Draft CPI terminology

The purpose of the following text is to suggest what terms should be used to ensure mutual comprehension in international discussions rather than to provide lexicographically precise definitions.

The conceptual **scope** of a Consumer Price Index is the range of **products** (consumer goods and services) consumed by the **target population** within the geographical area to which the index relates, i.e. all the consumption whose prices should ideally be measured in the index. The actual **coverage** of the index may be narrower than its scope for practical reasons.

Consumption may be defined in terms of **payments**, **transactions** [1] or **use** with prices correspondingly regarded as prices paid, prices agreed or accepted and as prices some of which are **user costs.**

The target population consists of:

households, i.e. groups of resident people who share living expenses,

and may also include

non-profit institutions serving households, which provide non-market services to households and whose main resources are contributed by households,

non-resident households in respect of their consumption purchases within the area,

collective households such as rest homes for the aged.

All of the above within scope can also be called **consumer units**. **Index households** are those resident households that are covered by an index whose scope does not include all such households.

The **product components** of the range of products within scope or covered are categorised hierarchically into **product groups**, each of which may consist of a number of **product subgroups**, known as "sections" in some countries**.** These may be further subdivided. The set of all products is sometimes called the **basket**.

The geographical area within scope or covered may be broken down by **region**, this meaning either a contiguous area of the country or, for example, rural areas or large towns. Sub-areas selected for price collection within a region, such as a shopping area, are called **locations.**

---

[1] "Acquisition" was he term used in the 1989 ILO Manual.

Aggregates and their **sub-indexes** are all defined in terms of product components and, in many cases, also by region and by type of **outlet**, an outlet being anywhere where products are purchased. At the lowest level of aggregation they are **elementary aggregates** and **elementary aggregate indexes,** all of these being defined by product category and many of them by region and by outlet type as well. Elementary aggregate indexes have a weight but contain no internal weights, elementary aggregate indexes being computed using only (sample) price data. The **overall index,** which may be loosely called the **all products index** is a weighted sum of all elementary aggregate indexes.

A subindex for the ith component of consumption is denoted $I^{i}_{\_/\_\_}$ where the part after the slash indicates the current period and the part of the subscript before the slash indicates the **price reference-period.** This is the period or date with whose prices current prices are compared in computing the index. Thus, if this is denoted 0, and the current period is denoted t, the overall index is $\sum_{i} w^{i} I^{i}_{0/t}$ where the w^{i} are the weights. These are shares of the components of consumption in the **weight reference-period,** and may be **price-updated** from the weight reference-period to the price reference-period by multiplying each by its sub-index and then rescaling so that their sum again equals unity. Most Consumer Price Indexes are not true Laspeyres indexes as their weight and price reference-periods do not coincide; they may be called **Laspeyres-type indexes.** A **Cost of Living index**, on the other hand, is the change in household expenditure required to maintain a constant level of utility, a theoretical economics concept.

Rescaling means multiplying or dividing all values of an index series by a constant. If an index series is rescaled by dividing all values by its value in period t, this makes t the **index reference-period** with a value of unity or 100. Linking and Chaining both mean joining together two consecutive time-series by rescaling one of them to make its value for a **link period** equal to the value of the other for that period, thus combining them into a single time-series. **Linking** should be used when individual price series are joined together, while **Chaining** should be used when indexes are joined together.

Prices **collected** and used in compilation relate to one or more **product-types[2],** types of product chosen to represent a class of goods or services for which prices are to be collected,

---

[2] "Representative product" was the term used in the 1989 ILO Manual. "Generic product" would be a better term had it not unfortunately acquired a specialist meaning with regard to pharmaceutical products.

whose **specifications** are laid down by **Head Office.** In each outlet, the particular product selected, either **purposively** or **randomly,** as falling within the specification of a product-type is a **sampled product**[3] whose **description**[4] necessarily contains more detail that specification. Its price, when the meaning is not otherwise clear, is a **sampled price**[5]**.** In the case of cars and some other durable goods, the concepts of **model range, model and variant** are appropriate terms. **Local price collection** is the collection of prices from retail outlets by **collectors** living in the region who transmit the data they obtain to Head Office either directly or via a **Regional Office.** **Centrally collected prices** are those collected directly by Head Office, for example postage rates.

In describing formulas used in computing elementary aggregate price indexes, the division of one price or average by another should be called a **ratio,** not a "relative"[6]. Two sets of sampled prices for one set of sampled products constitute **matched prices**.

When a **replacement product** is selected to supersede a **replaced product** because the latter is no longer available, a **forced replacement** is made. **Optional replacement** is undertaken to increase or preserve the representativity of a product sample, usually when a sample is revised. In such cases there is usually an **overlap** period. Overlap exists when the prices of replaced sampled products and of their replacements are both collected in at least one month. It is convenient to distinguish replacement, *which is done by the statistician*, from **substitution**, which is *what consumers do* in response to changes in price ratios.

The process of scrutinising reported prices is called **editing.** An **estimated price** may be imputed for **missing observations** or to replace recorded **price observations** which are judged to be in error and are rejected. Imputation methods include **carry forward** of the last price observed and, in the case of products which have been on Sale, **reversion to the last regular price**.

The **characteristics** of a sampled product are those attributes which serve to identify it and/or which are **price-relevant**. If none of the differences between the characteristics of two sampled products are judged to be price-relevant, these sampled products are **essentially**

---

[3] The term "Variety", formerly used, does not translate well; in French "Variété" is used to signify "Representative product".

[4] Note that in some countries "specification" or "definition" and "description" are used the other way round.

[5] Sometimes called a "quote".

[6] "Relative" does not translate well.

similar.[7] Essential similarity is sometimes said to allow "direct comparison" between their prices, but this term should be avoided as it also has two other meanings:– (1) "making a quality judgement." and (2) "comparison with price reference-period price", as opposed to comparison with last month's price. When the meaning is (2), the term **fixed base index** or **fixed base price ratio** should be used instead.

When there is a forced replacement a **quality judgement** may be made, meaning either the *process* of estimating the market value of any differences in the characteristics of sampled products, or the resulting *estimate.* A **quality adjustment** is the *process* or the *result* of using such an estimate to add to, subtract from or multiply a price by a coefficient to impute what its market price would be if it had the characteristics of the other sampled product. When the quality judgement is one of essential similarity, a zero adjustment counts as a quality adjustment. When a quality judgement is made the procedure is **judgmental**, otherwise a **mechanistic** procedure is followed instead. Examples of mechanistic procedures are:

— **Link-to-show-no-change**:- from t to t+1 price change is treated as zero; the whole price difference is eliminated as if it reflected a quality difference. (This is a very common method in practice);

— Use of overlap price ratios.

— **Extrapolating** forward the last sampled price of the replaced product to create a fictional overlap with the replacement and treating the resulting overlap price ratio as reflecting the value of the quality difference;

— **Retrapolation** of the price of the replacement product to estimate a price reference-period price for it.

---

[7] The term "comparable" has been used in this sense, but since it can also mean "capable of comparison" it is ambiguous and should be avoided.