Chapter 4: Sampling

Introduction

1. In order to construct a perfectly accurate CPI, the price statistician would need to record the price of every variety of all goods and services that are in scope of the CPI. Because it is in general too costly and in practice impossible to regularly record all the prices of the universe in a timely manner, sampling techniques are used to select a subset of prices that eventually enter the index compilations. Consequently, a CPI is a sample statistic. There may be some exceptions, such as tariffs, where all prices can readily be collected, for instance, directly at a central level and, consequently, no sampling is involved.

2. The sampling universe for a CPI has three dimensions:

- The product dimension, that is, all goods and services available for purchase.
- The geographical and outlet dimension, that is, all places and outlets where a product is sold.
- The time dimension, that is, the sub-periods of the index.

3. In practice, sampling in a CPI follows a multi-stage approach. The universe of products is structured by first selecting the items within the different sections of the expenditure classification. For each item of the CPI classification, one or more representative items can then be sampled. In order to address the geographic dimension, first specific locations for price collection are selected. In a second step, outlets are then selected within the sampled locations. The specific varieties to be priced are in general only selected in the sampled outlets. Finally, time can be considered as another sampling dimension as it must be decided at which moment during the reference month prices are observed. In general, less attention is given to the time dimension as long as price variation between collection periods can be considered to be small.

4. Within each of the different sampling stages, either probability or non-probability sampling can be considered, possibly in combination with some kind of stratification. The sampling frames required for probability sampling are not always available, so the price statistician either has to create one or rely on non-probability sampling techniques. For this reason non-probability sampling has been heavily used to draw samples for price collection in the CPI. However, the use of some form of probability sampling is in most cases the preferred option as it avoids the need for arbitrary decisions and ensures unbiased results.

5. The samples should be representative of the universe. When sampling designs are planned, the full universe of locations, outlets and outlet types, items and product varieties belonging to the scope of the CPI should be taken into account. All significant parts of that universe should be appropriately represented. An additional complication is that representativity is not something static but evolves over time. The samples that were initially designed for the price reference period may not be fully representative anymore of the current period. That is why it is important that samples are continuously monitored and updated as needed. Chapter 7 deals more specifically with the challenges of a dynamic target universe.

6. The overall objective of the sampling procedures is to minimize the Mean Square Error (MSE). The variance due to sampling should be minimal and the indices should not be subject to any bias. The sample design itself should allow publication of sub-indices at all levels which have been decided upon, such as regional indices or separate sub-indices for urban and rural areas. Sampling

designs should be efficient so that a maximal sampling precision is obtained for a minimum fieldwork and processing costs. Even without any formal variance estimation, sample allocation should be reasonably optimized by taking into account the weight and the magnitude of the price change variance of the sub-indices.

7. The design of CPI samples is in practice a complex exercise. The imperfections or sometimes absence of adequate sampling frames and the practical constraints of price collection make it difficult to systematically apply proper probability sampling at all stages. The multi-stage nature of the sampling process and the relationships between the different sampling dimensions must be well understood by the index compiler. In addition, there are often practical considerations that must be taken into account when it comes to organizing the price collection in the field. Consequently, some judgement is often needed and compromises must be made. In the end different sampling approaches may be adopted for different parts of the CPI basket. It is this thus essential that sampling procedures are clearly defined and well documented.

Sampling techniques

8. In survey sampling, there is a distinction to be made between the parameter and the estimator. In the context of a CPI, the parameter is the target price index number that is based on prices and quantities covering the whole universe. The estimator is the price index that is actually compiled using the sampled data as inputs. The result of the estimator depends on the particular price index formula, that may or may not use weights, and on the sampling scheme that has been adopted for selecting the items for which prices are collected. In practice, the parameter is unknown although simulations can for instance be conducted on Scanner Data to study the performance of different sampling strategies.

9. Several different estimators may be proposed for the same population parameter. In assessing the quality of a sample estimator, i.e. how well it estimates the parameter, two measures are often considered in the probability sampling paradigm. The first measure is the bias of an estimator, which is the difference between the universe parameter and the average of the estimator over all possible samples that could be drawn under the specified sample design. An estimator is unbiased if it has zero bias. The second measure is the variance of the estimator with respect to this sampling distribution.

10. An estimator is considered accurate if both its bias and variance are small; that is, the estimator is on average very close to the parameter and does not vary much from its mean. This is what is measured by the Mean Square Error (MSE) that can be defined as the sum of the variance and the squared bias. In addition, samples should be efficient. This means that maximum sampling precision should be obtained for minimum fieldwork and processing costs. It is likely that in a CPI, bias is a much greater problem than sampling error. The deterioration in precision caused by small samples can thus be offset by spending sufficient efforts and resources in selecting and maintaining representative samples.

11. There are two approaches that can be used to sample units: probability sampling techniques, which require that the probability of a unit being part of the sample is known in advance and that this probability is strictly positive for each unit, and non-probability techniques. Units can represent the locations and the outlets where households shop or the items and product varieties that households buy. Often a mixture of probability sampling and non-probability (purposive) sampling is used at the various sampling stages of a CPI. These techniques are often applied together with some kind of stratification.

Probability sampling techniques

12. Probability sampling amounts to selecting a sample of n units from the universe of N units by attaching a non-zero inclusion probability π_j to each unit j. The inclusion probabilities π_j for each unit in the sample are assumed to be strictly positive and known in advance. This requires a sampling frame, that is, a list of all the units eligible to be sampled. A frame may have over coverage to the extent that it includes units that are not in the universe or includes duplicates of units. It may have under coverage to the extent that some units in the universe are missing from the frame.

13. In simple random sampling (SRS) and systematic sampling each unit is sampled with equal probability and we have $\pi_j = n/N$. In simple random sampling (without replacement), a random number is assigned to each unit in the frame and the n highest (or lowest) values are selected. In systematic sampling, the sampling units are selected at equal distances from each other in the frame, with random selection of only the first unit. These techniques are usually recommended in situations where the units are relatively homogeneous.

14. In probability proportional to size (PPS) sampling the inclusion probability $\pi_j = \frac{nx_j}{\sum_{k \in N} x_k}$ is proportional to some auxiliary variable x_j. Units for which the probability is larger than one are selected with certainty, while the inclusion probabilities for the remainder of the units are calculated after excluding the large ones. This technique is recommended if the units are of different sizes, so that larger units are more likely to be selected than smaller units.

15. In a CPI context, the auxiliary variable is typically the household final monetary consumption expenditure on the goods and services covered by the CPI. In practice, this variable is often not available and so alternative variables must be used that are supposed to be correlated to expenditure. The extent to which this assumption holds eventually determines the quality of the sampling design.

16. In a CPI, the size of a sample is often fixed a priori. It is therefore impractical to let the sample size be randomly fixed by the sampling procedure. Several techniques exist for drawing fixed sample size samples, one of which is systematic sampling.

17. In systematic PPS sampling, the list of units is first ordered randomly and the cumulative total of the auxiliary variable x_i is calculated. Selection of units then takes place using interval sampling with the interval value calculated by dividing the cumulative total $\sum_{i \in N} x_i$ by the sample size n. This is done by generating a random starting point between zero and the interval value in order to select the first unit. The second random number is generated by adding the interval value to the starting point. This is then used to select the second unit. This process of adding the interval value to the previous random number, and selecting the corresponding units, is repeated until the requisite number of units has been sampled.

18. If the size of a unit is larger than the interval value, then it is selected with certainty. These units are then removed from the sampling frame and the whole process of cumulating the size variable and compiling an interval value is repeated with the remaining units. Systematic sampling can also be applied to SRS by simply assuming that all units have exactly the same size.

19. Systematic sampling is best explained by an example. In Table 4.1 we show how a sample of 3 outlets can be drawn from 10. In a first step, the outlets are put in a random order. Although turnover would be the preferred selection variable, it is often not available from the sampling frame. An alternative might be the number of employees as a proxy for turnover. We look at the list, where we have included the cumulative sizes and the inclusion intervals. We take the total number of our size measure, which is 90 in this case, and divide it by the sample size, 3. This gives us a sampling interval of 30. We next choose a random number between 1 and 30. Say that we get 25. The sample will then

consist of the outlets whose inclusion intervals cover the numbers 25, 25 + 30 = 55 and 25 + 2x30 = 85.

Table 4.1. Systematic sample of 3 out of 10 outlets, based on probability proportional to size

Outlet	***Number of employees = x***	***Cumulative x***	***Inclusion interval under PPS***	***Included when starting point is 25***
1	13	13	0–13	
2	2	15	14–15	
3	5	20	16–20	
4	9	29	21–29	X
5	1	30	30	
6	25	55	31–55	X
7	10	65	56–65	
8	6	71	66–71	
9	11	82	72–82	
10	8	90	83–90	X

20. PPS sampling has the advantage of selecting the sample in proportion to the relevant variable. It therefore ensures a sample which reflects the heterogeneity of the population. So in this sense the sample does not need to be re-balanced by re-weighting. Alternatively, if each unit could be given an equal chance of selection in the sample, then re-weighting may be necessary. For instance, SRS would give a large outlet the same chance of selection as a small independent shop despite the enormous differences in turnover and so re-weighting will be needed.

21. Let us assume that the prices that underlie the index compilations have been obtained under a specific probability sampling design. It can then be shown how common price indices are estimators for certain population price indices. If the size variable in PPS sampling corresponds to the expenditure of the item in the base period, then the sample Jevons price index is an approximately unbiased estimator for the population Geometric Laspeyres price index. Similarly, if the size variable corresponds to the quantities in the base period, then the sample Dutot price index is an approximately unbiased estimator for the population Laspeyres price index. These results suggest that items should be sampled based on their expenditures if using a Jevons index whereas for a Dutot index, items should be sampled based on their sold quantities. Finally, the Carli index is an unbiased estimator for the population Laspeyres price index if the size variable corresponds to the expenditures in the base period.

22. These estimators are “approximately” unbiased in the sense that the bias tends towards zero if sample size increases. It can be shown that the finite sample bias of the Jevons index is always positive. In particular, under SRS, the sample Jevons price index is on average larger than the population Jevons prix index. For the Dutot index, the sign of the finite sample bias is undetermined although its magnitude turns out to be often negligible. For the Jevons index however, the bias may be significant if sample sizes are very small. Consequently a caution against very small samples in a stratum may be warranted when the geometric mean is applied. Increasing the sample sizes will not only reduce finite sample bias, but also improve the variance due to sampling. As a rule of thumb and unless price change variance is very low, at least 8 to 10 observations should be selected for each elementary aggregate.

Non-probability sampling techniques

23. In many circumstances, formal sampling frames do not always exist. That is why non-probability sampling techniques are often applied because they do not necessarily rely on sampling frames. Moreover, such techniques allow, to a certain extent, more control over the samples than random approaches. For instance, the costs and the practical feasibility of price collection can play a role when designing samples for a CPI. In addition it has to be kept in mind that the sample has to be priced not only once but continuously over the next periods. Examples of non-probability sampling techniques that are often applied in a CPI context are cut-off sampling, quota sampling and the representative item method.

24. Cut-off sampling refers to the practice of choosing the *n* largest sampling units with certainty and giving the rest a zero chance of inclusion. As with PPS, the sampling frame must contain an appropriate variable that measures the "size" of the unit. Moreover, the quality of that variable has to be judged when performing cut-off sampling. The term "cut-off" refers to the threshold value between the included and the excluded units.

25. The disadvantage of cut-off sampling is that it does not produce unbiased estimators, since the small units may display price movements which systematically differ from those of the larger units. Any estimator resulting from cut-off sampling has zero variance because for a given threshold, there is only one sample that can be drawn. In terms of MSE, cut-off sampling might thus be a good choice if the variance reduction more than offsets the introduction of a small bias.

26. Cut-off sampling can be considered when the size variable is highly skewed. The goal is to select the largest units that are not so numerous and to disregard the many small units. This design is also useful when the "below cut-off section" of the universe is considered insignificant and perhaps difficult to measure. A particular CPI practice that can be seen as cut-off sampling is for the price collector to select the most sold product in an outlet, within a centrally defined specification. In this case, the sample size is one (in each outlet).

27. In the example in Table 4.2, the 3 largest outlets have been selected using cut-off sampling, the size of the outlet being measured by its number of employees. This cut-off sampling technique allows covering 54% of the total number of employees. If the cut-off rule is defined so to cover at least 70% of the total number of employees, then two additional outlets must be selected. Instead of fixing the sample size, one may also define the cut-off rule in terms of the cumulative size variable for a certain threshold. Contrary to PPS sampling, an outlet with a small number of employees will systematically be excluded in cut-off sampling. At the same time, it ensures that all the largest units are covered.

Table 4.2. Cut-off sample of 3 out of 10 outlets

Outlet	***Number of employees = x***	***Cumulative x***	***Cumulative x (%)***	***Included in cut-off-sampling when n=3***
6	25	25	28%	X
1	13	38	42%	X
9	11	49	54%	X
7	10	59	66%	
4	9	68	76%	
10	8	76	84%	
8	6	82	91%	
3	5	87	97%	
2	2	89	99%	

5	1	90	100%	

28. Quota sampling is another non-probability sampling technique used in a CPI. Many product groups, even rather small ones, are quite heterogeneous in nature, and the price varies according to a large number of subgroups or characteristics. There may well be different price movements going on within such a product group, and a procedure to represent the group by just one or a few tightly specified product types may then carry an unnecessarily great risk of bias.

29. In quota sampling, the actual selection of the sample is done using judgmental procedures with respect to known and relevant characteristics, such as the product group, when choosing products to price or the type of outlet to select. The sample is drawn in such a way to have the same proportions as the total population or universe to ensure that the sample is representative. Consequently the sample becomes self-weighted. Quota sampling is a stratified sampling with a sample allocation that is proportional to the stratum weights and where the sampling within a stratum is conducted in a judgmental way.

30. The following example illustrates the concept of quota sampling. If there is a characteristic that allows dividing a product group into three product types, each of which represent 30%, 20% and 50% of household expenditure, then the sampled prices will also represent these same proportions. If there are 10 prices to be collected, then a quota sampling approach recommends that 3 prices are collected for the first product type, 2 for the second product type and 5 for the third product type. If the Dutot index is used, then quantities should be used instead of expenditures for defining the importance of the strata.

31. In the “representative item method”, one or several item specifications are defined that are meant to represent the whole item. For instance, “Spaghetti” may be a representative item for “Pasta”. In this example, the price changes for pasta will be measured using the sampled price changes for spaghetti. According to this method, only products matching the specification are priced and no products falling outside the specifications will enter the index. The sampling of representative items for a CPI is usually purposive or judgmental because of the lack of adequate sampling frames that formally lists all the possible product types.

Stratification

32. A common sampling technique used in a CPI is to divide beforehand the universe into disjoint homogeneous sub-populations or strata. An independent sample of appropriate size is then selected for each stratum using any of the probability or non-probability sampling techniques described so far. A particular probability sampling technique is stratified systematic random sampling. This means that systematic random sampling is conducted independently within each of the pre-defined strata.

33. Stratification can be applied to all the sampling dimensions relevant for a CPI. In practice, a CPI is typically stratified by geography (e.g., region, city, rural or urban), by type of outlet or by outlet, and by item or product. Samples are set up within each geographic area, for each of the different shop types, and for each item/product within shops. Suppose that the CPI classification contains the item “refrigerators” which could be stratified by region. In a given region, refrigerators are primarily sold in larger specialized shop chains and in a number of small independent shops. In addition, refrigerators could be classified according to their capacity. Such a stratification structure is illustrated in Table 4.3. Within each cell, prices must then be sampled using for instance a two-stage approach. In a first step, specialized chains and independent shops that sell this type of goods are selected in the different regions. In the selected outlets, specific refrigerator models that meet the product type specifications are then identified for continuous pricing. In addition, expenditure weights need to be estimated for each stratum using for instance HBS data, retail trade statistics or other data

sources. These weights can then be used to aggregate the stratum indices into a price index for refrigerators.

Table 4.3. Stratification by region, outlet type and product type

	Region 1		Region 2	
	Specialized chains	Independent shops	Specialized chains	Independent shops
Small capacity				
Large capacity				

34. Formally, let us assume that a sub-component of the CPI is composed of several strata and that the sub-index for stratum k is labelled I_k. The sub-component's estimate is then obtained using the strata's weights and indices :

$$I = \frac{\sum_{k \in K} w_k I_k}{\sum_{k \in K} w_k}$$

where w_k is the weight of stratum k. The weight of the stratum corresponds to the whole expenditure of that stratum, and not only to the (smaller) expenditure of the locations, outlets or varieties that have been sampled within that stratum.

35. Ideally, the stratification should be designed in a way that the sampling error is minimized. The idea is to construct the strata in a way that the variance of the price changes within a stratum is low, whereas the variance between the strata can remain large. A low within stratum price change variance also has the advantage that results of different price indices are likely to be similar. Consequently, the choice of the elementary aggregate price index matters less.

36. At the same time, stratified sampling avoids selection bias by ensuring that the whole universe is appropriately represented. It controls for the fact that prices must be collected for each and every specified stratum. Without stratification certain parts of the universe may be completely ignored. In addition, the use of explicit weights guarantees that the prices are weighted according to the importance of each stratum. In price statistics, stratification thus helps reducing the bias of the estimates.

37. Stratification is a useful strategy to make the sample more efficient. The allocation of the number of prices across the strata is further discussed in the section on variance estimation and optimal allocation**.** Furthermore, the availability of detailed sub-indices can be convenient for meeting specific publication needs.

38. Although the universe can exhaustively be divided into a set of strata *K*, there are circumstances where only a sample of these strata *S(K)* is used in actual index compilation. Only prices belonging to the sampled strata *S(K)* will eventually be collected. The strata can be selected purposively. A strategy could consist in selecting only the most important strata, which is equivalent to running a cut-off-sampling. The bias that results from this depends on the weight of the omitted strata and on the difference between the price change of the omitted strata and the other sampled prices. If the weight of the omitted strata is small or if their price change is similar to the price change of the sampled prices, then the bias may be small as well.

39. If the strata composing the index structure are selected using probability sampling, then in principle the aggregation of the strata indices should be adjusted using sampling weights. In particular, if the expenditure is used as size variable in PPS sampling, the higher-level estimate is compiled as the

simple arithmetic average of the sampled stratum indices. In the example shown in Table 4.1, three outlets were selected using PPS sampling. Let us assume that the outlets 4, 6 and 10 each represent a stratum in the index structure. The weight associated with each stratum does not correspond to the actual weight of that outlet. Instead, an unbiased estimate is obtained by equally weighting the three stratum indices:

$$I = \frac{30}{90} I_{Outlet\_4} + \frac{30}{90} I_{Outlet\_6} + \frac{30}{90} I_{Outlet\_10}$$

40. A complication arises in systematic PPS sampling when the size of a stratum is larger than the interval value. In such a situation, a distinction must be made between the larger strata that are selected with certainty and the other strata that are sampled after the larger units have been removed. The higher-level index is then obtained as a weighted average of two sub-indices. The first sub-index covers all the certainty strata whereas the second sub-index contains a sample of some of the lower weighted strata. In the first sub-index, the actual weights of the certainty strata must be used in the aggregation as these strata only represent themselves. In the second sub-index, the remaining sampled strata are aggregated with equal weights.

Sampling stages in a CPI

41. In practice, national statistics institutes often adopt four levels of sampling in a CPI: locations; outlets within locations; items within different sections of the expenditure classification; and product varieties. These four steps will be discussed below. The product dimension is related to the outlet dimension. At some point, the sample of items must be matched to the sample of outlets. It must be decided which items are priced in which outlets. Scanner data can be used as a sampling frame for both outlets and products. Time is an additional sampling dimension that must be considered. If timing is relevant, it may be necessary to incorporate timing elements in the outlet or in the product specifications.

Sampling of locations

42. The objective of the first sampling stage is to select the locations where prices are going to be collected. A location is an area where outlets can be found that sell products to households. Ideally a large part of the CPI product basket should be available to purchase within each location. In order to obtain a full basket it may be necessary to create locations by combining out-of-town shopping areas that mainly sell non-food items with a neighbouring city in which food is available. The sampling of locations can be omitted if outlets are directly sampled independently of where they are located.

43. In order to define the list of possible locations, a decision must be made on the geographical scope of the price collection areas. Price collection can for instance be limited to the capital city only or to the main cities of a country or of a region. Some countries restrict price collection to urban areas because of the challenges to collect prices in rural areas. Alternatively, it can be decided to select locations belonging to both urban and rural areas. Under the broadest geographical scope, locations must be sampled so that they are representative of the whole national territory.

44. The sampling of locations is typically conducted using some kind of regional stratification. Location selection can then take place separately within each region. For instance, the strata used for sampling can be aligned on the administrative sub-divisions of a country and possibly take into account the distinction between urban and rural areas. In principle, stratification works best if the locations of a stratum are homogeneous. In that sense, the territory could also be partitioned by grouping together the locations that are situated close to each other or that have similar socio-

demographic characteristics. If no regional stratification is used, locations are sampled directly at the national level.

45. Within each region, some locations, such as the main cities, may be selected with certainty, in which case there is no randomness involved in the selection process. Locations can also be selected randomly using for instance systematic PPS sampling. To apply this kind of technique, a variable that measures the size of the locations must be available and the number of sampled locations within a region must be fixed beforehand. As information on expenditure made by households in outlets that are situated in a given location is rarely available directly, proxy measures must be used. The number of households living in a location can for instance be used as a size variable, although the residence of a household does not necessarily coincide with the place of purchase. Population size can be obtained from population registers or from a recent Census. The number of locations to be selected in each region can be determined as the proportion of national expenditure, income, or a corresponding proxy measure such as regional GDP, in that region, multiplied by the total number of locations to be visited nationally.

46. There are of course operational implications for setting up a new price collection in the sampled locations. That is why the process described above may need to be modified to take into account practical considerations. For example, collection costs are an important consideration and it may not be efficient to send a price collector to a small isolated location where many of the items in the CPI basket are not available for pricing. Location re-sampling is likely to be conducted on a less frequent basis than for instance outlet or product re-sampling. Sample coordination techniques can be used to adjust the initial selection probabilities in order to increase the likelihood of an overlap between the old and the new sample of locations.

Sampling of items

47. Cut-off sampling can be applied to select the items that compose the CPI basket. Expenditures for items are typically obtained from Household Budget Surveys (HBS) (see Chapter 3). However, not all items for which expenditures are available must necessarily be included in the CPI basket. Thresholds can be defined in order to identify low expenditure items that that do not justify a costly price collection although they do belong to the scope of the CPI.

48. In practice, an expenditure share is estimated for all the items that could potentially be included in the basket. If this share is below a given threshold, the item could be excluded. It might be decided, for example, to exclude from the index calculations items with an expenditure share lower than, say, 0.1 per cent for food items and 0.2 per cent for non-food items. A lower minimum threshold for the food items might be set because the prices for these products tend to display greater variability and because prices for food products are normally less expensive to collect. If an item is excluded, its expenditure may be redistributed to another expenditure group that is similar in terms of content and price development. Alternatively, the expenditures may be completely excluded from the calculation of the weights.

49. The cut-off procedure can be conducted at the national level or separately within a region so that the list of items used in each region may slightly differ. At the same time such an approach ensures that the regional baskets are representative of the consumption structure in each region. Manual adjustments may also be needed so that an item that falls below the threshold can nevertheless be included in the basket. For instance, it may not be so burdensome to collect a few additional prices if these are collected together with prices for more important items. The addition costs may thus be acceptable given the advantage of achieving wider product coverage.

50. The descriptions of the items that are eventually included in the CPI basket can be more or less broad. This largely depends on how detailed expenditures were recorded in the HBS. In any case, the universe of products and varieties that fall within the scope of an item is often quite large. Consequently additional sampling is required at this level for the selection of the specific varieties to be priced over time.

51. If information is available from other sources, a finer stratification can of course be developed below the item level. The strata may be constructed in a way to group together products with common properties or which are deemed to serve a common purpose. They may or may not be organized in a hierarchical manner. The product strata may form a partition of the universe. Alternatively, only a subset of the universe can be covered, in which case the selected product strata represent the non-selected ones.

52. In order to provide effective control over the sample of prices collected in the field, a list of product types can be drawn up by the central office using specifications. This is what is referred to as the "representative item" method. The whole product group is thus represented by a sample of specified product types. The fact that representative items are selected at a central level ensures that they are representative of consumer behaviour as a whole rather than one particular shopping location. Consistent with the principle of a fixed basket they are normally reviewed periodically at the same time as the weights and chain-linking is carried out. The problem with this method is that the national specification may often be unavailable in a number of locations.

53. A particular representative item can be framed either by loose or by tight specifications. For instance, it could be decided to collect prices for "Tea", which is a rather loose specification as it encompasses many different tea varieties that could eventually be selected to be priced. Alternatively, the representative item could be specified very tightly as a precise variety of tea defined maybe by a brand, a flavour and a package size. The discussion between loose and tight specifications also has implications for the choice of the elementary aggregate formula (see Chapter 8). In practice, a compromise with semi-tight definitions is often the preferred option.

54. If specifications are too tight, the price collectors may have difficulty in finding the products and this will lead to fewer price quotes and a deficit in the sample. There is also a risk that tight specifications will miss important parts of the market and consequently results can be biased. That is why the representative item method is more suitable for homogeneous product groups. In heterogeneous groups, it is more likely that important segments of the product universe, with different price movement, will be left out.

55. Loose specifications on the other hand give the price collector some freedom in choosing locally popular products and varieties and to adjust the sample to match local conditions. It will normally lead to greater representativity of the sample as a whole. At the same time, such an approach can be more challenging for price collectors who have to be appropriately trained to select in the outlet a particular variety to be priced that matches the loose specifications. Consequently this requires more formal procedures dealing with the selection of varieties within an outlet. Even with loose specifications, the variety eventually selected in the outlet has to be recorded sufficiently tight by the price collector in order to ensure constant quality pricing.

Sampling of outlets

56. There are different data sources that can be used as sampling frames for selecting outlets. An outlet should be understood here in a wide sense as any retail channel that sells goods or services to households, including different types of physical shops but also for instance the offices of service providers of utilities where they may not be a physical shop which customers can visit. Outlets may be selected only in those locations that were previously identified as price collection centres.

57. A starting point for selecting outlets is business registers which are data bases recording the business units that are resident in a given area. They may contain information such as location, type of business activity, turnover, and employment. Business registers should be updated regularly both for new businesses and just as importantly for closures. One limitation is that business registers may only collect data at the level of an establishment or an enterprise so that the information at the local level is not always available or regularly updated. Provided that some size measure is included in the information, such as number of employees, business registers can nevertheless provide a useful basis for preparing a PPS sample. Similarly, VAT registration records may be another source that has a wide coverage and that contains timely information on outlets. These tax records typically refer to a legal unit which can be composed of one or several outlets. With VAT data, the selection of random samples can then be based on gross sales as the measure of size.

58. Administrative records kept by local government, associations of businesses or market place managers are another potential source for outlet sample selection. These records can be used to create frames and might be particularly useful for sampling local markets. Depending on the presence of information on size, these sources might be able to provide a frame for PPS sampling. Telephone directories of businesses normally contain less information, for instance just business name, address and business activity. They do not include any size information. Therefore they are useful for simple random sampling or systematic sampling but not for PPS sampling unless additional information is sought, for example by visiting the outlets.

59. Information on outlets can also be obtained from specifically designed surveys. In particular, the objective of point-of-purchase surveys is to collect information on the place where purchases are made by households and on the amount of these purchases. HBSs are sometimes designed to indicate, on the survey form, the actual outlets from where the households purchased the goods and services recorded. Surveys that do include this information could be used as a source of relevant outlets for survey frames for various regions. As these surveys are based on a sample, they are however unlikely to provide an exhaustive list of all possible outlets and the conclusions of such surveys may become quickly out-of-date.

60. In the absence of existing sources as suggested above, it is necessary for the outlets in a location to be enumerated to provide supplementary information for the sampling frame or to construct a sampling frame from scratch. This enumeration could be carried out by price collectors or their supervisors, visiting each location and noting details of all retail outlets found. It can be a very costly activity and the price statistician will need to weigh the costs against the benefits in terms of a more representative sample. To reduce costs the price statistician may decide to limit the number of outlets enumerated per location, for instance by enumerating only a sub-district of a location. The ideal size measure of an outlet would be turnover. But a proxy may be used where an estimate of this is not readily available. For instance, the approximate net retail floor space, as estimated by the outlet enumerators, may provide an adequate alternative. It is also important that the assortment sold by the outlet is recorded in order to match the outlets with the CPI product basket.

61. If none of these more formal sampling techniques are feasible, outlets must be selected in a more judgmental way. Field staff can be asked to identify the largest retail and service outlets where households typically shop for a given type of product. Additional guidance may be provided to the

price collector, for instance by specifying the type of outlet or by delimiting the area where outlets are to be selected. This ensures that the overall sample of outlets remains representative of where households actually shop. This method of sampling often involves the price collector visiting the outlet to gather the necessary information on the outlet's characteristics prior to decisions being made on its inclusion in the sample and before the first price collection is conducted. For instance, if prices for children's clothes are to be collected, a price collector may visit a clothes shop initially selected from a central list to check whether it sells children's clothes and not just adult's clothes.

62. The sampling strategy for outlets typically relies on stratification by outlet type. Within each stratum, outlets are sampled either randomly or purposively if no appropriate sampling frame is available. The number of outlets to be sampled in each stratum could for instance be proportional to the weight of that stratum. In any case it is important that all significant outlet types are covered. If relevant, the Internet can also be seen as a particular outlet type or stratum from which specific websites are then selected.

63. As similar price trends can be expected in the outlets that belong to the same chain, it may be a good idea to also stratify by retail chain. For instance, one could decide to design a first stratum that contains all the major retail chains as sub-strata and a second stratum that covers the remaining independent shops. Outlets can then be selected for each retail chain separately and from the remaining universe of independent shops. In the end, the index should be designed so that the retail chains and outlet types are properly represented according to their market share.

Sampling of product varieties

64. In a final step, the selection of specific products or varieties within outlets is in general made by the price collector in a purposive way. The objective is to select the most sold products that fit the item specifications. This procedure essentially corresponds to a kind of cut-off sampling. In addition to being a high volume seller, the product is also expected to be available over time so that the same product can easily be priced during the following periods.

65. A compromise must sometimes be made between representativity and comparability. On the one hand, it is important to select products with high sales as this will ensure that the samples are representative. On the other hand, future availability of the product has to be taken into account in order to improve comparability over time by minimizing the need for replacements. There is however a risk of bias if the choice of varieties is mainly driven by the fact that they are convenient to collect. Price collection is further discussed in Chapter 5.

66. With broad item specifications, there can be many products available in an outlet which can possibly be selected for inclusion in the CPI sample. A more formal approach could then be adopted by designing a process that approximates PPS sampling with respect to the sales of each product. To apply this technique, all possible products or product types must first be enumerated and selection probabilities have to be determined for these products based on their sales. It may be possible to obtain information on sales directly from the respondents. Alternatively, shelf space can in certain circumstances be used to estimate sales proportions. Even with no information on sales, random sampling can still be conducted by simply assuming that all products have the same selection probabilities.

67. If the sales value measure is taken during a very short period, it may coincide with a special campaign with temporarily reduced prices. It could then happen that a product with a temporarily reduced price is given a large inclusion probability. Since this price will tend to increase more than average, an overestimating bias may result. It is thus essential that the sampling of the item takes place

at an earlier point in time than the first price collection or that sales values from an earlier period are used. Generally, annual sales values would be the best measure.

68. If no explicit weights are used in the elementary aggregate compilations, then the sample will be implicitly weighted by the number of observations. For instance, if prices of the same 2 varieties are collected in 5 outlets, then each outlet and each variety implicitly has the same weight. The elementary aggregate is often the lowest level for which weighs are available. However, efforts can still be made to improve the accuracy of this detailed sub-index using implicit or explicit weights. The sample could be designed in a way that the proportions within an elementary aggregate represent those of the universe for instance with respect to outlet type or outlet, if these aspects are not explicitly weighted in the CPI aggregation structure. Similarly, the sample should be well balanced with respect to product type or product characteristics. For instance, if local beer has an approximate market share of 80% and imported beer of 20%, these proportions could also be reflected in the elementary aggregate.

69. For some product categories, it may be possible to access additional data sources such as official registers or transaction data in order to select the individual varieties. If information on sales or on the number of items sold by individual product or by product type is available, formal sampling techniques such as PPS or cut-off sampling can be applied. For instance, data on car registration can be analysed in order to identify the most sold car models. For tariffs, it is often possible to obtain additional information on products directly from the providers, authorities or regulatory bodies. Such data can then be used to construct a detailed stratification and to sample the specific products to be priced. For a given item, the standard approach is that first outlets are selected and then varieties are selected in these outlets. If data sources are available in advance with all the varieties, this order may be reversed so that first varieties are selected which can then be priced in the different outlets.

Scanner data as a sampling frame

70. There are many ways how scanner data can be used for CPI compilation (see Chapter 10). In the first place, scanner data can be seen as a potential sampling frame for selecting outlets and individual products based on their sales. A small sample of products and outlets is drawn that can be easily monitored and maintained over time and that can be integrated with prices obtained through other price collection modes. The prices for the selected products can still be collected manually in the outlets. Alternatively, scanner data can also be used to replace the manual price collection while the overall CPI methodology remains largely unchanged.

71. In scanner data, both the product and the outlet dimension are closely related. A two-stage sampling approach can be adopted where first outlets are sampled, possibly belonging to the same retail chain or outlet type, and then products are selected within the sampled outlets. If price levels are similar between outlets or if data is not available by outlet, products can also be sampled at a more aggregated level, for instance at the level of an entire retail chain. Some product categories are characterized by high churn rates and some products may only be available for a very short period of time. The sampling frame could thus be restricted to more stable products that are available continuously during several months. This procedure must be adapted in order to sample seasonal products that are only available during certain periods of the year.

72. In principle, PPS sampling can be an appropriate technique to select products within a certain product category. A more pragmatic approach consists in using a cut-off sampling technique that selects the most sold products in each product category. The use of non-random sampling techniques can be motivated by the fact that the distribution of sales is often highly skewed, with a few products accounting for a large share in total sales. Moreover, such bestseller products are likely to be kept by

the outlets in their assortment in the future periods, which will reduce the need for imputations, replacements and quality adjustments.

73. After the initial selection, scanner data can also be used to monitor the quality of the samples over time. Typically, the sales of a product are continuously decreasing before the product disappears from the market. At the same time, a successor product may enter the market with sales that show an increasing trend. Such patterns can in principle be identified with scanner data, which then helps in selecting appropriate replacements. Some countries apply a more dynamic approach by re-sampling products every month using a specific cut off procedure. Expenditure shares are compiled for products that are available in two consecutive periods. Products then enter the sample if the average expenditure share in the two periods is above a certain threshold.

Sampling in time

74. In most circumstances, the moment of price collection is chosen purposively. The main principle is that prices should be collected each month at the same moment, during the same week or the same day of the month. If there is some price variation within a day, then it is important that prices are collected always at the same time of the day. Price collection can be limited to a certain week of the month, for instance during the middle of the month, or it can be spread over several weeks of the month according to some pattern, for example different weeks in different regions or for different product categories.

75. If prices are known to fluctuate a lot during a month, then a more frequent price collection should be considered. For such products, it can make sense to collect prices for instance once per week or sometimes once per day instead of once per month. This could be the case for energy products or for fresh food. Prices collected on the Internet can sometimes be highly volatile with some websites applying dynamic pricing strategies.

76. Some products have a time-dependent component that should be incorporated into the product specifications. For instance for airfares or other transport services, the price can be highly dependent on the day of the week, the times of the flight and how long in advance the ticket was purchased. Such timing elements should be held constant as this will ensure the comparability of the collected prices over time. Moreover, the specifications including their time-dependent components should be defined in a way to be representative of consumer behaviour. The flights should also be spread over the whole month in order to capture the price changes occurring within a month.

Sample maintenance

77. The samples are drawn at a particular point in time based on the information available at that moment. This typically happens when new weights are introduced. However in reality, the situation is seldom static. Products and outlets gain or lose in importance over time. New products constantly enter the market whereas other products may no longer be available. New outlets can open in a certain location whereas previously popular outlets are closed down. This eventually means that the sample that was set up at an earlier period may not be representative anymore of household expenditure at the current period.

78. In order to avoid that the quality of the sample deteriorates over time, it is recommended that samples are kept continuously under review and adjusted if needed (see Chapter 7). If resources do not permit a systematic full review of products and outlets, a rolling review could be set up where every year another part of the basket is examined. The focus should be put on those product and outlet areas that are likely to be affected by major developments.

79. In practice, feedback received from field staff can be taken up in order to maintain the on-going representativity of the samples. Price collectors and their supervisors should be encouraged to report if the item specifications used for price collection are no longer in line with the products on offer. Through their local knowledge, the staff in charge of price collection may identify new outlets that are becoming popular and that deserve to be included in the CPI samples. An increasing number of missing prices can also be an indication that the outlets or the product specifications are not appropriate anymore and action must be taken to adjust the samples. Some countries have dedicated product specialists that are in charge of monitoring the markets of complex product groups and can assist with sampling and quality adjustments.

80. One way of preserving the representativity of the samples over time is through replacements. Sometimes price collectors will find that an outlet at which they have been collecting prices closes or no longer stocks one of the products being priced. Disappearing outlets and products within a specific outlet are handled by replacing on a one-to-one basis when the outlet closes or the product ceases to be sold by the selected outlet. A more proactive approach can also be adopted by anticipating the expected disappearance of an outlet or a product so that the number of forced replacements is kept to a minimum.

81. The criteria for selecting replacements will vary. An outlet can be replaced on a like-for-like basis by another outlet of the same type, in the same or similar location, selling the same or similar range of items and products. In replacing a product one of two strategies is usually adopted. If the initial selection rule was “most sold” or “probability proportional to (sales) size” then the replacement could follow the same rule with the advantage of maintaining the representativity of the sample. However, the replacement product may not be comparable anymore to the product that it replaces. That is why quality adjustment methods have to be applied in order to measure pure price changes (see Chapter 6). There is a severe risk of bias if the differences in quality are not properly adjusted. Alternatively, the replacement may be made based on the product most similar to the one which disappeared, thereby reducing the need for quality adjustment. When replacements are performed, there is thus an inherent tension between representativity and comparability.

82. The dynamic nature of the target universe can also be addressed through sample rotation involving a full or a partial re-sampling. For instance, sample rotation is a technique where the length of time that outlets and products are included in the price surveys is limited by dropping a proportion of them after a certain period of time and selecting a new sample of outlets and products thereafter. The process of re-sampling can be based on any of the methods and on any of the sampling frames that are used for initial sample selection. Re-sampling can apply to products, outlets or even locations. Technically, re-sampling involves an overlap period where the first period of the new sample overlaps with the last period of the old sample. For instance, in an annually chained CPI, the basket can be re-sampled every year during the overlap month which links the series of two consecutive years.

Variance estimation and optimal allocation

83. A CPI is a statistic with a complex sampling design that involves different stages and that often relies on non-probability sampling methods. It is thus not a trivial task to estimate the variance of a CPI. Strictly speaking, variance estimation is only doable in case of probability sampling, as the sampling mechanism is known. To the extent that samples in a CPI are not probability based, variance estimates must rely on models in which some type of random sampling is assumed. Several countries have developed different approaches for variance estimation in a CPI.

84. It is useful to know the sampling error of a CPI as this can inform the users about the level of precision of the published indices. The interpretation of such error margins depends of course on the specific approach adopted for variance estimation. From an operational point of view, variance estimation is a useful tool for deciding on the optimal allocation of prices. If a variance and a cost function are known, an optimization program can be defined that minimizes the variance given the budget that is available for price collection. Alternatively, costs can be minimized under the constraint that a minimal variance must be achieved.

85. The final precision of a sample estimate depends only on the size and allocation of the sample and not on the size of the country, so in this sense there is no need for a larger sample in a larger country. Larger samples are called for if regional differences in price change are of interest and if the amount of product disaggregation that is desired in presenting the indices is very high. Of course, in large countries, the regional dimension may be more important and the budget allocated to CPI work may be larger, thus allowing for larger samples.

86. A very general form of a CPI is $I = \sum_k w_k I_k$, where *k* denotes products, w_k is the expenditure share of this product and I_k the product index. If the estimation of each product index were independent of each other the variance would be $V(I) = \sum_k w_k^2 V(I_k)$, where $V(I_k)$ denotes the variance of the product index *k*. However, product indices are not statistically independent. For instance, the price change of different products collected in the same outlet can be related to each other. The sampling errors of the product indices are therefore correlated, which means that this expression probably underestimates the total sampling errors to some extent.

87. If the representative item method is used for sampling products, the number of items must be sufficiently large to properly represent the diversity of products that can be found in a given product category. Prices that belong to the same representative item may be relatively homogeneous both in terms of price level and price change. What matters is the variance in price changes between the different representative items that could be selected. Sampling variance can be reduced by specifying more representative items in those product categories where the price change variance between the representative items is large.

88. Many resources are spent on price collection for producing a CPI. It is therefore worthwhile to devote some effort to allocating these resources in the most efficient way. A better allocation of prices can often help to improve the sampling error and consequently the quality of the CPI without an increase in the human and financial resources associated with price collection activities.

89. The general approach to distribute the number of observations across strata was established by Neyman. We assume that each stratum has a known weight w_k and that samples are drawn randomly and independently in each stratum. The standard deviation of the individual price changes (not the price levels) within a stratum is denoted by s_k. In addition, let c_k be the cost to collect one unit in stratum *k*. If the total sample size is *n*, the following allocation is optimal:

$$n_k = n \frac{w_k s_k / \sqrt{c_k}}{\sum_k w_k s_k / \sqrt{c_k}}$$

90. This expression formally shows that the number of observations in a particular stratum should be proportional to the weight of that stratum times the standard deviation of the price changes and inversely proportional to the square root of the unit cost. The expression can be simplified by assuming that costs are identical in all strata. The idea is that more price observations are needed for

those strata with a higher weight and an important variance in price changes. At the same time, it is sufficient to allocate fewer prices to the strata with a smaller weight and more similar price changes.

91. There are however challenges to systematically apply this type of allocation in practice. Whereas stratum weights are in principle available, the estimation of the standard deviation of price changes may be more problematic. One inconvenience of this approach is that it depends on the target price change. For instance, the prices can be compared to those of the previous month or to those of the same month of the previous year. The optimal allocation can be different if it is derived from monthly price changes or from annual price changes. Moreover, the standard deviation of price changes of a stratum is not necessarily constant over time.

92. Data sources, such as previous CPI surveys, must be accessed to estimate the standard deviations of the price changes of each stratum. There is however a risk that an estimation based solely on past observed prices data is biased. Consider the situation where a stratum is represented by a unique product with very tight specifications. The prices collected in the different outlets for such a tightly specified product may all change at a similar rate. This will consequently lead to a low sample price change variance. The true population price change variance may however be much larger because of the unobserved price change behaviour of the varieties that belong to the stratum but are not included in the CPI sample.

93. Consequently alternative approaches must be adopted if price change variances cannot be approximated in a satisfactory manner. Instead of allocating the same number of price observations to each stratum, an alternative approach consists in linking the number of prices to the weight, collecting more prices in those strata that have a larger weight. This will improve the sampling error of the sub-indices that have more impact on the higher-level aggregates because of their weight. In addition, the sample size per stratum should not drop below a fixed threshold to avoid formula bias or to guarantee a sufficient precision if stratum indices are separately published.

94. In a second step, the allocation can be adjusted by making judgements on the magnitude of the variance. This means that the sample sizes should be increased (decreased) for strata with an expected higher (lower) price change variance. For instance, prices for fresh food such as fresh meat, fish, fruits or vegetables are typically known to fluctuate a lot, even from one day to another. That is why larger samples are required for this type of products. On the other hand, prices for some services may only be adjusted once per year and so smaller samples can be sufficient. Finally, if the variation in prices across outlets is large, there is a need to draw a sufficiently large outlet sample in order to reduce the sampling variance

95. Costs can also play a role when defining sample sizes. It may be more time-consuming and complex, hence costly, to collect prices for some goods and services than for others. For instance for products where quality adjustment plays an important role, the price collector can be asked to record, in addition to the price, many technical characteristics. If the total budget is fixed, the higher data collection costs can imply having to accept smaller sample sizes.

96. In the example in Table 4.4, there are a total of 100 prices that have to be distributed over 5 items. A first strategy consists in collecting the same number of prices for each item. This allocation can be improved so that the number of price observations for each item is proportional to its weight. Because price change variance is expected to be low for item 3 but high for item 5, it may be appropriate to further adjust an allocation solely based on the item weights. For instance, it could be decided to decrease sample size for item 3 by 10 prices whereas the number of prices to be collected for item 5 increases by the same amount. Out of these three options, the third allocation aims at reducing the sampling variance while keeping the total number of prices collected unchanged.

Table 4.4. Different allocation strategies

			Number of prices		
Item	***Weight***	***Price change variance***	***Option 1***	***Option 2***	***Option 3***
1	10%	Medium	20	10%*100= 10	10
2	40%	Medium	20	40	40
3	25%	Low	20	25	15
4	20%	Medium	20	20	20
5	5%	High	20	5	15
TOTAL	100%		100	100	100

97. In order to design optimal samples, it must be well understood how the different sampling dimensions contribute to the CPI variance. Consider for instance a two-stage sampling, where first outlets are sampled and then varieties are selected within the sampled outlets. The sampling variance of the price index can be formally decomposed into a sum of two terms that are linked to these two stages. The first term relates to the variance of the price changes between the outlets, whereas the second term relates to the variance of the price changes of the varieties that are available within an outlet.

98. If the price changes between the outlets are rather homogeneous but the price changes of the different varieties that are sold in an outlet are more diverse, then the preferred strategy is to select a smaller number of outlets and a larger sample of varieties within each outlet. In addition, such a strategy is cost-efficient as it is often cheaper to sample an additional price in an already selected outlet than to increase the outlet sample. For instance, in some countries, car prices do not vary much across outlets for the same model and specification. On the other hand, there are a large number of models with different prices. One car retailer could easily provide prices for a number of models during the same collection period.

99. If prices within a big retail chain are similar across the country, then the number of sampled outlets belonging to that retail chain can be reduced. Instead, there is the opportunity to select more retail chains or to define larger samples that cover a wider range of products sold by that big retail chain. In any case, the prices must be appropriately weighted in the CPI. If a price quote from a big retail chain is combined with prices from other outlets within an elementary aggregate, this can for instance be achieved through implicit weighting by replicating the price of the retail chain an appropriate number of times. Alternatively, specific elementary aggregates must be constructed that are weighted according to the market shares of the different outlets.

100. The time dimension must also be taken into account. If prices vary a lot over time, it is preferable to collect more prices for the same product during a month. For example, petrol prices normally show small variation among outlets on the same day whereas prices can vary considerably even over a month. It would thus make sense to have a relatively small sample of outlets but follow prices several times in the same month. The time dimension is also highly relevant for purchases made on the Internet. As it is probably cheaper to consult websites than to visit physical outlets, it can be feasible to improve the temporal coverage of prices collected on the Internet.