

### 3 EXPENDITURE WEIGHTS AND THEIR SOURCES

#### Conceptual basis of the weights

1. A consumer price index (CPI) is usually calculated as a weighted average of the price change of the goods and services covered by the index. The weights are meant to reflect the relative importance of the goods and services as measured by their shares in the total consumption of households. The weight attached to each good or service determines the impact that its price change will have on the overall index. The weights should be made publicly available in the interests of transparency, and for the information of the users of the index.

2. The expenditure weights that are used in a CPI have to be consistent with the conceptual framework of the index. Expenditure data can be obtained from different data sources and in particular from Household Budget Surveys (HBS). When exploiting these data sources, the scope and the concept of the index will determine what should be included or excluded in the CPI weights. In order to meet multiple user needs, it is of course open to countries to compile supplementary CPIs. Separate weighting structures must then be estimated that reflect different concepts or coverages.

3. The geographical coverage of a CPI may follow either the “national” concept or the “domestic” concept. Under the national concept, the objective is to measure price changes experienced by the resident population, whether within the country or abroad. Under the domestic concept, the scope of the CPI (in terms of both prices and weights) extends to the whole economic territory and includes the consumption of both residents and visitors. The distinction between the two concepts is especially relevant if cross-border purchases are significant in a country. This also includes cross-border internet purchases. Under the domestic concept, decisions must therefore be made on the location of an e-commerce transaction. The choice for a domestic concept or a national concept depends primarily on the main users’ needs.

4. A HBS is generally capable of identifying all relevant expenditures made by resident households. These may or may not include expenditures made abroad. If the main purpose of the index is to measure price changes experienced by the resident population, the weights should in principle include their expenditures abroad. For practical reasons, even if weights cover the expenditures made both at home and abroad, prices may be collected only for the goods and services acquired in the economic territory of the country. Such an approach assumes that the price changes of the goods and services acquired abroad are the same as those for the same goods and services acquired at home. Alternatively, it may be possible to use CPI sub-indices compiled by the respective countries to measure the price changes of the goods and services acquired abroad. However, HBSs do not cover expenditures made by non-resident households. Sources other than HBSs must be used if the aim is to include the expenditures made by foreign tourists and to reflect all purchases of consumer goods and services made by resident or non-resident households within the country.

5. In principle, the weights should be representative of the whole country and all regions should be covered. Separate expenditure weights can also be derived for each region, in which case the expenditures must be sufficiently representative at the regional level. This is especially important if the expenditure pattern differs between the regions. For instance, it is common that the consumption habits vary between the urban and the rural parts of the country. The same principle applies to harmonized CPIs that cover a group of countries. Each country can then be seen as a “region” with its own national weights.

6. The discussion on the national versus the domestic concept likewise applies to the regional subdivisions. It can sometimes be the case that a household lives in one region but does most of its shopping in an adjacent region, particularly when a household lives close to a regional “border”. The question of whether the expenditure weights (and the prices) should be allocated to the region of expenditure or the region of residence is usually dictated by practical issues. In any case, the treatment should be consistent across all regions in order to avoid missing or double-counting parts of household expenditure.

7. The target or reference populations will be defined depending on the main purpose and use of the index. In principle, all types of households should be covered, irrespective for instance of their position in the income distribution. However, in some countries the wealthiest households are excluded because their expenditure may be atypical or the HBS information may be unreliable because of low response. Other countries may exclude the expenditures of the very poor. Another practice is to compile a CPI which excludes both extremes of the income distribution. If the CPI is used to compensate a certain subgroup of the population for increases in the cost living, then in principle, expenditure weights could be constructed for this subgroup only. In practice one CPI, that serves various users to the best extent, may be sufficient.

8. The weights may or may not include the expenditures made by people living in institutional households. Institutional households refer to people living permanently in an institution or who may be expected to reside in an institution for a very long time, such as members of religious orders living in monasteries or convents, long-term patients in hospitals, prisoners serving a long sentence or persons living permanently in retirement homes. Such people are treated as belonging to an institutional household when they have little or no autonomy of action or decision in economic matters. Many countries exclude such expenditure in their CPI because of the difficulty of obtaining reliable expenditure information, or because the expenditure associated with such households is unlikely to be very significant in comparison with non-institutional households. The choice to include or exclude the expenditures made by people living in institutional households may depend on the main source for the weights. Weights primarily based on National Accounts consumption information may include consumption by institutional households. However, if weights are primarily based on an HBS where this population group is excluded, it may also be excluded from the CPI coverage.

9. Depending on the use of the CPI, a decision must be made if non-monetary expenditures, such as production for own-consumption or remuneration in-kind, should be included in the weights or if it is more appropriate to limit the scope of the index to monetary transactions only. The production for own-consumption is treated by some countries as within the scope of the CPI, and in other countries as out of scope. In some countries, own-account production can account for a significant portion of household consumption. It can be argued that while it is definitely part of GDP (which values consumption regardless of its market mechanism) and should therefore be included to improve consistency with National Accounts and for the purpose of producing deflators, it is not necessarily appropriate for a general CPI or a CPI used for indexation where the narrowest concept of consumption that can be used is based on monetary expenditure. If own-account consumption is included in the CPI, the weights should include a valuation of the physical quantities of such products, the latter often derived from the HBS.

10. The treatment of owner-occupied housing is not straightforward and different approaches are discussed in Chapter 11. In principle, owner-occupied housing can be measured under the acquisition approach, the payment approach, or the use approach which is in practice often implemented as the rental equivalence approach. The weight given to owner-occupier housing depends on the conceptual approach that has been adopted. For instance, under the rental equivalence approach, a rent must be imputed which would be payable by owner-occupiers on the basis of the market rents payable for

accommodation of the same type and location. The weight would then reflect the imputed rentals of all the households that occupy their own dwelling.

11. In its role as an indicator of total consumer inflation, the CPI should in principle cover all types of goods and services that are consumed by the reference population. Some types of products may be excluded for practical reasons. These may include goods that are illegal, black market sales, gambling or prostitution. Some products may also be excluded because of policy reasons. For instance, it can be decided not to reflect price changes from tobacco in a CPI that is used for specific indexation purposes.

12. Only consumption expenditures are relevant for the construction of CPI weights. As explained in Chapter 2, outlays such as payments of social security contributions or income taxes, interest payments or repayments of debts, are irrelevant and should be ignored because they are not consumption expenditures.

13. Business related expenditures are explicitly excluded from the scope of a CPI. Consequently these expenditures must also be excluded from the CPI weights. In National Accounts, business related expenditures are considered either as Intermediate Consumption or as Gross Fixed Capital Formation. Households may engage in business activities from their home whereby part of the expenditures for some products are used partly to operate the business and partly for household consumption. In principle, only the portion that is used for household consumption should be included in the CPI weights

14. Given that the prices collected for the CPI are final prices paid by the purchaser, i.e. inclusive of the impact of all taxes and subsidies, the expenditure weights should also be inclusive of the impact of taxes and subsidies. Thus, the weights should be based on the price paid after any general subsidies by government are deducted, and after any taxes such as VAT, are added, i.e. the expenditure weights are based on actual purchase prices paid by the consumer.

15. The use of aggregated expenditure to derive weights reflects the principle that each household contributes to the weights an amount proportional to its expenditure. This is referred to as plutocratic weighting and means that the expenditure patterns of high-spending households have more influence on the index. The use of plutocratic weights is generally considered more appropriate particularly for consumer price indices which have been constructed to be a general macroeconomic indicator. In principle, it is possible to derive democratic weights, where each household is given equal weight, but these are not considered appropriate for an index used as a general measure of inflation. The use of equally weighted average household expenditure shares may be better suited to measure the inflation experience of different subgroups of households. Such an analysis is typically conducted by classifying households according to some socio-demographic variable such as income, age or educational level.

### **The weighting structure of the consumer price index**

16. The calculation of a CPI usually proceeds in two stages. In the first stage, elementary indices are estimated for each of the elementary aggregates. Elementary indices are constructed by (a) collecting a sample of representative prices for each elementary aggregate, and then (b) calculating an average price change for the sample. In the second stage, a weighted average is taken of the elementary indices using the expenditure shares of the elementary aggregates as weights.

17. Elementary aggregates are usually the smallest groups of goods and services for which expenditure data are available. They may cover the whole country or separate regions within the country. Likewise, elementary aggregates may be distinguished for different types of outlets. The nature of the elementary aggregates depends on circumstances and the availability of expenditure data. Elementary aggregates may therefore be defined differently in different countries. In general:

- Elementary aggregates should consist of groups of goods or services that are as similar as possible in terms of their purpose.
- They should also consist of goods or services that may be expected to have similar price movements. The objective is to minimize the dispersion of price movements within the aggregate.
- The elementary aggregates should be appropriate to serve as strata for sampling purposes in the light of the sampling regime planned for the data collection.

18. The weights are usually classified according to the *Classification of Individual Consumption by Purpose* (COICOP), used also in both the National Accounts and the HBSs. The aggregation structure for a typical CPI is organized as follows (see Figure 8.1 in Chapter 8) :

- First, the entire set of consumption goods and services covered by the overall CPI is divided into *divisions*, such as “food and non-alcoholic beverages”.
- Each *division* is then divided into *groups* such as “food”.
- Each *group* is further divided into *classes*, such as “bread and cereals”.
- Each *class* may be divided into more homogeneous *sub-classes*, such as “rice”.
- Finally, a *sub-class* may be further subdivided to obtain the *elementary aggregates*, by dividing according to region or type of outlet. In some cases, a particular sub-class cannot be, or does not need to be, further subdivided, in which case the sub-class becomes the elementary aggregate.

19. In some countries, a more detailed product classification is used that refines the products defined at sub-class level. These product categories are not part of COICOP itself but more detailed breakdowns of COICOP classes that are needed for CPI purposes and that depend on national circumstances.

20. The weights for the groups, classes and sub-classes are their shares in the total consumption expenditures of the reference population. The data sources used to derive these shares are further discussed in this chapter. In addition, the weight for a sub-class can be further stratified by region, by outlet or outlet type, or by a combination of both region and outlet. The elementary aggregate weights are the stratum weights according to expenditure class or sub-class, region and type of outlet. If no breakdown by region or outlet is carried out, then the sub-class as a whole becomes the elementary aggregate.

21. It is important that the weight of an elementary aggregate reflects the expenditure on the entire elementary aggregate and not the weights of the outlets and varieties that have been chosen to represent it. For instance, the weight for the sub-class “Rice” has to be based on the total expenditure made on rice, although the rice varieties selected for regular price collection only represent a fraction of this expenditure. Likewise, if an expenditure category is divided into two elementary aggregates according to outlet type, say, open markets and supermarkets, with corresponding market shares of food sales, 60% and 40% respectively, then these proportions have to be used to estimate the stratum weights, whatever the importance of the specific outlets eventually sampled.

22. Within each elementary aggregate, one or more products are selected to represent the price movements of all the goods and services in the elementary aggregate. For example, the elementary aggregate consisting of rice sold in supermarkets in the northern region covers all types of rice, from which parboiled white rice and brown rice are selected as a representative item. If the elementary product group is defined more broadly, more representative items might be selected in practice. Finally, for each kind of representative item, a number of individual varieties can be selected for continuous price collection, such as particular brands of parboiled rice. Again, the number of sampled varieties may vary depending on the nature of the representative item.

23. The methods used to calculate the elementary price indices from the individual price observations collected within each elementary aggregate are explained in Chapter 8. Working upwards

from the elementary price indices, all indices above the elementary aggregate level are described as higher-level indices that can be calculated from the elementary price indices using the weights of the elementary expenditure aggregates. The aggregation structure is consistent, so that the weight at each level above the elementary aggregate is always equal to the sum of its components. The price index at each higher level of aggregation can be calculated on the basis of the weights and price indices for its components, that is, the lower-level or elementary price indices. The individual elementary price indices are not necessarily sufficiently reliable to be published separately, but they remain the building blocks of all higher-level indices. Above the level of the elementary aggregate, therefore, no new information is introduced into the calculation of the CPI.

### **Regional weights and outlet type weights**

24. Regional weights may or may not be introduced into the CPI, depending on the size and structure of the country, data availability, resources and the purpose of the index. The rationale for introducing regional weights is to create more homogeneous entities which are likely to experience similar price movements and have similar consumption patterns. It may be necessary to distinguish different regions in federal countries because CPIs for the provinces or local states may be required for administrative or political purposes. In addition, in federal countries indirect taxes and hence price development may differ between the provinces.

25. In some countries, there may be quite large differences in consumption patterns and price developments between urban and rural areas. A common practice is then to introduce a stratification that distinguishes between urban and rural areas. This approach assumes that data sources are available that allow deriving expenditure weights separately for these two types of areas. Moreover, prices must be collected in both rural and urban areas in order to compile the respective stratum indices.

26. Where the weights derived from a HBS are available for rural as well as urban households and if price collection is limited to urban areas, one approach is to use the weights for urban and rural households combined, as this will normally improve the representativity of the index. This means that prices faced by rural households will implicitly be imputed by those collected in urban centers. This can be an acceptable assumption if for instance most of the monetary transactions made by households living in the rural areas take place in urban areas. Alternatively, if prices are only collected in urban centers, it could also be decided to likewise restrict expenditure weights to urban households and compile in that way an accurate urban CPI.

27. Within a given sub-class, the regional weight shows the consumption expenditure in the region in proportion to the expenditure in the whole country for that sub-class. For example, if 60% of the total expenditure on fresh fruits occurs in the North region and 40% in the South region, then these proportions can be used to derive the regional stratum weights. Let us suppose that at the country level, the expenditure share for fresh fruits is 5%. This share can then be split between the regions so that  $5\% \times 60\% = 3\%$  of the total national expenditure relates to fresh fruits in the North and  $5\% \times 40\% = 2\%$  to fresh fruits in the South.

28. Regional weights may typically be obtained from the HBS if the regional dimension has been appropriately incorporated in the sampling design of the HBS. When such estimates are not available from a HBS with a satisfactory degree of precision, population statistics are sometimes used to split household expenditure across regions. This approach is not ideal as it assumes that expenditures per capita or per household are the same in all regions. For instance, there are usually large differences between the urban and rural populations in the level and pattern of items that they consume. Finally, National Accounts data, if available at the regional level, can also be used to estimate regional weights.

29. There are different strategies to derive regional weights in practice depending on the availability and the quality of data sources. A bottom-up approach consists in first estimating household expenditures by region. These expenditures can then be summed up to obtain the expenditures at the country level. In a top-down approach, expenditures are first estimated at the country level before being distributed over the different regions.

30. Consistency should be ensured between the product and the regional dimensions. In the example in Table 3.1, expenditures are estimated for 3 items in 2 regions (Table A). The expenditure share of each stratum in the national index corresponds to the stratum expenditure divided by the national total (Table B). According to this example, region 1 has a total weight of 38.8%. At the same time, item 1 has a weight of 24.5% at the national level. If detailed product expenditures are only available at the national level, the regional product expenditures must be estimated. This can be done by using shares derived from the total expenditures that are made within a region. In Table C, we suppose that only item expenditures at the country level and total expenditure by region are known. The item expenditures for the country are then split up by region, assuming a weight of  $1900/4900=38.8\%$  for region 1 and a weight of  $3000/4900=61.2\%$  for region 2. The regional item shares in Table D now differ slightly from those in Table B.

*Table 3.1 Deriving expenditure weights by region*

<i>Table A</i>	Region 1	Region 2	Country
<b>Item 1</b>	400	800	1200
<b>Item 2</b>	500	1000	1500
<b>Item 3</b>	1000	1200	2200
<b>All items</b>	1900	3000	4900

<i>Table B</i>	Region 1	Region 2	Country
<b>Item 1</b>	$400/4900$ =8.2%	16.3%	24.5%
<b>Item 2</b>	10.2%	20.4%	30.6%
<b>Item 3</b>	20.4%	24.5%	44.9%
<b>All items</b>	38.8%	61.2%	100.0%

<i>Table C</i>	Region 1	Region 2	Country
<b>Item 1</b>	$1200^*$ $1900/4900$ =465	$1200^*$ $3000/4900$ =735	1200
<b>Item 2</b>	$1500^*$ $1900/4900$ =582	$1500^*$ $3000/4900$ =918	1500
<b>Item 3</b>	$2200^*$ $1900/4900$ =853	$2200^*$ $3000/4900$ =1347	2200
<b>All items</b>	1900	3000	4900

<i>Table D</i>	Region 1	Region 2	Country
<b>Item 1</b>	$465/4900$ =9.5%	15.0%	24.5%
<b>Item 2</b>	11.9%	18.7%	30.6%
<b>Item 3</b>	17.4%	27.5%	44.9%
<b>All items</b>	38.8%	61.2%	100.0%

31. Prices are collected from a variety of outlets and outlet types. In addition to the regional dimension, the sub-class may also be stratified by outlet or by outlet type. This can be especially useful if price levels and price changes significantly differ across different types of outlets. Information about the sale or market share of the outlets may be used to form elementary aggregate weights specific to a given outlet type. This type of information may for instance be obtained from retail trade statistics. In some countries, the HBS directly collects expenditure data by type of outlet which can then be used to estimate the stratum weights. Other potential sources to derive outlet or outlet type weights are point-of-sales surveys, scanner data or business registers. It may only be possible to establish outlet or outlet

type weights for a broader category of products so that the same proportions have to be used to disaggregate the expenditures of the different products.

32. In the example in Table 3.2, let us suppose that it was estimated that outlet types 1 and 2 have a market share of respectively 60% and 40 %, and that this same breakdown holds for both regions and for all three items. If a division is made according to type of outlet as well as region, then each item within a given region comprises two elementary aggregates: one for outlet type 1 and one for outlet type 2. In this example, the weight for the elementary aggregate for item 1 in region 1 sold in outlet type 1 would then be  $8.2\% \times 60\% = 4.9\%$ .

*Table 3.2 Deriving expenditure weights by region and by outlet type*

	Region 1		Region 2		Country
	Outlet type 1	Outlet type 2	Outlet type 1	Outlet type 2	
Item 1	8.2%		16.3%		24.5%
	4.9%	3.3%	9.8%	6.5%	
Item 2	10.2%		20.4%		30.6%
	6.1%	4.1%	12.2%	8.2%	
Item 3	20.4%		24.5%		44.9%
	12.2%	8.2%	14.7%	9.8%	
All items	38.8%		61.2%		100.0%

#### **Data sources**

##### ***Household budget surveys***

33. The HBS is in general the primary data source for deriving expenditure shares for categories of the goods and services covered by the CPI. As the HBS may have been designed to serve more than one purpose, it is desirable to ensure that the survey design also meets the requirements for the CPI. Coordination between the CPI and HBS staff is essential to ensure that CPI specific needs are taken into account.

34. The main requirements are that the survey should be representative of the CPI reference population and that it should include all types of consumption expenditures by households that are covered by the CPI. The sample size (number of households) for the HBS when used for the purpose of the CPI should be such that the expenditure data yielded are able to ensure statistically reliable weights at the elementary aggregate level. The HBS item list can be designed so that the information obtained maps directly into the CPI classification system. Ideally, the survey could also collect the data needed for deriving net weights for second-hand goods (see paragraph 85). The interview and recording periods should be appropriately distributed over time in order to ensure that annual estimates can be obtained which take into account possible seasonal patterns.

35. The expenditure values used for deriving weights should be consistent with the conceptual approach (acquisition, use or payment – see Chapter 2) adopted for the CPI. For some services, the moment when the service is actually consumed and acquired can differ from the moment when the service is paid for. The approach used in the HBS for determining the timing of consumption should be the same as the one used in the CPI compilation. For most goods, the moments of acquisition, payment and consumption coincide. Sometimes, the HBS focuses on the consumption of food products, which includes food purchased by the household but also own-produced food or food received for free. For durable goods, there is no difference between the acquisition approach and the payment approach as long as the purchase has not been financed with a loan. Once acquired, the use of such goods can span over several years. HBSs may collect different types of information regarding durable goods.

36. The HBS can theoretically provide information on regional breakdowns of expenditures and hence the weight data by geographical stratification. For this purpose, the regional dimension must be appropriately incorporated in the sampling design of the HBS. Depending on its design, the HBS may also be able to provide information on the types of outlet where purchases are made and the brands purchased. In particular, it can be possible to identify those purchases that are made on the Internet. Disaggregation of expenditure by type of outlet can be useful to construct elementary aggregates at a finer level of detail and to improve the sample design for shops and items for price collection.

37. Similar to a HBS, national food surveys are special surveys in which the primary emphasis is on collecting information on family expenditures for food products. These surveys provide a very detailed breakdown of food expenditures that can be used to derive the weights for detailed elementary aggregates.

38. The detailed expenditures obtained in a HBS can be subject to measurement errors. There is often under- or overreporting for those items that have a high prestige or that suffer from social stigma. Moreover, the respondent household may not correctly remember the details of all the expenditures that took place during a given recall period. Therefore, HBS estimates should be reviewed and evaluated for completeness of coverage with comparisons of secondary source data.

39. Given that most HBS's are probabilistic samples, statistical quality indicators (e.g. standard error and coefficient of variation) for the weights can be calculated. Such information can provide the index compiler with some direction on the structure of the elementary aggregates. If an analysis of the HBS shows, for example, that the expenditure data for lemons is of poor statistical quality, then a more broadly defined elementary aggregate such as citrus fruit could be considered instead. In a similar way, suppose that the HBS provides an unreliable expenditure for an item at a detailed local level. It may then be preferable to derive an estimate based on the expenditure share of that item within a broader geographical area. In order to find the best compromise between lowering the variance of the item expenditure estimate and biasing it towards the spending pattern of the broader geographical area, a composite share could be computed that averages the initial direct share with the share obtained for the broader geographical area.

40. Updating of weights depends on how often an HBS is conducted. For the purposes of the CPI, it is desirable for the HBS to be conducted annually. This will allow countries to revise and update their expenditure weights more frequently. Some countries conduct continuous HBSs with gradually rotating samples. A program of annual surveys with samples large enough to provide the type of estimates required for CPI weights can, however, be very costly. For this reason, some countries conduct large-scale surveys at ten-year or five-year intervals, perhaps supplemented with a smaller annual sample. Other countries distribute a large sample over several years. The average of the results over several successive years of smaller-scale surveys may provide a set of satisfactory annual estimates. The weights derived as the average of two or three years will also smooth any erratic consumer behavior over a short period, for example as a result of events such as droughts or floods, civil strife, oil price shocks, or exceptionally mild or cold winters.

### ***National Accounts***

41. National Accounts can be an alternative source for deriving CPI expenditure weights if reliable estimates for Household Final Consumption Expenditure (HFCE) are available. The practical advantage is that HFCE is updated every year, whereas a HBS may only be conducted on a less frequent basis. However, National Accounts may only be available at the national level so deconstruction of these data to provide a finer-level of detail or to produce regional expenditure weights may be necessary using other available sources of information.



42. The index compiler must understand the differences in scope and definition of consumption before using National Accounts data for CPI weights. If the CPI is restricted to monetary expenditures, then only a subset of HFCE must be used, excluding its non-monetary components. The CPI may adopt a different treatment for owner-occupied housing than the imputed rentals approach used in the National Accounts. The product coverage in the CPI may be smaller than the one in the National Accounts. In National Accounts, the household sector consists of all resident households, including people living in institutional households. HBSs, however, do not usually cover persons living permanently in institutional households.

43. National Accounts data may be used to improve HBS weights for products that are under-reported in the HBS. Note that National Accounts' figures for households' final consumption are usually based on statistics from the HBS and from several other sources such as domestic production, retail sales, tax information or import and export data. This means that National Accounts estimates are likely to be more accurate and reliable for estimating weights for consumption categories that tend to be wrongly reported in the HBS, and where results from the HBS suffer from a significant and distorting partial or total non-response rate.

44. In practice, weights for the main consumption groups can be obtained from the National Accounts down to a certain level of disaggregation. Each of these weights can then be further subdivided by applying the detailed HBS expenditure groups to the National Accounts consumption groups or classes. The combination of National Accounts and HBS data ensures consistency between the CPI and the National Accounts data on consumption expenditure of households at the level of the main consumption groups. Such an approach also facilitates more frequent weight updates. For instance, CPI weights can be updated at regular intervals from National Accounts data for higher level aggregates. These updated expenditures are then distributed using the shares obtained from HBS or other sources that may only be updated less frequently.

45. National accountants apply an element of discretion and judgment when making operational decisions relating to the construction of National Accounts. Some of the details of these decisions are not always readily available to users. Moreover, the preliminary results from National Accounts are in general revised several times before they can be considered as final. The most recent available data may not be sufficiently stable. Consequently, compilers of the CPI should consult with their National Accounts counterparts regularly before using their data for weights in order to ensure that they are consistent with the objectives of the CPI.

### ***Retail trade statistics***

46. Statistics on retail sales by region and type of outlet may be available for broad groups of items. One disadvantage is that some of the sales may be to groups outside the reference population, perhaps to the business sector or to the government. The corresponding purchases do not form part of household private consumption. Some sales may also be to non-residents, who may or may not be part of the reference population. Furthermore, for regional sales data, it needs to be kept in mind that sales may include purchases made by people living in other regions.

### ***Population censuses***

47. Population censuses provide statistics on the geographical distribution of the population and households, as well as on the regional differences in household size and composition. Combined with estimates of regional levels of household expenditure, these statistics can be used to estimate regional expenditure weights, especially when such estimates are not available from an HBS with a satisfactory degree of precision. In the absence of any expenditure statistics, population statistics might be used as the basis for regional weights. However, such an approach should be avoided because it assumes that

expenditures per capita or per household are the same in all regions, and ignores the fact that there are usually large differences between the urban and rural populations in what and how much they consume.

### ***Scanner data***

48. Scanner data can be a good source for deriving detailed weights. Such data contain the value of sales for individual products or product categories by outlet or outlet-type for a given reference period. This type of information can then be used to derive weighting structures especially at the lower levels of the index hierarchy. For instance, a detailed product and outlet stratification can be introduced by disaggregating the expenditure for a broader product category that was obtained from the primary data source. However, scanner data may only have limited outlet coverage. Moreover, scanner data may not be fully consistent with the scope of the CPI as in general no distinction can be made between sales made to businesses or households.

### ***Other data sources***

49. If the product categories are sufficiently important, it might be worthwhile to consult additional data sources. A small survey may be conducted with a selection of outlets in order to get a general sense of the breakdown of sales for a specific product category. Existing market intelligence information can also be an option. If there is an association of importers or distributors, other industry groups or marketing agencies and boards, then they will likely have at their disposal at least some general information on the breakdown of sales for specific products.

50. If the objective is to construct a CPI following the domestic concept, the expenditure of non-resident households within the national territory must be included in the weights. In countries where tourism is important, tourism expenditure surveys can be conducted to estimate non-resident household expenditures that can be added to the expenditure made in the country of resident households obtained through a HBS. Foreign visitors will generally have very different expenditure patterns from those of national residents (e.g. they will spend more on hotels and restaurants).

### **Deriving the weights in practice**

51. Once the reference population and the coverage of goods and services have been decided, the weights need to be derived. The weights are calculated as the proportions of the total consumption expenditure of all goods and services included in the index basket for the reference population during the reference period. The reliability of the CPI weights will obviously depend to a large extent on the reliability of the household expenditure data. In practice, the derivation of weights involves a series of steps.

### ***Arrange the data according to the classification and coverage of the CPI***

52. The detailed expenditure items identified in the HBS must be mapped to the CPI expenditure classes. If HBS classes do not match CPI expenditure classes, the HBS results must be transformed to match the CPI categories. This can be done by aggregating or disaggregating the relevant HBS headings over the relevant CPI expenditure classes. Such transformation is achieved much more easily and more reliably if the coding list for expenditure items in the HBS is coordinated with the corresponding list of items used for collecting price observations for the CPI.

53. The HBS may include payments that are outside the scope of the CPI: for example, payments of income taxes or social security contributions, life insurance premiums, remittances, gifts and other transfers, investments, savings and debt repayments are irrelevant and should be ignored because they

are not consumption expenditures. These should be excluded from the total used to calculate the expenditure shares that serve as the basis for estimating the CPI weights.

### ***Correct for over- and under-reporting***

54. The results from the HBS need to be carefully examined and adjusted to take account of under- or over-reporting of consumption expenditures on certain types of products. This is probably the most serious and common problem affecting HBSs. Evidence suggests that the understatement of expenditures can be significant for certain goods and services in the HBS if they suffer from social stigma, such as smoking and drinking of alcohol. It appears that expenditures on durables also tend to be underreported. Other expenditures are not reported because the purchases are small, and therefore easy to forget. Therefore, to the extent possible, results from the HBS should be compared and/or combined with the statistics from other sources when constructing CPI weights, especially when the HBS sample is small.

55. The usual strategy is to use supplementary information from other relevant sources such as tax data, other independent surveys or consumption expenditures from the National Accounts and to apply correction factors. For instance, assume that tax revenue indicates that annual sales for cigarettes may be twice as high as the total annual expenditure estimated from the HBS. The CPI weight is thus obtained by multiplying by 2 the HBS expenditure on cigarettes. It must however be emphasized that these alternative sources are not always perfectly comparable with the coverage of the CPI and may suffer from their own errors. For instance, tax revenue also covers sales made to non-resident households whereas the CPI may only be limited to resident households. At the same time, the fiscal data source does not capture cigarettes sold on the black market.

56. The commodity flow approach can be used to adjust unreliable figures from the HBS. This approach uses domestic production and import/export figures obtained from the National Accounts to adjust consumption expenditures which have been under or over reported in the HBS. In its basic form, the production plus imports less exports of a product is assumed to be equal to its consumption. Estimates of the proportion of domestic consumption which relates to non-household consumption are then used to estimate household consumption. The latter can then be compared with the corresponding estimates from the household budget survey to provide conversion factors to adjust HBS expenditure data for under or over-reporting.

57. The balancing of the Supply and Use Tables in the National Accounts relies on the commodity flow approach. These tables form an integrated framework where supplies of different kinds of goods and services originating from domestic industries and imports are allocated between various intermediate or final uses, including household final consumption. The product balance for any product recognizes that the sum of output at basic prices plus imports plus trade and transport margins plus taxes on products less subsidies on products is equal to the sum of intermediate consumption, final consumption and capital formation, all expressed at purchasers' prices, plus exports.

58. There are some practical limitations to applying the commodity flow approach. Often the balance can only be established for a category of products that may be broader than the product categories used in the CPI classification. Moreover, data are usually only available for the entire country and there is no detailed regional break-down. Finally, household consumption estimated for instance from the HBS may not necessarily be the only weak data source as the estimation of other components of the equation can be likewise unreliable.

59. Some products, such as vehicles or other major durable goods, are purchased only infrequently. If they are purchased, the amount spent on such products can be considerable. As the HBS is a sample

survey, the estimates are bound to be subject to sampling errors, which may be relatively large for such infrequent expenditures. For major durable goods, the HBS estimates should be compared with other sources such as import statistics or administrative files.

60. Expenditures made by households for utility services can often be obtained directly from the providers of such services or from administrative and regulatory data sources. The expenditures derived from these secondary sources do not rely only on a sample of households and can therefore be more accurate than those derived from a HBS.

61. Expenditures for financial services and insurances obtained from a HBS may not always be reliable. It can be difficult for households to properly report explicit charges paid for financial services. Moreover, the derivation of weights for insurance services is special (see paragraph 82). It may be preferable to use regulatory data sources or National Accounts estimates to derive expenditures for financial services and insurances.

62. Even if expenditures obtained from the HBS results are supposed to be accurate, adjustments might still need to be made to take into account any significant changes in expenditure patterns in the period between the time that the survey was conducted and the time that the new weights are introduced. Adjustments will typically be made for products belonging to fast evolving markets and which are significantly losing or gaining importance during this period. It can be that expenditure on some products may not be available from the HBS because the products appeared on the market after the survey had been completed. Additional data sources must then be accessed to estimate expenditures for such new products. Expenditures should also be reviewed if there are known changes that have occurred following administrative decisions, such as changes in taxation, that entered into force only after the HBS was conducted.

#### *Treat expenditures for items that are excluded*

63. The HBS, which in most cases is the main source for deriving the detailed weights, usually includes observations on a much larger variety of goods and services than it is practical to collect prices for. Some products may have a weight which for all practical purposes is negligible. The prices of very minor products may not be worth collecting if they contribute almost nothing to the CPI. In practice, a cut-off threshold can be defined in order to select the products to be included in the CPI basket (see also Chapter 4).

64. Among the consumption expenditures, there are also likely to be a few products for which the prices, or price changes, cannot be directly or satisfactorily measured, such as gambling. It may be difficult if not impossible to compile a reliable price index for such products. A decision must then be made on how to treat the expenditures allocated to such difficult-to-measure products.

65. Even if the product weight is small or if there are measurement problems, the product still remains within the scope of the CPI. Some price change has to be explicitly or implicitly assumed, or imputed. In practice, there are different options:

- The weight of the product is combined with one or several related products. For instance, if the weight for “cheese” is small, it can be combined with the weight for “milk”. This means that the elementary aggregate is now weighted using expenditures of both milk and cheese although only prices for milk will eventually be collected. This is based on the assumption that the price index for cheese would change in the same manner as the observed price index for milk. In practice, the weights for the products could also be kept separated but then price changes must

be explicitly imputed. To continue the previous example, the price index attached to the weight for cheese would simply be identical to the price index for milk.

- The alternative is to simply set the weight for which no representative prices exist equal to zero. This essentially removes the item from the scope of the CPI and is equivalent to assuming that the price of the excluded product would have moved in the same way as the overall CPI for all the products actually included in the index.

66. In the example in Table 3.3, the expenditure for item 3 is too small. It can be decided to add this expenditure to item 1. Alternatively, it may be added to the two items 1 and 2 proportional to the weights of these two items. Finally, it can also be simply removed, which implies that the total expenditure for the three items is now reduced. In general, the last option should be best avoided and expenditures should be allocated to products with similar price behavior. Because of the negligible size of the weight value involved, the consequence on the overall index will in general be negligible whichever method is used.

*Table 3.3 Treatment of products for which no prices are collected*

	Initial expenditures	Allocate to item 1	Allocate to items 1 and 2	Remove expenditure
Item 1	7	8	7.7	7
Item 2	3	3	3.3	3
<b>Item 3</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>0</b>
TOTAL	11	11	11	10

### **Weight reference period**

67. The weight reference period is the time period to which the estimated expenditures relate. For a fixed basket CPI which keeps weights constant over several periods, the weight reference period is typically a twelve-month period, such as a calendar year. A month or quarter is too short a period to use as a weight reference period, since any one month or quarter is likely to be affected by accidental or seasonal influences. This is especially important in countries where composition and size of expenditures can fluctuate significantly throughout the year. In some instances, data for a single year may not be adequate either because of unusual economic conditions or because the sample is not large enough. An average of several years of expenditure data may then be used to calculate the weights.

68. As the CPI is sensitive to the selection of the weight reference period it might be preferred to use a “normal” consumption period, if possible, as the basis for weights and to avoid periods in which there are special factors of a temporary nature at work. To achieve this, it may be necessary to adjust some of the values, to normalize them and to overcome any irregularities in the data. It might be considered to smooth particular erratic observations, for example by taking an average of several years. All available information concerning the nature of consumption in a weight reference period should be taken into consideration.

69. During periods of high inflation, multiple year weights may be calculated by averaging value shares rather than averaging actual value levels. Averaging the value levels will give too much weight to the data for the most recent year. Another option is to update the values for each year to a common period and then to compute a simple arithmetic average of adjusted yearly data.

70. In general, there is a time lag between the weight reference period and the moment that the weights are introduced in the CPI. The overall principle is that the implementation lag should be minimized so that the weights used in the CPI are as up to date as possible. New weights should be

introduced in a timely manner, as soon as the required source data are available. For instance, expenditure weights referring to year  $y-2$  could be introduced with the January index of year  $y$ .

71. As the weight reference period usually proceeds the price reference period, the expenditure weights may or may not be price updated to take account of the relative price changes from the weight reference period to the price reference period. Price updating of weights is discussed in more detail in Chapter 9.

#### **Frequency of weight updates**

72. In most countries, weights are kept fixed for a certain period. The expenditure weights should be updated regularly, at least once every five years, in order to ensure that they do not become unrepresentative and irrelevant. In addition, weights also need to be updated to reflect the introduction of new items and the removal of old ones. Countries which are experiencing significant economic changes and thus more rapid changes in the consumption pattern should update their weights even more frequently, such as annually. In general, increasing the frequency of updating weights minimizes the distorting impact that changes in consumer preferences and item substitutions might have on the index.

73. In the short run, consumers may change consumption patterns in response to shifts in relative prices, mostly between products included in the same class or sub-class. Over longer time periods, consumption patterns are also influenced by factors other than price changes. Most importantly, under rising incomes, changes in the level and distribution of household income will cause a shift in demand for goods and services towards goods and services with higher income elasticities. Demographic factors such as ageing of the population, and technological changes are examples of other factors that affect spending behavior in the longer run. Furthermore, new products will be introduced and existing ones may be modified or become obsolete. A fixed basket will be unresponsive to all these changes.

74. As a result of both relative price changes and long-term effects, the weights may become out of date and less representative of current consumption patterns. As shown in Chapter 15, the bias in a fixed basket index is likely to increase with the age of the weights. At some point, it therefore becomes desirable to use the weights of a more recent period to ensure that the index is weighting appropriately the price changes currently faced by consumers.

75. Whenever the weighting pattern has been updated, the new index using updated weights should be calculated for an overlapping period with the old one so that the two can be linked (see Chapter 9).

76. Frequent updating of weights and chaining can lead to chain drift. For a Laspeyres or a Lowe index, the drift can be upwards if there are systematic fluctuations in consumption and prices. Suppose that the expenditure for petroleum products increases between the old and the new weight reference period because prices went down for these products, causing an increase in the quantities consumed. If, after the new weights are introduced, petroleum prices begin to rise again, the aggregate consumer price index will rise more rapidly than before since these products now have more importance. However, at the level of broad product categories, expenditure data analyzed over longer periods often follow trends. Chaining can have a downward impact if for instance there is a gradual shift from one product to another product whose prices are rising less. A chained CPI compiled with for instance annually updated weights is often found to be lower than the direct index compiled with unchanged weights.

77. The decision when to update the weights depends, for the most part, on the differences observed between the current weighting structure and that for the new weight reference period. Changes in the relative importance of each item can be observed through expenditure survey results. If these statistics are available only at irregular intervals, the frequency of weight revision may necessarily be linked to

the availability of results from the HBS. When the weights are to be fixed for several years, the objective should be to adopt weights that are not likely to change much in the future rather than precisely reflect the activity of a particular period that may be abnormal in some way.

78. Even if weights are updated only every five years, it is desirable to carry out in between a review of the weights in order to ensure that they are sufficiently reliable and representative. The review, which may be confined to weights at the level of sub-indices and their major components, should examine whether or not there are indications that important changes may have taken place in the consumption pattern since the weight reference period.

79. The selection of the level in the index hierarchy at which the structure and weights are fixed for a period is particularly important. The main advantage of setting the level relatively high is that the actual samples of products and their prices below this level can be adjusted and updated as needed. New products can be introduced into the samples, and the weights at the lower level re-established on the basis of more recent information. There is thus a greater opportunity to keep the index representative, through an ongoing review of the sample of representative products.

80. If the level is set relatively low in the index structure, there is less freedom to maintain the representativeness of the index on an ongoing basis, and there will be a greater dependence on the periodic index review and reweighting process. In such circumstances, the arguments for frequent reweighting become stronger.

### **Items requiring special treatment**

#### *Seasonal products*

81. In practice, there are two types of approaches for the treatment of seasonal products:
- a fixed weights approach, which assigns the same weight for the seasonal product in all months, using an imputed price in the out-of-season months. Seasonal products are treated in the same way as other consumption products;
  - a variable weights approach, in which a changing (or moving) weight is attached to the product in various months. In this method, the weights of the seasonal products change monthly according to changes in consumption expenditure during the months of the weight reference period. The principle of a fixed basket – i.e. fixed weights – should, however, be maintained at least at some level of aggregation. One disadvantage of such an approach is that the monthly changes in the index may be more difficult to interpret as they may reflect not only price changes but also quantity changes.

The treatment of seasonal products is further discussed in Chapters 11 and 20.

#### *Insurance*

82. As explained in the section on insurance in Chapter 2, the weights for non-life insurance could be based on either (i) the gross premiums paid, consisting of the payment for the insurance itself, or (ii) on the implicit service charges payable to the insurance enterprise for arranging the insurance. The implicit service charges for administering the insurance and providing the insurance services are estimated by the gross premiums plus the income from investment of the insurance reserves less the amounts payable to policy holders in settlement of claims. This definition can potentially lead to negative weights for instance if there are irregular and unexpected large fluctuations in claims due to natural disasters or large-scale accidents. To reduce the risk of negative weights, an average service charge covering several years could be considered.

83. If the weights are based on the gross premiums, expenditure weights should normally exclude goods and services provided for or reimbursed by the insurance company on the basis of claims. If the weights are based on the implicit service charge, expenditure weights include goods and services that households buy and that are reimbursed by insurance companies, and also goods and services that are paid for and provided by insurance companies on the basis of claims.

84. In general, it seems preferable to base the weights for non-life insurance on the service charges. However, a case can also be made for basing the weights on the gross premiums. This is a difficult area in which there is not yet a consensus. Given the difficulties in separating the services charges from the price of the insurance itself, and because the information is not available each month, the gross premium is followed each month.

#### *Second-hand goods*

85. The prices of used or second-hand goods purchased by households are included in the CPI in the same way as the prices of new goods. However, households also sell used goods, such as cars. If the price of a second-hand good rises, a purchasing household is worse off, but a selling household is better off. From a weighting perspective, sales constitute negative expenditures, which implies that price changes for used goods sold by households implicitly carry a negative weight in the CPI. In effect, purchases and sales of second-hand goods between households, whether directly or indirectly through a dealer, cancel out (except for the dealers' margins, see Chapter 2) and carry no weight in the CPI. However, households also buy from, and sell to, other sectors. For the reference population as a whole, namely the entire set of households covered by the CPI, the weight to be attached to a particular kind of second-hand good is given by households' total expenditures on it less the value of the households' receipts from sales to/from outside the household sector, including the rest of the world. There is no reason why these should cancel out on aggregate. For example, many of the second-hand cars purchased by households may be imported from abroad. The difference between total expenditures and total sales is usually described as households' net expenditures. This is the weight to be attached the second hand good in question.

86. There can be second-hand markets for a whole range of durable and semi-durable goods. Except in the case of used cars, however, it is often practically impossible to estimate the net expenditure because most HBSs do not collect the data that would allow for a comparison between expenditures and receipts from sales of individual kinds of second-hand goods. Usually, only the total amount received from the sale of second-hand goods is collected. This information does, however, give an idea of the volume and significance of these transactions in the national economy.

87. In countries where this volume is small, second-hand goods (except used cars) may be ignored when calculating the weights of the index. In countries where second-hand purchases are important and their prices are believed to change at different rates from those of new goods, separate weights are needed for them. The information could be obtained, at least for some major durables, from HBSs, if the surveys ask about expenditure on second-hand and new goods. Because the amounts spent on purchasing second-hand cars are usually large, they should be included in the CPI basket if the data are available.

88. Most countries include expenditure on second-hand goods in the estimation of CPI weights, but second-hand goods may or may not be covered in the price collection. By excluding them from the price collection, it is implicitly assumed that the prices of new and second-hand goods move in the same way. If they are included, the price determining characteristics of the second-hand good must be kept constant over time so the same good is priced each month. If the goods are different, appropriate quality adjustments must be made.