from a universe of foreign trade activity in which each producer and product has a known chance of selection.

6.43 Nonprobability sampling is known as judgmental or purposive sampling, or expert choice, and samples are chosen by experts to be representative. In practice, however, different experts would rarely agree on what is representative, and the samples are subject to biases of unknown size. Judgmental sampling may be justified when sample sizes are small, but concern about their biases increases with sample size.

6.44 Alternatively, in cutoff sampling, all establishments and/or products over a certain threshold of size are selected for inclusion in the sample. Many countries have found cutoff sampling to be an efficient way to identify the import units for their establishment surveys. Not only does cutoff sampling save resources, it has been shown to approximate probability samples. Cutoff sampling is discussed in more detail below.

6.45 Using a probability sample comes with two well-known advantages. First, it ensures that the items to be priced are selected in an impartial and objective fashion. In the absence of probability sampling, a danger exists that only items that are easy to price will be selected, resulting in biased estimates (indices). In particular, there is likely to be poor coverage of technologically advanced items, like machine tools, electronic equipment, aircraft, or home electronics. These are difficult to price because of rapid changes in specifications. There is also a tendency to place too much emphasis on simpler products, like food items, cement, textiles, or steel bars, for which a comparable series of price quotations can easily be provided.

6.46 The second advantage is that a probability sample permits the measurement of the quality of the survey results through estimates of the variance or sampling error. The quality of results in this context relates to the chance of a difference between the results obtained from the sampled observations and the result that would have been obtained in a complete enumeration of all reporting units in the universe. The use of a probability sample, of course, does not permit the measurement of errors arising from nonresponse, inaccurate reports, obsolete weights, unrepresentativeness of the commodities priced, or any other non-sampling source.

6.47 Probability sampling conceivably could be used at all stages of the selection process. For example, a random sample of products could be selected from a comprehensive list of all goods produced by all mining and manufacturing firms. For each selected commodity, a random sample of producers could be picked using a comprehensive list of producers; for each selected producer, a random sample of specific brands could then be chosen for regular price reporting from a complete list of each producer's output. A less rigorous approach

might involve random choice of producers or retailers, followed by a purposive selection of individual products or items; alternatively, the producers or retailers might be selected on a nonprobability basis using cutoff sampling (described next), while a random sample is picked from all items made by the selected producers. This mixture of nonrandom with random selection procedures and cutoff sampling procedures narrows the interpretation that may be placed on estimated sampling errors but still will retain the advantage that a certain amount of objectivity is imparted to the selection process.

6.48 Optimal sample design requires, for all units in the population, information that will allow effective stratification and increased efficiency due to selection by PPS. Different variants of probability sampling can be used by statistical agencies:

Simple random sampling– every possible unit has an equal chance of being drawn;

Systematic sampling–every k^{th} unit is selected, after a random start. This sampling is affected by any ordering or pattern in the sampling frame. Ordering leads to a form of implicit stratification, and a pattern in the frame can lead to biased samples;

PPS–each unit has a probability of selection in proportion to it size (or some other indicator of importance, but size is commonly used). Once these probabilities of selection are assigned, either simple random or systematic sampling techniques can be used.

6.49 Despite the attractions of probability sampling methods, there will be situations where it is neither necessary nor desirable. Price indices are an area of statistics where the risks in not having a probability sample are relatively low. The potential diversity of the change in prices charged by various producers of a given commodity over many time periods is relatively low. Compare this to the potential diversity for sales or capital expenditures of firms making the same product over the same period of time. The largest firm may become the smallest, and vice versa. Some may even abandon production of the commodity, and new firms may enter. In summary, the measurement of price changes appears to require less rigor with respect to probability sampling than do other areas of statistical measurement. The additional costs that may be involving probability sampling can be allocated to other areas in the survey, such as price data collection or improvements to source data on weights.

6.50 That said, without probability sampling, statistical agencies will not be able to produce meaningful measures of sampling error to guide users in distinguishing between real changes in prices and those due to statistical noise. They also will experience difficulty in statistical decision making to improve the sample design and allocate resources more efficiently. Good measures of sampling error provide statistical offices with data for reallocating the sample to areas with high variance to reduce statistical error.

6.51 In several countries, the range of domestically produced mining and manufacturing goods is so limited and the number of firms producing them so small that there is no point in making a selection; the survey should try to cover all products and all producers.

6.52 In other cases, there may be no practical way of determining the universe in advance. A basic requirement for probability sampling is to define the universe (or population) and to

identify all units in the universe. For XMPIs the univers of goods is usually defined by the customs data for goods and the BOP information for services.

6.53 The cost of installing and administering a probability sample may be judged too high. There clearly are high costs involved in the design, selection process, control, and administration of a probability sample for collecting price observations.

6.54 Estimates of variability in price movements also are needed. This information is rarely available for all units in the population, certainly not at a detailed product or item level. One way of dealing with this is to use a two-phase sample, where certain information is collected from a sample of units, and then these units are resampled using this information.

6.55 Probability selection often will be inappropriate because the survey of producers' prices ideally should form part of an integrated program of price statistics. This means that the choice of items to be priced at the intermediate (that is, producers' prices) stage may depend on the items selected for pricing at an earlier (for example, imports) or at a later (for example, exports, consumption) stage.

6.56 Thus, for most countries a strict probability approach will not be possible or the costs will greatly outweigh the advantages, so a combination of probability and purposive sampling techniques is employed.

F.1.2 Cutoff sampling

6.57 Cutoff sampling is a strategy frequently used by countries to select samples. In this approach, a predetermined threshold is established with all units at or above the threshold included in the sample (selected with certainty) and units below the threshold level not included (zero probability of selections). Cutoff sampling generally results in a high degree of coverage among a small number of prospective units. This occurs because the distribution of the selection variable (for example, value of imports or exports) is concentrated in a small number of large establishments.[5]

6.58 The problem with such an approach is that the smaller establishments may have different price movements from the larger units and, thus, introduce an element of bias into the price index. The bias would be the difference between the average price change for the noncovered units and the price change for the overall population. If the importance of units excluded is very small or the bias is very small, the effect on the overall error may be very small. Usually the total error is measured by the root mean square error, RMSE, ($\sqrt{\boldsymbol{Variance} + \boldsymbol{Bias}^2}$), and the sample with the lower total error is deemed more efficient. Thus, the approach that produces the lowest total error or RMSE will be preferred. It is possible that a cutoff sample could be more efficient if the bias component of the excluded units is small. For example, if the noncovered units have substantial variation with regard to price change but small bias (that is, the average price change is not much different), the RMSE could be smaller using the cutoff sample, and the survey costs could be much lower.

[5]See Haan, Oppredoes, and Schut (1999) for an analysis of cutoff sampling in the CPI.

6.59 Cutoff sampling has a great deal of practicality for selecting the establishments and products in a multistage sampling scheme. For example, in selecting the importers of goods that will be included in the sample, a threshold can be established that all establishments covering up to, say, 70 percent of goods imports or exports will be chosen. Another aspect of sampling where the cutoff approach can be used is in the selection of the representative products in the sample where all products that represent up to, say, 70 percent of import or exports (at a say eight-digit HS classification) are included in the sample.

6.60 Cutoff sampling is not the same as probability sampling. Sampling errors for cutoff samples will not be as accurate because the sample is not necessarily representative of the index population. Statistical offices will need to make special efforts to measure bias among smaller firms in order to calculate the RMSE to get a meaningful measure of error.

F.1.3 Multitiered stratification

6.61 Alternatively, it may be useful to use stratified samples in which various classes of establishments are sampled separately. Often it is helpful to identify three or four strata based on their size, such as large, medium sized and small establishments, with each stratum having a different sampling rate. For example, large establishments (based on turnover or employment) may be sampled with certainty (that is, all selected in the sample), medium-sized establishments may be sampled at a rate of 25 percent (one out of every four), and small establishments may be sampled at a rate of 2 percent (one out of every 50).

F.2 Sampling frames

6.62 Whether selecting a sample using probability or nonprobability techniques, we need to define the universe (population) from which we wish to sample, that is, construct a sampling frame. In most countries it is possible to define the population using various lists of enterprises (business registers), compiled for administrative purposes. For the XMPI, these business registers can be constructed from customs' documents such as import and export declaration forms. An establishment will be included in the frame if it is named on a customs form as the recipient of an import or the declarer of an export.

6.63 The ideal **sampling frame** would

- *Be a complete list of all eligible units (producing and exporting) within the geographic and industry or product coverage required.*

6.64 Registers typically are compiled as the by-product of an administrative system such as tax collection, social security schemes, and customs records. Alternatively, lists can be compiled using records such as bank accounts. Tax and social insurance lists generally contain, as a minimum, information about geographical location and size (turnover or number of employees), but may not indicate the principal activity of an enterprise or identify it as an exporter. In trade statistics the main administrative source for assembling a list of establishments engaged in international trade is customs records, which only cover goods trade and the associated transport and insurance on imports. Supplementary lists from BOP sources will be needed for enumerating resident establishments purchasing services from and

providing services to nonresidents in transport and insurance services on exports, as well as passenger transport, life and property insurance services, business services, and others.

- *Be updated instantly with all births and deaths of units and changes in addresses, fax numbers, etc.*

6.65 Maintaining an up to date register is resource intensive. It generally is the case that information about the bigger units is more up-to-date than data on smaller units. This is a particular problem during periods of changing economic structure when some industries or residential areas are expanding, and new units may be starting up in large numbers. If units are not removed from the sampling frame when they no longer exist, they may be selected as part of the sample. This needs to be borne in mind when determining sample sizes. Also, a common error with systematic sampling is to substitute the next unit in the list when a dead unit is sampled, but this should be avoided since the probability of selection of that next unit is enhanced. The sampling interval should be repeated as usual, and dead units simply dropped.

- *Hold certain fields for each unit, allowing sorting of the list and stratification as required.*

6.66 For example, industry classification at the ISIC four-digit level and information about value of output would be maintained for XMPI purposes (ideally of each product, at the six-digit CPA level, produced by each unit, and for each transaction, the destination/source country). This information would be updated annually. Export and import values maintained using both customs and BOP records as part of the unique business ID maintained for each establishment.

6.67 Lists maintained primarily for tax collection purposes are likely to hold information on the values on which taxes are levied, for example, imports at cost-insurance-freight.

- *Identify each unit uniquely at the correct institutional level.*

6.68 In practice some units may be listed more than once, and others may be grouped under one listing. Ideally, a structure would identify enterprises and their corresponding establishment structure with separate classification and other stratification information for each establishment. If such information is not immediately available from the business register, additional steps or surveys may be needed to collect this information as part of the process of sample frame refinement.

F.3 Sample structure

6.69 The sample structure for XMPIs is straightforward because priority proceeds from classification by product at the highest level, to destination/source country, to industry supplying exports or purchasing imports.

6.70 Consider the XPI or MPI structure using the following example:

(i) We require XPIs (MPIs) for products (five-digit CPC), but also want to publish

time series by industries (four-digit ISIC) and destination (source) country;

(ii) There are establishments trading in a range of products characteristic of more than one industry and with more than one destination (source).

(iii) The shares of exports (imports) by product are the most stable over time, followed by the shares by industry, while the shares by destination (source) may be stable in some instances and volatile in others, particularly within a product class.

(iv) Seasonal availability affects the sample size as it will be predictably lower during off season periods; this should be taken into consideration in the design of minimum sample strata, so that the several similar products included within the stratum have overlapping and, collectively, year-round availability. Also, sample sizes for these strata should be increased because of the higher variability in price movements among seasonal products.

6.71 Given the uses of XMPIs and the relative stability of trade by product, the first step in this process involves selecting the products across industries and destination/source countries that will be represented in the XMPIs. In most countries some products account for an extremely small part of trade–for example, products that comprise less than 0.02 percent of total trade in a major product group for either exports or imports. (If this is not the case, then all products could be included for estimation.) It would be possible to use a cutoff approach where those products below the threshold level (in this example, 0.02 percent of sales) are excluded from the sample of products, but their weight is allocated to another closely related stratum or distributed across a number of other strata. Sampling frames for exports or imports then are built for cells cross classified by product.

6.72 The statistical office should review the products that fall below the cutoff point and determine if any traditionally important products should be included anyway. Also, newly emerging products that are expected to grow in importance might be included because they will eventually exceed the threshold. Finally, for the products not selected, the statistical office should determine if groupings of these products can be made that reach the threshold level.

6.73 Because most uses of XMPIs focus on the product breakdowns there should be a product orientation to the sample design. To construct *product XMPIs,* for exports we would need export or sales information and import or purchase information for each establishment for each six-digit HS product that it produces, enabling us to form a list of all producers for each six-digit product. From each list we would sample transactions and weight them accordingly to give product XMPIs.

6.74 To bring in industry detail, a compromise is to employ a two-stage[6] sampling scheme–that is, the frame is stratified first by product, then stratified by amount of trade

[6]A distinction is made between *two-stage* sampling, where a sample of establishments is selected and then a sample of transactions is selected from each, and *two-phase* sampling, where a sample of establishments is selected to provide detailed output data, and this sample then is used as a new sampling frame. This new frame

(continued)

within each product cell. Next, establishment samples are selected to represent each of these cells or strata. The establishment samples are pooled to determine the selected establishments representing more than one product category. The establishments are visited, coded by industry, and transaction samples selected from each product stratum the establishment was selected to represent. Each transaction selected must then be classified under a product heading, and coded as well with a destination (source) country. Product XMPIs can be compiled using all prices for each product, regardless of the industry in which the establishments are classified or the destination (source) of the transaction. With two-stage sampling of this sort, some accuracy of the industry and source and destination XMPIs will be sacrificed.

F.4 Stratification

6.75 It is a well-known principle of sampling that stratification into segments for which the dispersion of price changes is lower (more homogeneous) than the overall dispersion tends to increase the efficiency of the sample by reducing variance.

6.76 For example, in the two-stage sample described above, the list of price-forming units is first stratified by product, for example, the at the five-digit SITC or six-digit HS code. Each product stratum then can be further stratified by variables appropriate for that product group. The ideal variant for stratification is the value to be measured in the survey—that is, price change—but in practice we use proxy variables that we assume to be correlated with price change. For example, the size of the production unit may cause differences in production technologies and, thus, different responses to changes in demand or input costs.

6.77 In the U.S. XMPI, the sample design ensures that all units (that is, products or producers) above a certain size are included. The remaining units are sampled with probability of selection proportionate to size (PPS). The alternate approach of setting broad strata, such as those with value of sales of 1 million to 5 million, 5 million to 10 million, etc., will result in units within each stratum having an equal chance of selection and, when selected, an equal weight. In a PPS sample design, a unit with 5 million in sales will have roughly a five times greater chance of selection than a unit with 1 million in sales. Further, the unit falling into the sample on a PPS selection would have a weight proportionate to its size, an additional improvement over broad stratum sampling.

6.78 Ideally, stratification should be optimized to minimize sampling errors. For example, the number of strata (L) can be optimized based on a relationship such as

$$(6.1)\quad V(\overline{y_{st}}) = \frac{S_y^2}{n}\left[\frac{\rho^2}{L^2} + (1-\rho^2)\right]$$

where S_y^2 is the variance of the variable being estimated (y), in this case price change, n is sample size; and ρ is the correlation between y and the variable used for stratification, in this case a proxy for price change such as output or sales.

---

can be sorted and stratified much more effectively than the original frame as a result of the information collected in phase one.

F.5 Sample allocation

6.79 Given that there is always an upper limit on the amount of data that can be collected because of resource constraints, decisions must be made about how to allocate the data collection between the strata—that is, we must decide how many establishments to sample in each stratum and how many prices to collect from each. It is generally more expensive to increase the number of establishments sampled as opposed to increasing the number of prices collected from each establishment, although simply increasing the latter may add little to accuracy when intra-establishment (within establishment) variance is low. So, it is generally the case that the number of establishments to be sampled is the constraint, rather than the total number of prices collected.

6.80 Ideally, the sample allocation would be optimized so that accuracy is maximized within the cost constraint, according to some equation linking sample size with accuracy. For example, the simplest form of optimal allocation is to make the sampling fraction (f_h) in a stratum (h) proportional to the standard deviation S_h in the stratum, and inversely proportional to the square root of the cost (c_h) of including a unit from that stratum in the sample—that is,

$$(6.2)\quad f_h \;\alpha\; S_h / \sqrt{c_h}$$

Thus more heterogeneous and cheaper strata are sampled at higher rates. Often, costs do not differ between strata, so the optimum allocation reduces to $f_h \;\alpha\; S_h$, the so-called Neyman allocation.

6.81 If probability sampling techniques have been used, it is possible, in theory, to estimate variances at each level. Typically, the frame first is stratified by six-digit product codes. Two-stage PPS sampling is employed to select establishments within each code and then transactions from each establishment. Establishment industry classification and the destination/source of the product transactions are recorded at initiation of the establishment.

6.82 The variance of each XMPI will depend on the variance between (inter) establishments producing a product, and the variance within (intra) each establishment in the sample. The intra-establishment variance might be because of differences in variety or terms of transaction but it likely will be relatively small compared to the inter-establishment variance. So, an optimization model will allocate the sample of establishments in proportion to the variance within strata, but will suggest collecting a fair low number of prices for each product from each establishment.

6.83 Calculation of the variances of the product XMPIs is complex, and thus the optimization algorithm also is complex. There are variances between establishments in each industry, and within each product stratum in each establishment in the sample.

6.84 The above examples assume that probability sampling techniques are used and that variances therefore can be estimated. In sample surveying, however, we usually assume very limited information about the frequency distribution followed by sample measurements. This means that in practice, optimization often is done using a variety of pieces of information, applied to more or less formal optimization models. Information that may be available includes the following:

- The total sample size that resources allow;
- The number of units in each product frame;
- The economics of each market, that is, the value of output, company and product composition, product dispersion, price-setting mechanisms, etc.
- Which XMPIs need to be published—it may be necessary to allocate larger sample sizes to some strata industries or products than simple empirical methods would indicate in order for XMPIs to be published at a detailed level without fear of breaching confidentiality guidelines; and
- Response rates.

6.85 The aim often is simply to produce product indices with comparable accuracy and to publish a reasonable amount of industry and destination/source detail. As for the number of prices collected from each establishment, it may be necessary to use a general rule, such as the average number of prices should be around 4 or 5 with no single establishment providing more than 15 or 20.

G. An Example of Sample Selection and Recruitment of Establishments

6.86 For sample selection to proceed, all of the earlier steps of sample design must have been completed. Decisions have been made on the sampling techniques to use at each stage of the sampling process. Assume for simplicity that manufactured goods have been chosen as the first area to be included in the XMPI. (Subsequently, mined ores, agricultural products, power and other public utility services, transport services, etc, may be added.) For this purpose, information on establishments such as industry, output, sales, name, and location is available from a recent Census of Manufacturing or a Census of Establishments. Products at the six-digit HS level have been selected using a cutoff sampling strategy. All product groups with imports greater than 0.02 percent of total imports, f.o.b. have been chosen. (The cutoff value–0.02 percent– is determined by the market share considered significant within the country. If the number of product groups is too large given the resources available, a higher cutoff threshold may need to be used.)

6.87 In addition, the production of quite a few products is concentrated among a few large enterprises, while others have less concentrated production. It would be helpful to stratify the industries by size of firm. In those industries where production is highly concentrated among a few large enterprises (for example, three firms represent 90 percent of production), the

large enterprises are selected. In those industries with a more disperse concentration, the largest firms could be selected with certainty (that is, chosen with a probability of 1.0) while a sample of smaller firms could be selected using random sampling techniques (for example, PPS sampling as described below). In general, the number of sampling units for the smaller firms should increase as the concentration ratio (percentage of industry output by large firms) becomes smaller. For example, for industries where the concentration ratio is 70 percent, a sample of four units among the smaller establishments might be adequate, but if the concentration ratio is less than 50 percent, the number of units might be twice that size. Using such a process also requires that appropriate weights be assigned to each selected unit. For the certainty units, the weight would be the firm output (sales), while for other units it would be the sampling interval (see example below).

6.88 At this point the frame is stratified, allocations of sampling units have been made, and the sampling technique has been decided upon. Usually, three phases are left to sample selection:

(i) Select establishments;
(ii) Recruit establishments; and
(iii) Select transactions.

G.1 Selection of establishments using probability techniques

6.89 The sampling frame of establishments has been stratified by four-digit industry and size for probability sampling (purposive sampling could be used instead, and some of the issues involved in this are discussed under "Selecting Products and Transactions in the Establishment"). In this situation, either systematic or PPS sampling could be used, or a combination of the two. A common application of PPS is to assign a probability of 100 percent to units in the largest size strata (as discussed above), and then select randomly from each of the other strata, with probability of selection proportionate to size.

6.90 A combination of systematic sampling and PPS is used in the United States, where a stratum frame would be ordered by size and cumulative totals calculated. For example, assume that we know the average cost per establishment for collecting price information, and that the costs will not vary significantly by industry. Based on this information, we determine that the number of establishments in the sample would be 400 (total data collection costs divided by average cost per establishment). If the industry for which we are drawing the sample represents 1.0 percent of the total sector output, then we would allocate four establishments to the industry (400 x .01), and we can proceed to draw the sample from the frame. Assume the information below in Table 6.3 is available from the sampling frame.

Table 6.3. Step 1 for Establishment Sample Selection

Establishment Identifier	Size (value of production in millions)	Cumulative Size	Cumulative Percent
E	200	200	34
C	100	300	52
D	80	380	66
B	60	440	76
G	50	490	84
F	40	530	91
H	30	560	97
A	20	580	100

The sampling interval is calculated:

$$\text{Sampling interval} = \frac{\text{cummulative grand total}}{\text{number of sample units}}$$
$$= \frac{580}{4} = 145.$$

6.91 All establishments with production values greater than the sampling interval (145) have 100 percent probability of selection and are known as "certainty units" (Establishment E). These selected units are removed from the frame, we recalculate the cumulative size, and a new sampling interval is calculated using the reduced frame and the remaining number of sample units to be allocated (as shown in Table 6.4).

Table 6.4. Step 2 for Establishment Sample Selection

Establishment Identifier	Size (value of production in millions)	Cumulative size
C	100	100
D	80	180
B	60	240
G	50	290
F	40	330
H	30	360
A	20	380

$$\text{Sampling interval} = \frac{\text{cummulative grand total}}{\text{sample allocation}}$$
$$= \frac{380}{3} = 127.$$

6.92 If there are new certainty units in the reduced sample, these are removed (not in this case) and the process is repeated until a sampling interval is calculated for which there are no certainty units. This sampling interval is used for systematic sampling. The remaining sample is sorted (largest to smallest as shown in Table 6.4), a random number between 0 and 1 is generated, and the sampling interval is multiplied by this random number to give the starting point for the sampling pattern.

Table 6.4. Step 3 for Establishment Sample Selection

Establishment Identifier	Size (value of production in millions)	Cumulative Size
C	100	100
D	80	180
B	60	240
G	50	290
F	40	330
H	30	360
A	20	380

Random number = 0.34128

Starting point: 0.34128 x 127 = 43

Sampling pattern:

43	(43 + 127)	(43 + 127 + 127)
43	170	297

Thus Establishments C, D, and F are selected, giving a total sample of C, D, E, and F.

6.93 The weights assigned to each establishment would be as follows. Establishment E will have a weight of 200. It was selected with certainty, and it will maintain the same weight because it is representing itself in the sample. Establishments C, D, and F will each have a weight of 127 because they are representing all the other establishments not selected in the sample. Thus, the total of their weights must be the total of all the noncertainty establishments, which is 380 in this example. Additional detail on the source of weights and methods for proportional allocation of weights within establishments to products is presented in Chapter 5 Sections D and E.

6.94 An alternative approach is used in some countries is to use cutoff samples so that a certain level of output or sales is achieved. For example, there may be a desire to have the sample represent 70 percent of the output in each industry in the sample. In such a case, a cutoff sample is used. Establishments in the product sampling frame are ranked in order of the output (largest to smallest). The percentage of output that each establishment represents to the total for the product group is calculated. The cumulative percentage then is derived. A cutoff of 70 percent is established, so that all establishments below this threshold in the cumulative rankings are dropped and the sample will consist of those remaining. This approach guarantees that the sample consists of large establishments.

6.95 In the previous example if one used the cutoff procedure, establishments E, C , D, and B would have been selected because their cumulative percentage of output is 76.

G.2 Selection of products and establishments in small developing countries

6.96 An alternative sampling strategy that statistical offices in developing countries have applied successfully is to use a cutoff approach for selecting products and establishments.

(i) The first step is to tabulate the import (export) data from customs records for one or two recent years to derive total value of shipments by eight-digit HS code or five-digit SITC. These represent the products for external trade in the country's recent experience. (One-of-a-kind shipments and shipments of personal effects should be deleted.)

(ii) Order the products that remain from highest to lowest in terms of their value of shipments and calculate the total value. This represents the value of regular recurring imports (exports).

(iii) Compute each product's (again at the 8-digit HS code) relative importance by dividing the value of shipments for the product by the total value of recurring shipments.

(iv) Calculate the cumulative relative importance for each product.

(v) Apply a cutoff threshold of 70 percent so that all products with a cumulative total less than or equal to 70 percent are accepted for the product sample.

(vi) Tabulate the same information by establishment and select the establishment sample using a cutoff sample with a threshold of 70 percent.

(vii) Select those establishments that import (export) the products selected in the product sample. The result is a sample list of important establishments that import (export) the selected products.

6.97 The resulting sample normally would represent approximately 50 percent of imports (exports). The statistical office should recruit the sampled establishments and select representative transactions within the establishment for the products identified.[7]

G.3 Recruiting establishments

6.98 Recruiting an establishment means securing the cooperation of its staff (particularly if the survey is voluntary), so that data will be of a high quality. It is highly recommended that each establishment receive a personal visit during which the purpose and function of the price survey are explained, and the sample of transactions or varieties to be priced is selected. Supplementary data for weighting transactions also can be collected during the visit. All these tasks can be more effectively carried out via personal visits rather than via telephone calls or mailed questionnaires.

G.4 Selecting products and transactions in the establishment

G.4.1 Probability and cutoff sampling procedures

6.99 The probability approach also can be used for selecting products and transactions by soliciting information from establishment records. Once in the establishment, however, the respondent may be reluctant to provide detailed records for selecting products and transactions. One alternative would be to ask the respondent to list the products produced and provide an estimate of the percentage each product represents of total sales. This information can be used to select the sample by ranking the products from highest to lowest and then making the selection using the same techniques discussed above.

6.100 Another alternative, if the respondent is unwilling to provide product percentages, is to ask him or her to rank the products in order of importance. Using the ranking information, estimated percentages can be established. Consider the information in Table 6.5 that is provided by a respondent in an establishment with eight products:

[7] In small countries, there may only be a few products that are exported. In such a case, all the large exporters are known to the statistical office and the sample of exporters would be those few key establishments. In order to publish product level indexes in this case, the statistical office may have to seek approval from the exporter due to confidentiality requirements.

Table 6.5. Selection of Products Using the Ranking Method

Product	Ranking	Importance	Estimated Percentage	Cumulative Percentage
G	1	5	33	33
H	2	4	27	60
I	3	3	20	80
J	4	2	13	93
K	5	1	7	100
Total		15	100	

The respondent was able to rank the products in order of importance. Each product can then be assigned its importance based on the reverse order of its ranking: Product G is assigned 5, Product H is assigned 4, etc. Next, an estimated percentage of sales is calculated using each importance as a percentage of the total of the assigned importances. Assume that the sample design indicates that three products are wanted for this establishment. These percentages can then be used to select a sample of products through the probability sampling procedures described above or through cutoff sampling procedures.

6.101 If probability procedures are used, the sampling interval is first calculated:

Sampling interval = 100 / 3 = 33.
A random number is selected to determine the starting point and the sampling pattern:
Random number = 0.45814
Starting point = 0.45814(33) = 15
Sampling pattern = 15, 48 (15 + 33), and 81 (48 + 33)

The selected sample will be Products G, H, and J. (Note that we do not select Product I because it is below the third interval in the sampling pattern.)

6.102 If the cutoff procedure is used, the first two or three products (G, H, and I) will be selected depending on whether we used a 50 or 70 percent threshold. With the cutoff procedure the most important products are selected.

6.103 In addition, representative transactions for continuous pricing need to be identified. The respondent should be asked to supply information on various transactions that apply to the selected products. Again, the data can be in the form of actual values from company records, estimated percentages, or by ranking. If two transactions per product are required, then the same procedures as those just described would be followed to select the two transactions. (Normally, we would expect to have two–four transactions per product depending on its importance and the number of transactions the establishment normally has

for each product. Because respondent burden may be problematic, we would not want to have much more than 20 observations per respondent.)

6.104 In the above examples, if the respondent could not provide any information or if he or she says that they are all equally important, then equal probability would be assumed. In such a case, each product or transaction would be assigned the same importance (that is, 100 divided by the number of products), and the selection procedure would continue as explained above.

G.4.2 Purposive sampling

6.105 Since the selection will be based largely on the judgment of the members of the establishment's staff present at the recruitment meeting (respondents), it is important that these people are knowledgeable and hold senior positions, probably from the marketing, sales, or accounting departments.

6.106 The first step is to stratify by products produced by the establishment selected for the product group sample. As a general guide, it is reasonable to have between 3 and 10 detailed product strata (depending on the size of the establishment) that are deemed representative of the establishment's output. It should be possible to obtain a sales figure or estimate for each stratum, or at least to order the strata by size. If the establishment is both an importer and exporter the product strata should be further stratified between imports and exports. Separate prices should be collected for exports and domestic products, as necessary.

6.107 Then for each stratum, one or two specific transactions should be chosen, bearing in mind the general rule that the average number of prices from establishments should be around 4 or 5, with no single establishment providing more than 15 or 20 (strata may have to be combined if the number is too large). The aim is to choose transactions and terms of sales that account for a significant proportion of sales, are broadly representative of other production, and are expected to be available for sale or stay in production at future price collections.

6.108 Weights for each transaction selected could be determined by proportional allocation of the establishment weight to each product and transaction selected. This procedure is discussed in Chapter 5 Section B.1

G.5 Recording product specifications

6.109 After transactions have been selected, the price-determining characteristics must be carefully discussed and recorded on the collection form. (See Chapter 7 for more details on recording product specifications.) Examples of such characteristics are as follows:

Product specifications:
Type of product

Brand name or model number; and

Main price-determining characteristics–size, weight, power, etc.

Transaction specifications for the XMPI:

- Type of buyer–exporter, wholesaler, retailer, manufacturer, government;
- Type of contract–single or multiple deliveries, orders, one-year, agreed volume;
- Unit of measure–per unit, meter, ton;
- Size of shipment–number of units;
- Delivery basis–free on board, sale with or without delivery to customer;
- Type of price–average, list, free on board, net of discount, cost-insurance-freight; and
- Type of discount–seasonal, volume, cash, competitive, trade.

H. Sample Maintenance and Rotation

6.110 Price surveys are panel surveys in that data are collected from the same establishments on more than one occasion. The general problems with such surveys are that the panel becomes depleted as establishments stop producing, the panel becomes increasingly unrepresentative as time passes and the universe changes, and some establishments may resent the burden of responding and leave the panel or provide poor quality data. All these problems cause bias.

6.111 A widely used method to alleviate some of these problems is to limit the length of time that establishments stay on the panel by using some form of panel rotation.[8] Rotation has two main benefits: (i) it ensures that most producers participate in the survey for a limited time and, therefore, the burden is shared among enterprises, and (ii) it helps to alleviate the problems caused by a sample being out of date—that is, sample depletion and not being representative of current trends. Recruiting new establishments helps to ensure that new products are represented in the price surveys.

H.1 Approaches to sample rotation

6.112 Obviously sample rotation has a cost since new panel members need to be recruited. There are several options regarding how rotation might be done. First, a rotation rate should be fixed. For example, if the whole panel is to be rotated every five years, then the annual rate is 20 percent. This could be implemented by dividing the industry headings into five groups and dealing with one group each year. Or 20 percent of all respondents, across all industries, could be dropped each year and replacements recruited. An establishment's

[8]In many countries, the rotation is limited to the smaller respondents, for whom it is felt that responding to surveys imposes a significant burden. This need not be the general case, and the use of full-panel sample rotation is encouraged.

rotation cycle could be related to its size, so that larger establishments stay in the sample for more than five years, and small establishments stay in for fewer than five years.

6.113 If sample rotation is done by product group, this provides a good opportunity to review the sample design and reallocate and select new establishments as necessary. Rotation and sample revision fit best within a system of annual chain linking in which the product structure and weights can be updated each year.[9]

6.114 Where cutoff samples are used, the same procedures as those described in Section G.2 could be completed as each years custom's records become available to identify new establishments and those that have become more important in external trade. New establishments that now fall within the cutoff threshold can be added and old establishments that fall outside the threshold could be dropped. Thus, the sample would reflect the most current establishments and products that important to the country.

H.2 Procedures for introducing a new sample of establishments

6.115 The procedures used to introduce a new sample of establishments are similar to the overlap procedure used for linking replacement price observations or introducing a new product structure in a weight update. Assume the rotation strategy calls for replacing 20 percent of all products. If the XMPI sample consists of 100 six-digit product groups, then each year the statistical office will replace the samples in 20 product groups. For each of the targeted product groups, a sampling frame is needed to select a new sample of establishments. The staff must then recruit the establishments, as discussed in Section E.2.

6.116 The new establishment sample will have new weights for the selected establishments, products, and transactions. The new sample and weights will be used directly to replace the old sample. During the same month, the data collection staff will have to collect price observations for both the new and the old sample. The old sample prices are used to calculate the index in the usual way, and the new sample will provide new base-period prices to calculate the index the next period using the new weights. For example, the old sample for a particular product market may consist of five establishments and 20 price observations, while the new sample may have eight establishments and 32 price observations. Both samples are collected during the overlap month, that is, 13 establishments with 52 price observations (assuming no establishment in the old sample is also in the new). The 20 observations from the old sample are used for the current period index calculation. The 32 price observations for the new sample provide basic data for setting new base prices in the new sample.

6.117 The index formula used will influence the relationship between the price reference period for the weights and the reference period for the base prices. If the statistical office compiles a Lowe or Laspeyres index, it will use the first set of prices collected in the new

[9] Annual weight update is not a requirement for sample rotations; it simply makes the process a bit easier because weights already are being updated at most levels of the index. When there is no system for annual weight updates, sample rotation does require a two-tier system of weights—fixed weights at higher levels of aggregation for aggregating to higher-level indices and separate weights for low-level indices that are updated periodically.

sample to set the base prices for the index. The base price reference period and the weight reference period need to align if the Laspeyres price index is used. If the weight reference period for the establishment and product weights are, for example, annual shipment values for 2005 and the prices collected for the new sample are for June 2007, then the new prices will have to be estimated backwards to the annual average for 2005. This is accomplished by applying the price change for the product between June 2007 and the annual average for 2005 to the June 2007 price observations. For example, if the prices in the product (over all establishments) rose by 10 percent between the annual average index for 2005 and the June 2007 index, then each price observation would be deflated by the factor 1.10. This calculation adjusts the new price observations for the average price change in the product market between the weight reference period and the current period.

6.118 Consider a similar example for the Lowe index. Again, assume that the weight reference period is for 2005 and that the base price reference period is December 2006. In this case, the statistical office will need to update the weights for price changes between the 2005 annual average and December 2006. The price index for the product group is used to calculate the price change between 2005 and December 2006 and this price change is applied to all the weights. Next, the June 2007 prices will need to be adjusted backwards to December 2006. The product price index is used to measure the price change between December 2006 and June 2007. This price relative then is used to deflate the June 2007 price observations to obtain December 2006 base prices.

6.119 If the estimation method is a Young index, the process is much simpler because the new weights are used directly in the computation of the index using the new prices without any adjustments. (See Chapter 16, Sections D.2 and D.3 for a discussion of the Lowe and Young indices.)

6.120 These procedures assure that the new prices and weights are consistent with the index number formula within each six- to ten-digit product group selected for sample rotation. For higher-level indices, the weight reference period may not be the same as for the products going through sample rotation. In practice, the aggregation weights used to combine group level indices may have a different price reference period than for the sample rotation groups. For example, the product group weights used to produce higher-level indices (three-digit, two-digit, etc.) may have a reference date of 2005 because they come from customs records for 2005. The index reference period might also be 2005 = 100, because of a statistical agency policy to re-reference index numbers once every five years. On the other hand, the weights from the establishment sampling frame used to draw the rotated sample may be for 2006, because the weights for the rotated product groups are taken from the most recently available customs' data or services trade survey (perhaps with a special supplement for product groups scheduled for sample rotation). The price index reference period could be December 2006 because the price information is readily available from sample respondents.

6.121 Thus, there can be a difference between the base price reference period for the new sample at the lowest level (elementary aggregate)—December 2006—and the index reference period for higher-level indices—annual average for 2005. In such cases, the price change from the lower level indices will be used to move the higher-level indices forward to the current period. For example, in product 8411.11.40 (8411 Turbojets, turbopropellers and

other gas turbines, and parts thereof: Turbojets: 8411.11 of a thrust not exceeding 25 kN: 8411.11.40.00 Aircraft turbines) the index level in December 2006 was 108.0, and in September 2007 it was 110.2 with an index reference period 2005 = 100. The sample of 10 establishments and 40 price observations for this product was rotated in January 2007 using base prices from December 2006. The elementary indices for the products in this industry have a price reference date of December 2006. To estimate the product index, the statistical office will have to use the price change from the new sample and link it to the level of the higher-level index. This can be done in two ways, depending on whether the statistical office uses a direct or chained price index formula (see Chapter 10, Section B.3). Assume a direct index is used where the current price for October 2007 is compared to the base price in December 2006, resulting in a price index of 102.96 (December 2006 = 100). The long-term price relative (1.0296), times the product 8411.11.40 price index for December 2006 (108.0), gives the October 2007 index level of 111.2. Alternatively, if the monthly chained index form is used, where the October prices are compared to the September prices, then the lower level index is linked to the September 2007 higher-level index. Assume the one-month price relative was 1.0091 in October 2007. The September 2007 product 8411.11.40 index (110.2, where 2005 = 100) is multiplied by this price relative to derive the October 2007 product index of 111.2. The results of the formulas should be the same. The advantage to using the monthly chained index form is that it facilitates making quality adjustments as discussed in Chapter 8, Section C.3.3.

I. Summary of Sampling Strategies for the XMPI

6.122 The approach to a sampling strategy in the XMPI requires a number of steps to gain enough information and design a survey that will produce reasonable estimates of price change within the level of resources provided. The following points provide a logical sequence to the sampling issue as presented in this chapter.

(i) *Determine the survey objectives, uses, coverage, and resources before determining the data to be collected, the periodicity of collection, and the type of sampling that will be employed.*

6.123 It is important to decide at the beginning of the process if price changes for product and/or destination (origin) will be needed as well as products and the degree of accuracy required. It will also be important to decide whether monthly or quarterly indices will be produced. These, in turn, will determine the level of resources allocated to the program. Alternatively, if there is a fixed level of resources available, it is possible to work with cost controls to determine affordable sample sizes and collection frequency at the expense of accuracy.

(ii) *Identify sources to use to develop a sampling frame for selecting the establishments and products for covered sectors and industries.*

6.124 The availability of up-to-date business registers with appropriate selection parameters (for example, industrial codes and measures of size) could serve as a source for developing sampling frames for selected industries. Many of the sources of weight data discussed in Chapter 5 also could be used to develop a sampling frame. For international trade in goods

tariff and export declaration administrative records are an excellent source for a list of establishments engaged in international trade. Others include industrial census, bank surveys, and administrative records.

(iii) *Use probability sampling techniques to the extent possible.*

6.125 While probability sampling throughout the selection process is a desirable goal, it may not be entirely affordable. An alternative is to use cutoff sampling at various stages in the process.

(iv) *To make the sample more efficient, use multiple levels of stratification within the sample design.*

6.126 In most cases, two strata will be identified within the sample—product and establishment. However, the sample could be more efficient and representative if additional strata are used, such as establishment industry code, establishment size (large, medium, and small), region or location (if there are price trend differences by location within country), and export versus domestic market production (if there are price trend differences for these markets). Additional strata will be helpful to the design wherever there might exist differing price trends or price variability within the chosen strata.

(v) *The price sample should be based on actual transactions with the characteristics of those transactions fully described.*

6.127 Often there is a tendency to use average prices or unit values (sales value ÷ quantity sold) as the price reported in the XMPI. These are not true transaction prices, in that they represent the average of a number of transactions for which there could be differences in quality or pricing characteristics. Therefore, it is important to select a sample of individual transactions with a detailed description of all of the characteristics that determine the price. These transaction prices and their characteristics then will be observed through time. So-called transfer prices between related establishments that may have little or no relationship to market values can be especially problematic for internationally traded items, particularly between countries with large trade volumes and integrated economies, such as the United States, Canada and Mexico, or the member countries of the European Union.

(vi) *Initial recruitment of establishments should be completed by personal visits.*

6.128 Initial sample recruitment should be conducted through personal interviews with establishment managers in order to accurately select representative products and transactions. The purpose of the survey must be explained, along with the need for the continuous reporting of price data for the selected transactions.

(vii) *Samples of establishments and products must be maintained so the reliability of the XMPI remains intact. A program of sample maintenance is needed for this purpose, and sample rotation may be also be desirable.*

6.129 Products produced by establishments will frequently change in response to market conditions. Also, establishments will cease operations and new ones will begin production. The XMPI sample sizes must be maintained in order for XMPI estimates of price change to be accurate. Therefore, it is necessary to have a program targeted toward keeping the sample intact and the products representative of current trade in terms of both the goods being produced and the establishments producing them. Fortunately, for international trade in goods, there is a large volume of current administrative information to draw upon for constructing the establishment frame.