# Guide on Multilateral Methods in the Harmonised Index of Consumer Prices

2022 edition


MANUALS AND GUIDELINES

eurostat

# Guide on Multilateral Methods in the Harmonised Index of Consumer Prices

2022 edition

Luxembourg: Publications Office of the European Union, 2022



Theme: Economy and finance
Collection: Manuals and guidelines

PDF ISBN 978-92-76-44354-4 2315-0815 doi:10.2785/873932 KS-GQ-21-020-EN-N

# Foreword

Multilateral methods are a specific type of index compilation method that can be applied to scanner data. The Harmonised Index of Consumer Prices (HICP) implementing regulation ([1]) specifies the use of multilateral methods for the compilation of elementary price indices. Given the current state of play with these methods, there is a need to summarise the main conclusions and to develop concrete guidelines and recommendations on this topic.

A price index measures the aggregate price change in the current period, compared with some base period. There exist many index formulas that take into account the prices and, possibly, the weights of products in the two comparison periods. In multilateral methods, the aggregate price change between two comparison periods is obtained from prices and quantities observed in multiple periods, not only in the two comparison periods. Multilateral methods have been applied for many years for making price comparisons across space (e.g. between different countries or regions), and they have been adapted to make comparisons across time.

Scanner data sets are typically characterised by being 'dynamic'. New products continuously enter the market while obsolete products previously available are removed from the assortments. With standard bilateral price index methods, prices of products in the current period are compared with the corresponding prices in some past base period. However, as they move away from the base period, the overlap of products declines, which makes the calculation of price comparisons more difficult. One way of increasing the overlap of products is to frequently update the base period and chain-link the resulting bilateral price indices. However, it has been found that such an approach can be subject to 'chain drift', especially if products are explicitly weighted. Frequently chained indices tend to systematically drift away, and therefore do not measure a reasonable price change over longer periods.

Multilateral methods have been found to be a solution to the problems encountered with bilateral methods. They take into account all the products that are available in the different periods. They allow each product to be explicitly weighted according to its importance in each period. Finally, they aim to avoid the chain-drift problems encountered with chained bilateral indices. Given these advantages, multilateral methods have been recommended as suitable price index compilation methods for transaction data, despite their additional complexity compared with bilateral methods.

This guide should support countries in understanding and implementing multilateral methods in the context of the HICP. Recommendations are provided on different aspects of implementation.

When applying a multilateral method, the following technical choices have to be made.

- How should the individual product that serves as input to the multilateral method be specified? (See Section 3.)
- Which multilateral index formula should be used? (See Section 4.)
- How should the time window over which the index formula is compiled be specified, and which splicing technique should be used to combine the respective indices? (See Section 5.)
- At what level should a multilateral method be applied and how should the index be integrated into the HICP? (See Section 6.)

The application of multilateral methods to seasonal products is discussed in Section 7. In order to validate and interpret the results, it is possible to calculate the contribution of an individual product to the final index (see Section 8). Multilateral methods can also be used to compile indices for the HICP at constant tax rates (see Section 9). A research agenda for multilateral methods can be found in the annex.

This guide has been drafted with the support of the HICP Task Force on Multilateral Methods. The mandate of the task force was to make proposals for recommendations on the use of multilateral methods and to serve as a forum for exchanging experiences on this topic. It was composed of experts from 12 countries (Belgium, Denmark, Finland, France, Germany, Italy, Luxembourg, the Netherlands, Norway, Poland, Slovenia and Sweden), from the European Central Bank and from Eurostat. The guide

([1]) In particular, see Article 12 of Commission Implementing Regulation (EU) 2020/1148.

was endorsed by the Price Statistics Working Group in its meeting of 11–12 November 2021. Claude Lamboray from Eurostat coordinated the work on the guide and was responsible for the editing of the manuscript.

# Contents

# 1 Data prerequisites

In order to use multilateral methods in the compilation of price indices, it is necessary to meet some data prerequisites. First of all, the national statistical institute (NSI) must have access to transaction data. The coverage and structure of such data sets may be different from one data provider to another. The data should contain the turnover and the number of units sold by item, for instance by Global Trade Item Number (GTIN) or stock-keeping unit (SKU), in one or more outlets during a specific time period ([2]). Secondly, the data should ideally contain some additional information on the item, such as a text string freely describing the item, some retailer-specific product classification code or some other structured product characteristics. Thirdly, as multilateral methods are defined with respect to time windows encompassing many months, data that go back far enough are required for testing and implementing these methods.

The received raw data need to be preprocessed and classified. Since multilateral methods are, in principle, based on all transactions, there is no need to sample products or to exclude products with a low turnover (i.e. apply a low-sales filter ([3])). Each product will be incorporated according to its importance. However, in practice, observations can still be excluded because important information is missing (e.g. turnover) or through application of filters, such as the following.

- **Outlier filter.** An observation is excluded if the price change compared with the previous month (or any earlier period) is implausible, or if the price or the quantity is unusual (too high or too low). This may point to coding errors or other mistakes in the data sets. Thresholds for identifying outliers should be set very carefully so that the number of outliers remains low. It is best to follow up any outliers detected during this phase.
- **Dumping filter.** An observation is excluded if both the price and the quantity sold go down considerably compared with the previous period. This is an indication of clearance sales. The observation could be removed, as such observations may inappropriately influence the resulting indices, depending on which index method is used (see Section 4.4.5). Product sales identified as a dumping in a month should also be excluded in the following months if price and quantity remain similar.

For the classification of data, each observation has to be mapped to the classification structure used in the Harmonised Index of Consumer Prices (HICP). There are different techniques to do this in practice. As a minimum requirement, the data need to be mapped to the European Classification of Individual Consumption according to Purpose (ECOICOP) subclasses, but very often more detailed national or retailer-specific classifications are used. It is important that all the price observations that enter the index compilation are classified. Misclassifications can have an impact on the resulting indices.

([2]) Multilateral methods may also be applied to web-scraped data. The main challenge with web-scraped data is the lack of weights and finding an appropriate proxy for them. For an example of obtaining proxy weights for this type of data see, for instance, Chessa and Griffioen (2019).

([3]) The dynamic-basket method (see Section 2) typically includes a low-sales filter. It is not recommended to apply a low-sales filter with multilateral methods. Removing products with relatively small expenditures in a preprocessing step could bias the results.

# 2 Index compilation methods

In order to aggregate the prices of the individual products and obtain elementary price indices, the statistician can choose from many methods. Any method takes as input the prices and quantities of individual products (see Section 3 for the specification of the individual product). The compilation methods can be distinguished according to (i) the weights and product universe, (ii) the index properties, (iii) the update and linking strategy and (iv) the sampling method. The framework is illustrated in Figure 1.

**Weights and product universe**

In each period, the set of available individual products may differ. Note that this set depends on the definition of the individual product. There are different product universes that can result from this. A key distinction must be made between, on the one hand, approaches that are based on a static product universe and fixed weights and, on the other hand, approaches that are based on a dynamic product universe and/or variable weights.

**Index property**

There are many properties (or tests) that a price index may or may not satisfy. In order to discriminate between the different compilation methods, we focus on transitivity, time reversibility and identity.

- Identity is a property that requires that, if each and every price remains unchanged between two periods, the price index must equal unity. This test is sometimes referred to as the 'strong' identity test. A corresponding weaker version of this test can also be defined. The weak identity test requires that, if all prices, and all quantities, are equal in the two comparison periods, then the price index must equal unity. The weak identity test is less controversial than the strong identity test ([4]).
- Time reversibility is a property that requires an index between periods *a* and *b* to be equal to the inverse of the same index between periods *b* and *a*.
- Transitivity is a property that requires that an index that compares periods *a* and *b* indirectly through period *c* is identical to one that compares periods *a* and *b* directly.

**Update and linking strategy**

With a bilateral index, at some point, the base period must be updated, together with an update of the underlying basket of individual products (resampling). This may happen every month, or less frequently, for example only once per year. There are trade-offs in designing an update strategy. More frequent updating is better at taking into account a dynamic product universe. At the same time, frequent chaining can lead to chain drift.

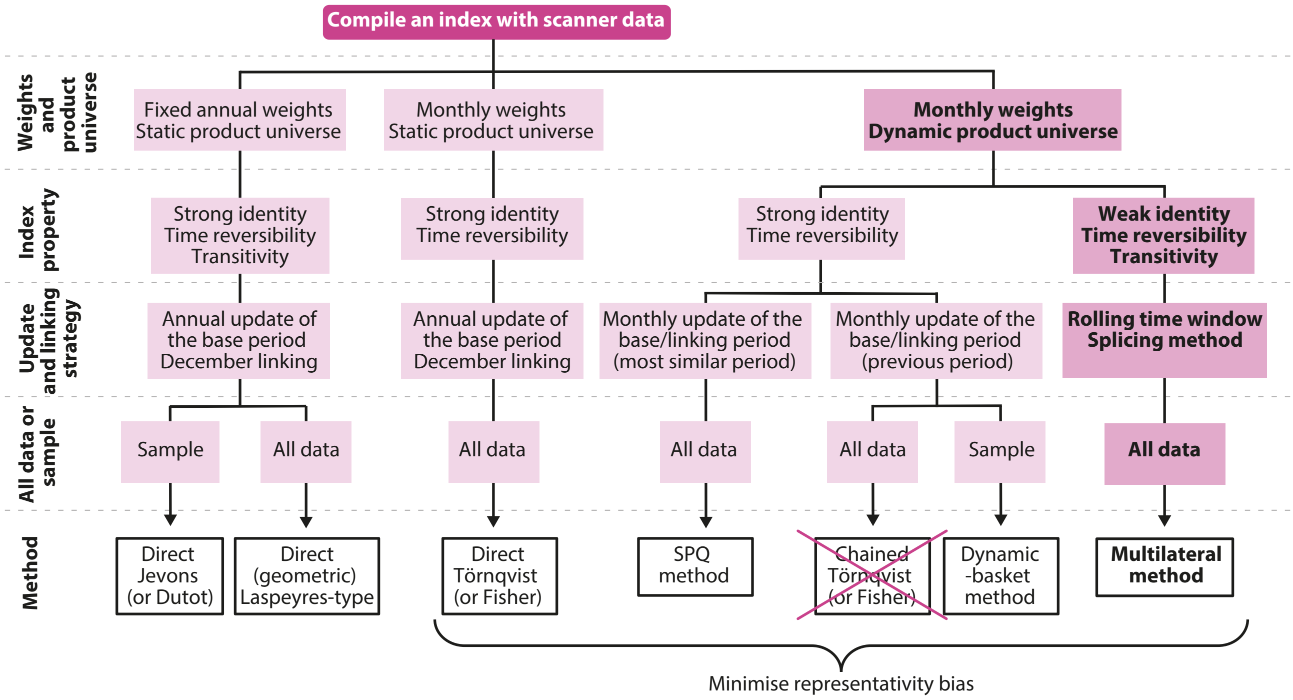
([4]) A multilateral version of the strong identity test could also be defined. If prices are constant over the entire time window consisting of several periods, then the price index must equal unity over that time window. This is in fact a special case of the proportional prices test discussed in Section 4.4.2.

With a multilateral index, the time window over which the index is calculated is updated every month. Most often, rolling time windows are used: each month, the time window is shifted forward by 1 month. Methods are needed to link (splice) the indices calculated over rolling time windows.

### All data or sampling

All the data of the product universe enter the index compilation and no sampling is carried out. Alternatively, a sample of the product universe is selected that represents the product universe and that limits the individual products that eventually enter the index.

**Figure 1: A framework for index compilation methods**




There are different index compilation methods that result from this framework:

- an annually updated direct Jevons or geometric Laspeyres-type index (see box below for more details);
- an annually updated direct Törnqvist index (see box below for more details);
- the linking based on relative price and quantity similarity (SPQ) method (not discussed further in this guide ([5]));
- a monthly chained Törnqvist index, which is, in general, not recommended at the elementary level, as it can be subject to chain drift;
- the dynamic-basket method (see box below for more details);
- multilateral indices, such as the Geary–Khamis (GK), the weighted time–product dummy (WTPD) or the Gini–Eltetö–Köves–Szulc (GEKS), applied over rolling or expanding time windows (these methods are the scope of this guide).

Broadly speaking, there are two approaches. Approaches based on fixed weights and static baskets are similar to the methods typically applied to other price data (e.g. prices collected in the field). They are essentially consistent with the Laspeyres principle and with the methods adopted to aggregate elementary indices. At the same time, these approaches may suffer from what can be called 'representativity bias' ([6]). Approaches based on variable weights, and possibly on dynamic product

([5]) The SPQ method is a method that relies on the relative price and the quantity dissimilarity measure in order to select the best linking period. This method is further discussed in Diewert, 2020.

([6]) The CPI manual (ILO et al., 2020) notes the following in paragraph 12.42:
*The concept of upper-level substitution bias has been derived and discussed in the context of COLI [cost-of-living index] theory, but an*

universes, attempt to reduce any representativity bias. The trade-off between dynamic- and fixed-basket approaches in the context of scanner data is not new. The Eurostat guide on processing scanner data discusses both approaches.

Multilateral methods have the following advantages.

- Multilateral methods take into account all of the products that are available in at least two periods of the time window. This is important because scanner data are typically characterised by being 'dynamic'. Every period, new products enter the market and old products are not available any more.
- With multilateral methods, the monthly weights attached to the products are explicitly used in the index compilation. Therefore, the products are always adequately weighted according to their importance.
- The use of transitive index formulas helps to address the chain-drift problem ([7]).

Multilateral methods are subject to some challenges.

- While multilateral (transitive) index formulas are free from chain drift with respect to a given time window, chain drift can appear when indices compiled over different time windows are spliced together (see Section 5).
- The longer the time window, the more product churn can potentially appear in the data, and the fewer products can be matched ([8]) across all periods of the time window.
- The aggregated price change between two periods may be influenced by data from periods other than the two comparison periods (this is sometimes referred to as loss of 'characteristicity').
- Contrary to other approaches, the multilateral indices described in this document satisfy not the strong identity test, but only the weak identity test.
- Multilateral methods are more complex than bilateral methods. These methods are more difficult to explain to users. It may also be more difficult for compilers to interpret and validate the results of multilateral methods. There are ways to estimate the impact of each individual product on results (see Section 8).

There are alternatives to multilateral methods, but they also have their own limitations.

- In the dynamic-basket method, the set of individual products is updated every month. However, this method is based on only a sample of available products and these products are not explicitly weighted. There could be an increased risk of chain drift in the dynamic-basket method.
- Methods based on the direct Törnqvist (or Fisher) index do use the weight of each individual product in each period. However, the basket of products that enters the calculations is updated only once per year. Moreover, the result can depend very much on the choice of the base period.
- The chained Törnqvist index should, in general, be avoided because it can be subject to significant chain drift. When prices return to their normal levels, quantities purchased are frequently below their normal levels and this can lead to a downward drift.

# Summary of key points

- Approaches based on variable weights and a dynamic product universe aim to minimise representativity bias in the elementary aggregate, whereas approaches based on fixed weights and a static product universe ensure better consistency with the compilation methods applied elsewhere in the HICP.

*equivalent bias may be defined from the perspective of the fixed-basket price index. If the Fisher ideal or other superlative index is judged preferable based on its symmetric treatment of base period and current period expenditure patterns, then the difference between that index and a Laspeyres could be interpreted as a measure of representativity bias. A similar argument could be applied with respect to lower-level substitution bias within elementary index aggregates.*

([7]) In practice, the degree of chain drift depends on both the compilation method and the specific data set. Chain drift can be measured in terms of the violation of the multiperiod identity test. This test requires that, when all prices and quantities in a period *T* revert back to their values observed in the base period 0, then the index should indicate no price change.

([8]) The degree of matching also depends on how the individual product is specified (see Section 3).

- Multilateral methods explicitly use the weights attached to the individual products in each period; are not based on a sample but cover all the data; and are usually less prone to chain drift. With multilateral methods, the set of individual products that enters the calculations is updated every month.

**Some bilateral approaches**

Let us denote by $p_i^t$ and $q_i^t$ the (average) price and the total number of units sold of the individual product $i$ in period $t$. Let $N_t$ be the set of individual products available in period $t$. We suppose that this set is not static, but changes over time. Let $b$ be a period prior to the first period 0.

***Fixed-base Törnqvist index***

Prices in the current period are compared with the prices of the same products in the base period. As weights are known for the two comparison periods, it is possible to apply a symmetrically weighted index formula, such as Törnqvist:

$$I_{Tq}^{0,t} = \prod_{i \in N_0 \cap N_t} \left(\frac{p_i^t}{p_i^0}\right)^{0.5 \times \left(\frac{p_i^0 q_i^0}{\sum_{j \in N_0 \cap N_t} p_j^0 q_j^0} + \frac{p_i^t q_i^t}{\sum_{j \in N_0 \cap N_t} p_j^t q_j^t}\right)}$$

The advantage of this approach is that chain drift is avoided because the current period is compared with some fixed base period, and hence no chaining is involved. Moreover, variable weights are attached to the products that allow the economic importance of each product to be captured in a timely manner.

A limitation of this approach is that it is based only on products that are available in both the current and base periods. It thus misses out on those products that are available in only one of the two comparison periods. As we move away from the base period, it is likely that the overlap between the two comparison periods decreases. That is why, at some point, the base period must be updated. Chaining is required at the moment when the base period is updated and a new fixed-base index is introduced. Some chain drift may still occur at that moment.

In order to maximise the overlap of products in the two comparison periods, an entire year can be used as a base period. An annual base period also has the advantage that it incorporates seasonal products that are available only for certain months of the year. With a 1-month base period, seasonal products that are out of season in that base period would be systematically ignored. In practice, a fixed-base price index is compiled for a sequence of 13 months (December of year $y - 1$ until December of year $y$). Each month is compared with the year $y - 1$. For the following year, the base period is updated so that the year $y$ becomes the new base period and the fixed-base index is compiled for another sequence of 13 months (December of year $y$ until December of year $y + 1$). The two series are chained using the December month of year $y$ as the link period. This ensures that the monthly price developments up to December of year $y$ are consistent with the first index, and the monthly price developments thereafter are consistent with those of the second index. This approach is thus consistent with the standard December linking applied in the HICP at higher aggregation levels.

A fixed-base Törnqvist index has been implemented in Finland (see Vartia et al., 2018, for more details on this method).

***Direct geometric Laspeyres-type and Jevons indices***

Prices in the current period are compared with the prices in the base period for a fixed basket of products denoted by $N_b$. The weights attached to the products refer to some past period (e.g. the previous calendar year) and are kept fixed, for example for 1 year.

$$I_{GL}^{0,t} = \prod_{i \in N_b} \left( \frac{p_i^t}{p_i^0} \right)^{\left( \frac{p_i^b q_i^b}{\sum_{j \in N_b} p_j^b q_j^b} \right)}, b < 0. \; I_L^{0,t} = \frac{\sum_{i \in N_b} p_i^t q_i^b}{\sum_{i \in N_b} p_i^0 q_i^b}$$

If a product becomes permanently missing, a replacement product is selected and a quality adjustment is performed. This ensures that the basket initially selected is kept fixed and remains representative. New products can continuously be incorporated into the index compilation as replacements for disappearing products. However, new products that are not used as replacements for disappearing products will be ignored in this approach. In practice, manual or automatic procedures must be designed that support the selection of the replacements.

Instead of calculating a geometric Laspeyres-type index over all individual products included in $N$, a Jevons index could be calculated over a smaller sample $S$. The Jevons index can be defined as a geometric average of price relatives:

$$I_J^{0,t} = \prod_{i \in S} \left( \frac{p_i^t}{p_i^0} \right)^{\frac{1}{|S|}} = I_J^{0,t-1} \prod_{i \in S} \left( \frac{p_i^t}{p_i^{t-1}} \right)^{\frac{1}{|S|}}$$

The set $S$ is selected for the base period and priced over time. The sample $S$ used is relatively small in order to allow the NSI to manually maintain the sample over time. There are different strategies to select the initial sample. For example, cut-off sampling or sampling proportional to size can be applied in order to take into account how often a product was sold in the past periods (e.g. the previous year). Note that, with probability sampling proportional to size, if the expenditure of an individual product in a base period is used as a size variable, then the sample Jevons price index is an approximately unbiased estimator for the population geometric Laspeyres price index (see Balk, 2005, for details).

The base period and basket must be updated at regular intervals. In line with the higher-level aggregation used in the HICP, the price reference period typically corresponds to the December month of the previous year $y - 1$. The price and weight reference periods are updated once per year and a continuous series is obtained by chain-linking over the December month.

A Jevons index has been implemented, for example, in Switzerland (Müller, 2010) and in Denmark (Mikkelsen, 2012). An approach based on a geometric Laspeyres-type index has been implemented in France (Institut National de la Statistique et des Études Économiques (INSEE), 2020).

***Monthly resampling and chained Jevons (index dynamic-basket method)***

Every month, the set of products that enters the index compilation is resampled. In practice, cut-off sampling is applied, which selects the most-sold products in two consecutive periods. A Jevons price index (unweighted geometric mean of price ratios) is compiled between the two periods taking into account only the selected products. The final price index is obtained by chaining the month-on-month Jevons indices. This approach is referred to as the 'dynamic approach' in the practical guide for processing supermarket scanner data (Eurostat, 2017).

This approach can be seen as a pragmatic compromise: on the one hand, it allows the basket to be updated every month, contrary to the fixed-base approaches; on the other hand, chain drift is avoided (at least to some extent) because the index formula does not explicitly rely on weights. In fact, weights are used only implicitly in the sampling procedure. Unfortunately, only products that pass the cut-off threshold will be taken into account in the compilations. Low-sales products are ignored. In that sense, the observed weights are transformed into some crude weights that are 1 for all selected products and 0 for excluded products. This works best if the expenditure distribution is highly skewed, with a few products accounting for a large share in the total category expenditure. Hence, the main disadvantage of this method is that the observed weights are not explicitly included in the index formula. Although, compared with a chained Törnqvist index, the likelihood of chain drift is significantly reduced, it cannot be fully excluded. Some downward bias can appear when products exit the sample at reduced prices. In order to mitigate this problem, a dumping filter can be implemented that excludes from the calculations those products that exhibit large decreases in both prices and quantities sold.

The dynamic-basket method has been implemented, for example, in Italy (Istituto Nazionale di Statistica (ISTAT), 2020), Spain (National Statistics Institute (Instituto Nacional de Estadística (INE)), 2020) and Slovenia (Republic of Slovenia Statistical Office, 2018).

# 3 Individual product specification

After the data have been received, treated and classified, it is necessary to specify the individual product. This is a critical step, and can have a significant impact on the final results. The individual product is the statistical unit for which a price is tracked over time. As in any other index number approach, the prices of the individual products will be the input of the multilateral index compilation method. First, the conceptual framework of individual products is discussed and then some indications are given on how to specify them in practice.

## 3.1. Conceptual framework

A 'product offer' is a product specified by its characteristics, the timing and place of purchase, and the terms of supply, and for which a price is observed ([9]). In traditional price collection, prices for product offers are observed at one point in time, in a specific outlet, for a specific product. In scanner data, we typically do not observe prices of product offers, but rather average prices (turnover divided by number of units sold) of some item codes ([10]).

When specifying individual products, one needs to consider the following three dimensions: (i) the time dimension, (ii) the outlet dimension and (iii) the product dimension. At each level, an average price (unit value) is calculated, as Figure 2 highlights. The unit value price is calculated by dividing total turnover by the related total quantity. Technically, the order of aggregation (first over time, then over outlets and, finally, over products) does not matter. However, in practice, it makes sense to think about the three dimensions in this order, as there is an increasing likelihood of unit value bias.

**Figure 2: From product offers to homogeneous products**

| | | | | | | |
|---|---|---|---|---|---|---|
| **Product offers** / transacted product | *Aggregation over time* → | Average price for a specific time period (e.g. a month), specific outlet and specific item code | *Aggregation over outlets* → | Average price for a specific time period and specific item code, but over several outlets | *Aggregation over item codes* → | Average price for a specific time period, but over several outlets and item codes = **homogeneous product** |

The individual product can be defined at any level of these successive aggregations. It may be defined in very narrow terms, referring, for example, to an item code in an outlet for a given time period. The main idea of creating broader individual products is to increase the matching over time. The number of individual products that will be taken into account in the index compilation will decrease when more of the data are grouped together. The specification of the individual products depends very much on

([9]) Article 2(4) of Regulation (EU) 2020/1148: ' "product-offer" means a product specified by its characteristics, the timing and place of purchase and the terms of supply, and for which a price is observed'.

([10]) Item codes are further discussed in the practical guide for processing supermarket scanner data (Eurostat, 2017).

the data characteristics (data may be supplied only at a more aggregated level), and on the product category (e.g. homogeneous products may be especially relevant for clothing).

Conceptually, the main principle is that transacted products / product offers can be combined as long as there are no significant quality differences between them. A 'homogeneous product' is defined as a set of product offers among which there are no significant quality differences ([11]). Quality differences must be evaluated with respect to the already mentioned time, outlet and product dimensions. If transactions are combined that are of different quality, unit value bias may occur.

This implies that there is full substitution between the items that belong to the same homogeneous product. This issue is discussed in Silver (2009). Consider the following example of two item codes, A and B (Table 1). Each item code has the same price in both comparison periods. However, there is a shift in the quantities towards the cheaper item, which makes the average price go down from 11 to 10.67. Any standard index number formula would give an answer of unity, (i.e. no overall price change). However, the answer for a homogeneous product would be a decrease of 3 % in the average price. If there is no significant quality difference between item codes A and B, then it is permissible to combine these two item codes into a homogeneous product and calculate an average price. However, if there are quality differences between the two item codes, then these item codes must be treated as separate individual products.

**Table 1:** Example of a homogeneous product and two item codes

| | $p^0$ | $q^0$ | $p^1$ | $q^1$ |
|---|---|---|---|---|
| Item code A | 10 | 6 | 10 | 8 |
| Item code B | 12 | 6 | 12 | 4 |
| Homogeneous product | 11 | | 10.67 | |

There is also a statistical argument. It can be shown (e.g. Diewert and von der Lippe, 2010) that if the prices of the item codes that are grouped together are all equal, and therefore identical to the average price for that homogeneous product, then grouping or not grouping gives essentially the same result. This property is illustrated in the example in Table 2. Item codes 1 and 2 and item codes 3 and 4 always have the same price. Therefore, we group the first two item codes and the last two item codes into two separate homogeneous products. The calculation of a Törnqvist index based on the four item codes or based on the two homogeneous products gives the same result (an increase of 30.12 %). Note that, from an economic point of view, if the products are truly homogeneous and perfect substitutes, and there is a well-functioning market, then prices should be similar.

**Table 2:** Impact on the resulting index of grouping item codes

| | | $p^0$ | $q^0$ | $p^1$ | $q^1$ |
|---|---|---|---|---|---|
| Item code 1 | | 5 | 10 | 6 | 13 |
| Item code 2 | | 5 | 12 | 6 | 15 |
| Item codes 1 and 2 | Homogeneous product 1 | 5 | 22 | 6 | 28 |
| Item code 3 | | 1 | 8 | 2 | 10 |
| Item code 4 | | 1 | 10 | 2 | 8 |
| Item codes 3 and 4 | Homogeneous product 2 | 1 | 18 | 2 | 18 |

([11]) Article 2(5) of Regulation (EU) 2020/1148: ' "homogeneous product" means a set of product-offers among which there are no significant quality differences and for which an average price is calculated'.

# 3.2. Specifying the individual product in practice

Subject to data limitations, the specification of individual products can be made on the following grounds (see Dalèn, 2017, for an initial discussion on this topic).

## 3.2.1. Time dimension

In principle, it is appropriate to calculate a unit value when an item is sold at different prices to different consumers, perhaps at different times within the same month. Conceptually, if all points in time during a certain time period are approximately equivalent to the consumer and there are no systematic price-level differences between weekdays or hours of the day, then the whole time period (month or week) can be considered homogeneous for the purpose of price aggregation. For example, depending on the data supply arrangements and the production and publication calendar, the individual product could cover the first 3 (or sometimes 4) weeks of the month. This is a valid strategy for most situations. In theory, it is important to cover as much as possible of the reference month. Diewert, Fox and de Haan (2016) showed that aggregation over only 1 week of the month can be upward biased, compared with aggregation over the full month.

In some cases, such as accommodation or transport services, the timing is an important quality dimension. Travelling on a Friday evening may be considered a different product (i.e. a different quality) from travelling on a Wednesday afternoon. The price may also depend on the moment of purchase. In such a context, one could argue for treating differences in the time of supply of the service, and possibly differences in the time of booking of the service, as differences in quality. The calculation of a unit value over services with differences in quality would not be correct. Instead, these different services must be treated as different individual products.

In the example in Table 3, the data are provided by week. The 4 weeks are aggregated into a single month. Note that in week 4 the price is significantly lower, for instance because of discounts. If there are no other quality-relevant aspects to consider, it makes sense to simply compile an average price for the whole month.

**Table 3:** Aggregation across time

| Item code | Outlet | Time | Turnover | Quantities | Price |
|---|---|---|---|---|---|
| 1 | 1 | Week 1 | 120 | 56 | 2.14 |
| 1 | 1 | Week 2 | 130 | 63 | 2.06 |
| 1 | 1 | Week 3 | 100 | 43 | 2.33 |
| 1 | 1 | Week 4 | 200 | 120 | 1.67 |
| **Item code** | **Outlet** | **Time** | **Turnover** | **Quantities** | **Price** |
| 1 | 1 | Weeks 1–4 | 120 + 130 + 100 + 200 = 550 | 56 + 63 + 43 + 120 = 282 | 550 / 282 = 1.95 |

## 3.2.2. Outlet dimension

The ideal solution would be to specify the individual product at the level of an outlet and to keep the data as disaggregated as possible. Quality differences between outlets can be associated with different opening hours, different assortments, etc. Even outlets that belong to the same chain may conduct different pricing strategies.

In practice, there may be reasons to combine outlets of the same retail chain or brand. The same product sold by different outlets can be viewed as homogeneous as long as the price of the product is found to be systematically the same across outlets. One reason for aggregating across outlets is that it significantly reduces the number of individual products that will be used in the index compilation. If the number of outlets is large, such a strategy reduces computation time and requires less data storage capacity, although such technical limitations can usually be overcome nowadays.

The impact of aggregating (or not) across outlets is an empirical matter. An extensive discussion of whether aggregation across outlets or not is warranted can be found in Ivancic and Fox (2013). The impact of the outlet dimension can be econometrically assessed. Regression can be used to test if price levels in an outlet are significantly different from the price levels in a base outlet. Another approach is to test the impact of aggregation (or not) across outlets on the final index. It may very well be that there is no large impact, and hence aggregation across outlets is acceptable. Should there be an impact, then it would be best not to aggregate across outlets.

In practice, the treatment of the outlet dimension also depends on the level of detail in the data supplied. For instance, data may be delivered only for an entire retailer, and hence disaggregation by outlet is not possible. The aggregation across outlets is illustrated in Table 4.

**Table 4:** Aggregation across outlets

| Item code | Outlet | Time | Turnover | Quantities | Price |
|---|---|---|---|---|---|
| 1 | 1 | Weeks 1–4 | 550 | 282 | 1.95 |
| 1 | 2 | Weeks 1–4 | 2 203 | 1 123 | 1.96 |
| **Item code** | **Outlet** | **Time** | **Turnover** | **Quantities** | **Price** |
| 1 | 1 and 2 | Month | 550 + 2 203 = 2 753 | 282 + 1 123 = 1 405 | 2 753 / 1 405 = 1.96 |

## 3.2.3. Product dimension

In practice, the GTIN is often found to be the most granular product level available in the data set. Besides GTINs, some retailers may use SKUs. These item codes can be slightly more stable than GTINs. There may also be other product identifiers in the data, depending on the product. In order to deal with relaunches and high turnover in the item codes, the following strategies can be employed: (i) linking, (ii) grouping or (iii) hedonics.

### 3.2.3.1. LINKING

For most supermarket products, it is sufficient to simply use the item code (e.g. GTIN, or preferably SKU) available in the data as a product identifier. This should ideally be combined with some procedure that identifies relaunches. This is best explained by an example (Table 5). Let us suppose that a first item has a price of EUR 20 in period 1 and becomes unavailable from period 2 onwards. A second item of the same quality (a 'relaunch') has a price of EUR 25 and is available only from period 2 onwards. A matched price index would miss the price increase from EUR 20 to EUR 25. In order to avoid this, the two products should be linked.

**Table 5:** Example of a relaunch

| | $p_1$ | $p_2$ | |
|---|---|---|---|
| GTIN A | 20 | – | No matching; therefore, price increase is not captured |
| GTIN B | – | 25 | |
| Linking | 20 | 25 | Price increase is captured |

The price for a homogeneous product is defined as a unit value price (sum of the turnover divided by the sum of the quantities). Relaunches may come with changes in the package size. For example, a chocolate bar previously consisted of 100 g and is now sold in a package of only 80 g. In order to account for a change in package size, one needs to adjust the compilation of the average price. This can be done by calculating an average price by unit of volume. Technically, this corresponds to a quality-adjusted unit value price.

An example of an implementation of the linking strategy in the context of clothing can be found in Bertolotto (2019). The individual products are constructed by linking any new item code to one of the existing item codes ('the closest match'). The closest match is identified by estimating a score that compares the description (text string) of the new item code with the description of possible existing item codes.

### 3.2.3.2. GROUPING

Chessa (2019a) argues that, when constructing homogeneous products, trade-offs must be made between homogeneity and stability over time. If homogeneous products are defined too broadly, there is a risk of a unit value bias. If they are defined too tightly, there is a risk that relaunches are not captured. The 'match-adjusted *R*-squared' (MARS) method (Chessa, 2019a) (see box below for a description of this method) can be used as a tool for finding a compromise between these two objectives.

In general, it is recommended that products with life cycles that frequently end with price reductions are combined into homogeneous products with a longer duration even when small differences in quality exist. This may be especially relevant for clothing.

Sometimes the price level is used to create homogeneous products. Using the price levels to measure quality should be done carefully, especially as price change is the primary measurement target. Identical prices do not necessarily imply identical quality. It is prudent to examine the item codes that are grouped together based on the price level. In general, price level should not be used as a criterion to assess the quality of a product. Price level may be used only as an auxiliary criterion once the data have been preclassified based on other characteristics.

The problem may be more difficult for services than for goods. Homogeneous products for services can be defined in various ways based on the different dimensions or characteristics according to which transaction data for services may be supplied (see Ståhl, 2019, for an example for dental services). Another example can be found in Henn et al. (2019), in which homogeneous products are specified for package holidays and used as an input for a GEKS index. The aggregation across item codes is illustrated in Table 6.

**Table 6: Aggregation across item codes**

| Item code | Outlet | Time | Turnover | Quantities | Price |
|---|---|---|---|---|---|
| 1 | 1 and 2 | Weeks 1–4 | 2 753 | 1 405 | 1.96 |
| 2 | 1 and 2 | Weeks 1–4 | 863 | 378 | 2.28 |
| **Item code** | **Outlet** | **Time** | **Turnover** | **Quantities** | **Price** |
| 1 and 2 | 1 and 2 | Weeks 1–4 | 2 753 + 863 = 3 616 | 1 405 + 378 = 1 783 | 3 616 / 1 783 = 2.03 |

### 3.2.3.3. HEDONICS

Instead of using product characteristics to form homogeneous products, thereby increasing the matching over time, one could work with narrowly defined individual products and directly integrate the product characteristics in hedonically adjusted multilateral indices (see Section 4.4.3). The decision whether or not to group depends very much on the type of product. For example, for clothing it may be more suitable to create homogeneous products, whereas for electronics it is more common to estimate hedonic functions.

In all cases, the practical construction of homogeneous products or the estimation of hedonic functions requires some kind of metadata on product characteristics. These can be obtained by extracting information from the text strings that come with the data. Another strategy is to use additional information on product characteristics provided by market research companies. Scanner data sets could also be complemented by product characteristics scraped from websites.

# Summary of key points

- The individual product must be carefully specified considering aggregation across the time, outlet and product dimensions, as aggregation can significantly impact end results.
- Concerning the time dimension, the advice is to aggregate across a period that covers as much as possible of the reference month.
- The aggregation across outlets is an empirical matter. In principle, the individual product should be specified at the most detailed outlet level available in the data. Aggregation across outlets can be envisaged if data are supplied only at a more aggregated level or if it can be shown that price levels are similar in the outlets (e.g. of the same type and chain).
- Relaunches and product churn can be treated by linking outgoing and incoming item codes. Changes in package size (volume) of relaunches should be adjusted for when linking.
- Another strategy to deal with relaunches and product churn is to create groupings of item codes. When creating homogeneous products, both unit value bias and matching over time must be taken into consideration.
- Hedonic techniques can also be an alternative to the aggregation across item codes (homogeneous products).

### The MARS method

The MARS method is based on the idea that the construction of a homogeneous product is a compromise between two criteria: homogeneity and product match. Metrics are given for these two criteria. This then makes it possible to find the 'optimal' product stratification. The MARS method requires categorical variables that describe the products.

The degree of product match is obtained by summing the quantities (in the current period) of products available in the current and base periods. The degree of homogeneity is defined as the proportion of explained variance in product prices, relative to the total underlying variance in product prices. These two measures are multiplied in order to get a score that combines the objectives. The optimal solution is the one with the highest score.

Consider the following simplified example of T-shirts with data from two periods. There are five item codes in the data. In addition there are two product characteristics: fabric (cotton/organic cotton) and sleeve (short/long). There are different ways to define the homogeneous product. For each solution, a measure of product match and homogeneity is computed, as illustrated below. There are, in this case, five different solutions, with an increasingly broad homogeneous product definition.

| Solution | Item code | Fabric | Sleeve | $p^0$ | $q^0$ | $p^1$ | $q^1$ |
|---|---|---|---|---|---|---|---|
| 1 | 1 | Cotton | Short | 1.90 | 103 | 2.00 | 53 |
| | 2 | Organic cotton | Short | | | 7.00 | 29 |
| | 3 | Cotton | Long | 14.00 | 15 | 15.11 | 18 |
| | 4 | Cotton | Short | 2.00 | 108 | 2.00 | 1 |
| | 5 | Cotton | Short | | | 5.10 | 50 |
| 2 | | Cotton | Short | 1.95 | 211 | 3.49 | 104 |
| | | Cotton | Long | 14.00 | 15 | 15.11 | 18 |
| | | Organic cotton | Short | | | 7.00 | 29 |
| | | Organic cotton | Long | | | | |
| 3 | | | Short | 1.95 | 211 | 4.26 | 133 |
| | | | Long | 14.00 | 15 | 15.11 | 18 |
| 4 | | Cotton | | 2.75 | 226 | 5.20 | 122 |
| | | Organic cotton | | | | 7.00 | 29 |
| 5 | | | | 2.75 | 226 | 5.55 | 151 |

Solution 1 is the most granular solution, whereas, in solution 5, there is only one product ('T-shirts'). The quantity for T-shirts in period 1 is the sum of the quantities of the five item codes: 53 + 29 + 18 + 1 + 50 = 151. The price for T-shirts in period 1 is the average price over the five item codes: ((2.00 × 53) + (7.00 × 29) + (15.11 × 18) + (2.00 × 1) + (5.11 × 50)) / (53 + 29 + 18 + 1 + 50) = 5.55.

With the MARS method, each solution can now be evaluated. It turns out that solution 3 gives the highest MARS score. This suggests that the homogeneous product should be defined by sleeve.

The application of this method to real data may be more complex than what is suggested in this illustrative example. The analysis must be conducted over more than two periods. It may be that the best product definition depends on the time period chosen. Moreover, the outcome could be overly sensitive to the specific data (product sample and prices) used in the analysis. The MARS method should be seen as a decision aid.

**Solution 1: The product is defined by 'item code'**

Total sum of squares: $(53 \times (2.00 - 5.55)^2) + (29 \times (7.00 - 5.55)^2) + (18 \times (15.11 - 5.55)^2) + 1 \times (2.00 - 5.55)^2) + (50 \times (5.10 - 5.55)^2) = 2\,397.1$

Observed sum of squares: $(53 \times (2.00 - 5.55)^2) + (29 \times (7.00 - 5.55)^2) + (18 \times (15.11 - 5.55)^2) + (1 \times (2.00 - 5.55)^2) + (50 \times (5.10 - 5.55)^2) = 2\,397.1$

Product homogeneity: 2 397.1 / 2 397.1 = 100 %

Product match: (53 + 18 + 1) / 151 = 47.7 %

MARS score: 100 % × 47.7 % = 47.7 %

**Solution 2: The product is defined by 'fabric' and 'sleeve'**

Total sum of squares: 2 397.1 (see solution 1)

Observed sum of squares: $(104 \times (3.49 - 5.55)^2) + (18 \times (15.11 - 5.55)^2) + (29 \times (7.00 - 5.55)^2) = 2\,147.6$

Product homogeneity: 2 147.6 / 2 397.1 = 89.6 %

Product match: (104 + 18) / 151 = 80.8 %

MARS score: 89.6 % × 80.8 % = 72.4 %

**Solution 3: The product is defined by 'sleeve'**

Total sum of squares: 2 397.1 (see solution 1)

Observed sum of squares: $(133 \times (4.26 - 5.55)^2) + (18 \times (15.11 - 5.55)^2) = 1\,868.3$

Product homogeneity: 1 868.3 / 2 397.1 = 77.9 %

Product match: (133 + 18) / 151 = 100 %

MARS score: 77.9 % × 100 % = 77.9 %

**Solution 4: The product is defined by 'fabric'**

Total sum of squares: 2 397.1 (see solution 1)

Observed sum of squares: $(122 \times (5.20 - 5.55)^2) + (29 \times (7.00 - 5.55)^2) = 75.5$

Product homogeneity: 75.5 / 2 397.1 = 3.1 %

Product match: 122 / 151 = 80.8 %

MARS score: 3.1 % × 80.8 % = 2.5 %

**Solution 5: The product is defined as 'T-shirts'**

Total sum of squares: 2 397.1 (see solution 1)

Observed sum of squares: $151 \times (5.55 - 5.55)^2 = 0$

Product homogeneity: 0 / 2 397.1 = 0 %

Product match: 151 / 151 = 100 %

MARS score: 0 % × 100 % = 0 %

# 4 Multilateral index formulas

Multilateral methods have been applied for many years for making price comparisons across space (e.g. between different countries or regions), and they have been adapted to make comparisons across time (Ivancic, Diewert and Fox, 2011) ([12]). A multilateral index is compiled over a given time window comprising a sequence of successive months. The index formula takes as input the prices and quantities of the previously defined individual products that are available in the months of the given time window. Below, three multilateral index formulas are described. These multilateral methods are also presented in Chapter 10 of the CPI manual (ILO et al., 2020). Each multilateral index will be illustrated on the following 'example data set' (Table 7), which consists of four individual products and four periods.

**Table 7: Example data set with four individual products**

| Individual product | $p^0$ | $q^0$ | $p^1$ | $q^1$ | $p^2$ | $q^2$ | $p^3$ | $q^3$ |
|---|---|---|---|---|---|---|---|---|
| **1** | 2.97 | 15 | 2.96 | 25 | 2.93 | 32 | 3.03 | 33 |
| **2** | 3.64 | 44 | 3.50 | 79 | 3.36 | 65 | 3.42 | 90 |
| **3** | 6.75 | 49 | 6.71 | 41 | 6.67 | 35 | 6.73 | 53 |
| **4** | 3.37 | 35 | 3.29 | 59 | 3.37 | 30 | 3.37 | 31 |

A larger data set, together with some multilateral indices, can be found online ([13]). The indices included in this file have been calculated using the R package *PriceIndices*. The data set is also part of the package.

The multilateral methods rely on both prices and quantities. Note that multilateral methods can also be implemented without weights (i.e. assuming equal weights). That way, technically, multilateral methods can, for example, be applied to web-scraped data. Applying multilateral methods without weights can yield quite different results. This shows the importance of using the weights in the index calculations so that each individual product impacts the result according to its importance. With web-scraped data, it may be possible to use some proxy weights (based, for example, on the value of the scraped data).

## 4.1. Gini–Eltetö–Köves–Szulc

The GEKS index is based on a bilateral index that is used to compare any two periods belonging to the time window. The initial GEKS method is defined with Fisher price indices. Here, the GEKS method will be presented using a matched Törnqvist index (GEKS–Tq). This multilateral index is also known as CCDI (Caves, Christensen and Diewert, 1982).

([12]) In addition to replacing the spatial dimension with the temporal dimension, there are some additional challenges when applying multilateral methods to scanner data, such as a dynamic product universe or issues with defining and updating the time window.

([13]) https://circabc.europa.eu/ui/group/7b031f10-ac19-4da3-a36f-58708a70133d/library/70222965-2064-47ba-9a42-32ecb6511d07/details

The GEKS–Tq index between the two periods **0** and ***t*** is calculated as follows for a given time window *W*:

$$I^{0,t}_{W(GEKS-Tq)} = \prod_{k \in W} \left( I^{0,k}_{Tq} \times I^{k,t}_{Tq} \right)^{\frac{1}{|W|}}$$

The first step is to calculate bilateral Törnqvist indices. The results of these two-by-two comparisons on the example data set can be summarised in the following four-period × four-period matrix (Table 8).

**Table 8: Period times period matrix underlying the GEKS calculations**

| | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | 1.0000 | 0.9810 | 0.9705 | 0.9820 |
| 1 | 1.0194 | 1.0000 | 0.9877 | 0.9998 |
| 2 | 1.0304 | 1.0124 | 1.0000 | 1.0140 |
| 3 | 1.0184 | 1.0002 | 0.9862 | 1.0000 |

For instance, the Törnqvist index in period 1 compared with period 0 stands at 0.9810. Note that this matrix has a diagonal containing only 1s. Moreover, the matrix is symmetrical: the index in period 0, compared with period 1, is equal to the inverse of the index in period 1, compared with period 0 (1.0194 = 1 / 0.9810).

The GEKS–Tq index can then be derived as follows:

$$I^{0,1}_{[0,3](GEKS-Tq)} = \left( \left( I^{0,0}_{Tq} \times I^{0,1}_{Tq} \right) \left( I^{0,1}_{Tq} \times I^{1,1}_{Tq} \right) \left( I^{0,2}_{Tq} \times I^{2,1}_{Tq} \right) \left( I^{0,3}_{Tq} \times I^{3,1}_{Tq} \right) \right)^{\frac{1}{4}}$$

$$= \left( (1 \times 0.9810)(0.9810 \times 1)(0.9705 \times 1.0124)(0.9820 \times 1.0002) \right)^{\frac{1}{4}} = 0.9817$$

$$I^{0,2}_{[0,3](GEKS-Tq)} = \left( \left( I^{0,0}_{Tq} \times I^{0,2}_{Tq} \right) \left( I^{0,1}_{Tq} \times I^{1,2}_{Tq} \right) \left( I^{0,2}_{Tq} \times I^{2,2}_{Tq} \right) \left( I^{0,3}_{Tq} \times I^{3,2}_{Tq} \right) \right)^{\frac{1}{4}}$$

$$= \left( (1 \times 0.9705)(0.9810 \times 0.9877)(0.9705 \times 1)(0.9820 \times 0.9862) \right)^{\frac{1}{4}} = 0.9696$$

$$I^{0,3}_{[0,3](GEKS-Tq)} = \left( \left( I^{0,0}_{Tq} \times I^{0,3}_{Tq} \right) \left( I^{0,1}_{Tq} \times I^{1,3}_{Tq} \right) \left( I^{0,2}_{Tq} \times I^{2,3}_{Tq} \right) \left( I^{0,3}_{Tq} \times I^{3,3}_{Tq} \right) \right)^{\frac{1}{4}}$$

$$= \left( (1 \times 0.9820)(0.9810 \times 0.9998)(0.9705 \times 1.0140)(0.9820 \times 1) \right)^{\frac{1}{4}} = 0.9822$$

In the GEKS calculations, apart from a Fisher index or a Törnqvist index, any other bilateral index can be used, as long as the index satisfies time reversibility. This property requires that the index between periods ***a*** and ***b*** is equal to the inverse of the same index between periods ***b*** and ***a***. In particular, the GEKS method could be applied with an underlying bilateral Walsh index.

The GEKS method requires that there is at least one product match for any two periods of the time window. If there is no product match, the bilateral Törnqvist index cannot be compiled, which means that the cell in the matrix is empty. If this is the case, it will not be possible to compile GEKS indices that are transitive. If there are empty cells, there are different solutions: reduce the time window, impute the missing prices or use another multilateral method.

# 4.2. Weighted time–product dummy

The WTPD approach consists of running a regression that includes dummy variables for the products and time periods that belong to the time window. Let *W* be a time window. Let $N_t$ be the set of individual products available in period ***t***. Let *N* be the set of products available in any period. The model can be described as follows:

$$\ln p_i^t = \alpha + \sum_{\substack{r \in W \\ r \neq 0}} \delta^t D_i^r + \sum_{\substack{j \in N \\ j \neq 1}} \gamma_j K_j + \varepsilon_i^t \ \forall\, t \in T, \forall\, i \in N_t$$

where $K_j$ is a dummy variable with the value 1 if the observation relates to the individual product ***j*** and **0** otherwise, and $D_i^r$ is a dummy variable with the value 1 if the observation relates to period ***r*** and **0** otherwise. Note that dummies for the first individual product (product 1) and the first time period (period 0) are excluded to identify the model.

The regression is estimated using the weighted least squares (WLS) estimator, which minimises the weighted sum of squared residuals. Each observation is weighted according to its share in a given period ***t***.

$$s_i^t = \frac{p_i^t q_i^t}{\sum_{j \in N_t} p_j^t q_j^t} \forall\, t \in W, \forall\, i \in N_t$$

The final index (14) is then obtained by taking the exponential of the estimated coefficient for the time dummy variables (15):

$$I_{W\,(WTPD)}^{0,t} = \exp\left(\widehat{\delta^t}\right)$$

We illustrate the WTPD index on the example data set. For the regression, the logarithm of the price will be used as the dependent variable. There are three dummy variables for the products ($K_2$, $K_3$ and $K_4$) and three dummy variables for the periods ($D^1$, $D^2$ and $D^3$). Table 9 gives the data used for the regression. Each observation is weighted according to its expenditure share in that period.

(14) A comparison between the WTPD index and a Törnqvist index can be found in Diewert, 2020. Unless there are no divergent trends in log prices and sales shares, it is likely that the WTPD and the Törnqvist will be different.

(15) From an econometric point of view, a bias is introduced when taking the exponential. However, from an index theory point of view, a bias adjustment may not necessarily be as desirable because of its impact on the properties of the index, as argued in de Haan and Krsinich (2018).

**Table 9:** Data underlying the regression of the WTPD

| ln(*p*) | $D^1$ | $D^2$ | $D^3$ | $K_2$ | $K_3$ | $K_4$ | Weight (%) |
|---|---|---|---|---|---|---|---|
| 1.088562 | 0 | 0 | 0 | 0 | 0 | 0 | 6.8 |
| 1.291984 | 0 | 0 | 0 | 1 | 0 | 0 | 24.5 |
| 1.909543 | 0 | 0 | 0 | 0 | 1 | 0 | 50.6 |
| 1.214913 | 0 | 0 | 0 | 0 | 0 | 1 | 18.1 |
| 1.085189 | 1 | 0 | 0 | 0 | 0 | 0 | 9.0 |
| 1.252763 | 1 | 0 | 0 | 1 | 0 | 0 | 33.7 |
| 1.903599 | 1 | 0 | 0 | 0 | 1 | 0 | 33.6 |
| 1.190888 | 1 | 0 | 0 | 0 | 0 | 1 | 23.7 |
| 1.075002 | 0 | 1 | 0 | 0 | 0 | 0 | 14.5 |
| 1.211941 | 0 | 1 | 0 | 1 | 0 | 0 | 33.8 |
| 1.89762 | 0 | 1 | 0 | 0 | 1 | 0 | 36.1 |
| 1.214913 | 0 | 1 | 0 | 0 | 0 | 1 | 15.6 |
| 1.108563 | 0 | 0 | 1 | 0 | 0 | 0 | 11.5 |
| 1.229641 | 0 | 0 | 1 | 1 | 0 | 0 | 35.4 |
| 1.905088 | 0 | 0 | 1 | 0 | 1 | 0 | 41.0 |
| 1.214913 | 0 | 0 | 1 | 0 | 0 | 1 | 12.0 |

The following estimates are obtained from the regression: $\alpha = 1.10754$, $\delta^1 = -0.01743$, $\delta^2 = -0.03019$, $\delta^3 = -0.01703$, $\gamma_2 = 0.15287$, $\gamma_3 = 0.81206$ and $\gamma_4 = 0.11488$. The index can be derived by taking the exponential of the estimate for the time dummies:

$$I^{0,1}_{[0,3]\,(WTPD)} = \exp\left(\delta^1\right) = \exp(-\,0.01743) = 0.9827$$

$$I^{0,2}_{[0,3]\,(WTPD)} = \exp\left(\delta^2\right) = \exp(-\,0.03019) = 0.9703$$

$$I^{0,3}_{[0,3]\,(WTPD)} = \exp\left(\delta^3\right) = \exp(-\,0.01703) = 0.9831$$

# 4.3. Geary–Khamis

The GK index is an example of a quality-adjusted value index ([16]). It is obtained by solving the following system of equations:

$$I^{0,t}_{W\,(GK)} = \frac{\sum_{i \in N_t} p_i^t q_i^t \,/\, \sum_{i \in N_0} p_i^0 q_i^0}{\sum_{i \in N_t} v_i q_i^t \,/\, \sum_{i \in N_0} v_i q_i^0}$$

$$v_i = \sum_{z \in W} \frac{q_i^z}{\sum_{s \in W} q_i^s} \frac{p_i^z}{I^{0,z}_{W(GK)}}$$

([16]) See also Chessa (2016) for a practical application of this method in the Dutch CPI. The family of quality-adjusted unit value methods (QU methods) encompasses the GK method explained in this document.

We illustrate the GK index using the example data set. The following adjustment factors are used for the four products: $v_1 = 3.029974$, $v_2 = 3.523584$, $v_3 = 6.820025$ and $v_4 = 3.392614$. This then leads to the GK indices presented in Table 10.

**Table 10:** Calculation of the GK Index

| Period | Value | Volume | Price index |
|---|---|---|---|
| 1 | (2.97 × 15) + (3.64 × 44) + (6.75 × 49) + (3.37 × 35) = **653.41** | (15 × 3.029974) + (44 × 3.523584) + (49 × 6.820025) + (35 × 3.392614) = **653.41000** | (653.41 / 653.41) / (653.41 / 653.41) = **1.0000** |
| 2 | (2.96 × 25) + (3.50 × 79) + (6.71 × 41) + (3.29 × 59) = **819.72** | (25 × 3.029974) + (79 × 3.523584) + (41 × 6.820025) + (59 × 3.392614) = **833.8977124** | (819.72 / 833.8977124) / (653.41 / 653.41) = **0.9830** |
| 3 | (2.93 × 32) + (3.36 × 65) + (6.67 × 35) + (3.37 × 30) = **646.71** | (32 × 3.029974) + (65 × 3.523584) + (35 × 6.820025) + (30 × 3.392614) = **666.4713966** | (646.71 / 666.4713966) / (653.41 / 653.41) = **0.9703** |
| 4 | (3.03 × 33) + (3.42 × 90) + (6.73 × 53) + (3.37 × 31) = **868.95** | (33 × 3.029974) + (90 × 3.523584) + (53 × 6.820025) + (31 × 3.392614) = **883.7440252** | (868.95 / 883.7440252) / (653.41 / 653.41) = **0.9833** |

For each individual product, its quantity-based share can be calculated as shown in Table 11.

**Table 11:** Calculation of the quantity shares

| Individual product | All periods | Period 1 | Period 2 | Period 3 | Period 4 |
|---|---|---|---|---|---|
| 1 | 15 + 25 + 32 + 33 = 105 | 15 / 105 = 14.3 % | 25 / 105 = 23.8 % | 32 / 105 = 30.5 % | 33 / 105 = 31.4 % |
| 2 | 44 + 79 + 65 + 90 = 278 | 44 / 278 = 15.8 % | 79 / 278 = 28.4 % | 65 / 278 = 23.4 % | 90 / 278 = 32.4 % |
| 3 | 49 + 41 + 35 + 53 = 178 | 49 / 178 = 27.5 % | 41 / 178 = 23.0 % | 35 / 178 = 19.7 % | 53 / 178 = 29.8 % |
| 4 | 35 + 59 + 30 + 31 = 155 | 35 / 155 = 22.6 % | 59 / 155 = 38.1 % | 30 / 155 = 19.4 % | 31 / 155 = 20.0 % |

The adjustment factors can now be derived from the price index obtained in the previous step. The price in each period is deflated by the price index. An average price over the whole time window is obtained using the quantity shares.

$$v_1 = \left(14.3\,\% \times \frac{2.97}{1}\right) + \left(23.8\,\% \times \frac{2.96}{0.9830}\right) + \left(30.5\,\% \times \frac{2.93}{0.9703}\right) + \left(31.4\,\% \times \frac{3.03}{0.98273}\right) = 3.029974$$

$$v_2 = \left(15.8\,\% \times \frac{3.64}{1}\right) + \left(28.4\,\% \times \frac{3.50}{0.9830}\right) + \left(23.4\,\% \times \frac{3.36}{0.9703}\right) + \left(32.4\,\% \times \frac{3.42}{0.98273}\right) = 3.523584$$

$$v_3 = \left(27.5\,\% \times \frac{6.75}{1}\right) + \left(23.0\,\% \times \frac{6.71}{0.9830}\right) + \left(19.7\,\% \times \frac{6.67}{0.9703}\right) + \left(29.8\,\% \times \frac{6.73}{0.98273}\right) = 6.820025$$

$$v_4 = \left(22.6\,\% \times \frac{3.37}{1}\right) + \left(38.1\,\% \times \frac{3.29}{0.9830}\right) + \left(19.4\,\% \times \frac{3.37}{0.9703}\right) + \left(20.0\,\% \times \frac{3.37}{0.98273}\right) = 3.392614$$

The adjustment factors are identical to those initially used. This shows that we have found a solution for the system of equations.

In practice, one may solve the system of equations by iteration. One starts with a given set of adjustment factors (for instance $v_i = 1$ for all products). This is then used to compile a price index, as outlined above. Given this price index, new adjustment factors can be compiled. The iteration stops when the values for the adjustment factors and the index do not change any more. In order to optimise the calculations, one should choose the initial values wisely.

# 4.4. Analysis of the index formulas

In order to select an index formula, there are different criteria that can be considered. Each method has its advantages and its limitations. Trade-offs must be made, depending on which criteria are considered to be more or less important. Some criteria are summarised in Table 12 ([17]). These criteria are further examined in the next subsections.

In addition to these methodological criteria, practical considerations can also play a role. For example, it is usually more time-consuming to run a program that calculates a GK index or a WTPD index than a GEKS index. This can become an issue for very large data sets.

**Table 12: Comparison of the multilateral index formulas**

| | GEKS | WTPD | GK |
|---|---|---|---|
| **Main principle** | It is closely related to an underlying bilateral index. It is the method that is most consistent with the economic approach to price indices | It assumes that prices are generated according to a model. This approach is grounded in the stochastic approach to price indices | It is an additive method that is closest to the calculations of unit values used at the level of homogeneous products |
| **Tests to distinguish between methods** | | | |
| Homogeneity in quantities | Satisfies | Satisfies | Fails |
| Basket test | Fails | Fails | Satisfies |
| Responsiveness | Satisfies | Fails | Fails |
| **Flexibility** | Possible to impute missing prices using a hedonic function | Possible to include product characteristics in the regression | Possible to estimate a quality-adjusted unit value index using a hedonic function |
| **Bilateral counterpart** | Weighted geometric average of price ratios; weights are arithmetic average of shares (Törnqvist) | Weighted geometric average of price ratios; weights are harmonic average of shares | Basket index; weights are harmonic average of the quantities (bilateral GK) |
| **Sensitivity/robustness** | | | |
| To the choice of splicing methods | Less sensitive | More sensitive | More sensitive |
| To clearance prices | More sensitive | Less sensitive | Less sensitive |

## 4.4.1. Main principles

### 4.4.1.1. GINI–ELTETÖ–KÖVES–SZULC

The GEKS method is closely related to its underlying bilateral Törnqvist index (or any other bilateral index used as a building block of the GEKS). The Törnqvist index is not transitive. It can be shown that the GEKS index is the transitive index that is closest to its underlying bilateral index. Formally, the GEKS index is a solution to the following optimisation problem (Caves, Christensen and Diewert, 1982):

([17]) Another strategy to be further explored (see Białek and Bobel, 2019) consists of generating artificial data sets from a known probability price distribution. Knowing the theoretical expected value of the price, we can choose the multilateral price index that is closest to the theoretical value.

$$\min_{I^1,\dots,I^T} \sum_{i,j \in [1,\dots,T]} \left( \ln(I_{Tq}^{i,j}) - \ln\left(\frac{I^j}{I^i}\right) \right)^2$$

The GEKS is also related to the economic approach to index number theory, which assumes that the quantities are functions of the prices. Households are assumed to optimise their purchases by maximising their utility under a budget constraint, or by minimising the cost of purchases for a given utility level. Let us assume that preferences can be described by a constant elasticity of substitution (CES) utility function. The parameter $\sigma$ is the elasticity of substitution. A parameter of $\sigma = 0$ corresponds to a Leontief utility function that exhibits no substitutability between products. A parameter of $\sigma = 1$ corresponds to a Cobb–Douglas function. A parameter of $\sigma = +\infty$ approximates a linear utility function that assumes full substitutability between the products. In a CES context, the elasticity of substitution is assumed to be constant between any two products. Diewert and Fox (2020) generated artificial data sets assuming different values for $\sigma$. It turned out that, for elasticities of substitution in the usual range of 1–4, the GEKS–Tq approximates CES preferences reasonably well.

There are some limitations of the economic approach in the context of scanner data. It can be argued that real scanner data are not fully consistent with the economic approach. For example, the economic approach does not take into account stock-keeping behaviour. The theory assumes that the periods of consumption and acquisition are the same. This assumption may not always hold. For instance, when a non-perishable product is on sale, consumers may purchase more of that product, but consume it only at a later point in time. Moreover, the dynamic universe is a key feature of scanner data. According to the economic approach, reservation prices must be estimated for new and disappearing products, which can be difficult in practice ([18]). However, for data that are consistent with the economic approach, the GEKS index works best.

#### 4.4.1.2. WEIGHTED TIME–PRODUCT DUMMY

The WTPD method assumes that prices can be estimated according to a certain model. The method can be easily understood and implemented by compilers and explained to users. Usually the model is not fully perfect, and hence there will be non-zero residuals (i.e. the difference between the observed price and the predicted price). A low value for the mean squared error of the regression can be an indication that the model fits the underlying prices data well. In fact, this method should be confined to situations in which the individual products are closely related, so that their prices can be reasonably well approximated by the model. This method should not be used at higher levels of aggregation.

The model can easily cope with situations in which individual products are not present in all periods of the time window.

When estimating the regression, it is important that the price observations are weighted and to apply WLS instead of ordinary least squares. The weighting ensures that, from a price index perspective, the individual products are adequately weighted.

It can be shown that the WTPD index can be described as a weighted geometric mean of the quality-adjusted prices for products that are available in period $t$:

$$I_{W\,(WTPD)}^{0,t} = \prod_{i \in N_t} \left( \frac{p_i^t}{\exp\left(\hat{\alpha} + \widehat{\gamma}_i\right)} \right)^{s_i^t} \Big/ \prod_{i \in N_0} \left( \frac{p_i^0}{\exp\left(\hat{\alpha} + \widehat{\gamma}_i\right)} \right)^{s_i^0}$$

([18]) Reservation prices are usually not part of the practice of official statistics. They can be complex to estimate and the outcome depends very much on underlying assumptions.

#### 4.4.1.3. GEARY–KHAMIS

The GK index is an additive method. Additivity means that a value index deflated by a price index leads to a volume index that has the form of a basket index whereby the quantities in the two periods are compared using constant prices.

$$\frac{\textit{Value index}}{\textit{Price index}} = \frac{\sum_i p_i^* q_i^t}{\sum_i p_i^* q_i^0}$$

The GK index is an additive index. This is an interesting feature because it gives a simple interpretation to the implied volume index. However, it is not clear if the constant prices obtained from the GK equations are the correct ones. Note that the WTPD is an approximately additive method (Diewert and Fox, 2020).

The GK index can also be seen as a special case of a quality-adjusted unit value index, in which the quality adjustment factors correspond to the coefficients determined by the GK equations. These coefficients correspond to an average deflated price and are implicitly determined by the data. In that sense, this method is consistent with the unit value calculations performed at the level of homogeneous products.

### 4.4.2. Test approach

According to the test approach, desirable tests are defined that a multilateral index may, or may not, satisfy. The test approach does not give an indication of how severe a failure is. These tests are defined for an index formula compiled over a given time window (no splicing is involved at this stage). The Australian Bureau of Statistics has examined various tests for multilateral comparison (Australian Bureau of Statistics, 2016). This work draws heavily on initial work done by Diewert (1999) and Balk (2001). Moreover, the axioamtic approach is discssed further in Diewert, 2020.

We use the following notations to describe the tests. In each period $t$, a vector of prices $p^t = \left(p_1^t,\ldots, p_n^t\right)$ and quantities $q^t = \left(q_1^t,\ldots, q_n^t\right)$ are provided. Let $I^{0,t}\left(p^0,\ldots, p^T, q^0,\ldots, q^T\right) = I^{0,t}(P,Q)$ be a multilateral price index that is a function of the prices and quantities vectors observed in a time window that spans from periods $0$ to $T$.

***Test 1: Transitivity***

The multilateral indices all satisfy transitivity. This means that the results are independent of the choice of base period. This is a key property that avoids the chain-drift problem. In the HICP implementing act, it is explicitly mentioned that index formulas are allowed if they satisfy transitivity. Therefore, the multilateral indices described here are all eligible with respect to the legal framework of the HICP.

$$I^{0,t2}(P,Q) = I^{0,t1}(P,Q) I^{t1,t2}(P,Q)$$

***Test 2: Identity***

This test requires that the index equals identity if all prices revert back to their initial levels. Let $p_i^t = p_i^0$ for all $i = 1, \ldots, n$. Then:

$$I^{0,t}(P,Q) = 1$$

***Test 3: Multiperiod identity test***

This test requires that, if the last period is identical to the first period, the chained index will revert back to unity at the end. If $p_i^T = p_i^0$ and $q_i^T = q_i^0$ for all $i = 1, \ldots, n$, then:

$$I^{0,1}(P,Q) \times I^{1,2}(P,Q) \times \ldots \times I^{T-1,T}(P,Q) = 1$$

***Test 4: Continuity, positivity and normalisation***

$I^{0,t}(P,Q)$ is a positive and continuous function of the prices and quantity data and $I^{0,0}(P,Q)$ is equal to 1.

***Test 5: Proportional prices test***

If prices in all periods are proportional, the price index depends only on these proportions. Suppose that there exists $\alpha^t \ (t = 0,\ldots,T)$ such that $p_i^t = \alpha^t \ p_i^0$ for all $i = 1, \ldots, n$, $t = 0, \ldots, T$. Then:

$$I^{0,t}(P,Q) = \alpha^t \ \forall \ t = 0,\ldots,T$$

***Test 6: Homogeneity in quantities***

Rescaling the quantities in a period does not alter the price index. For any period $k$, there exists $\gamma$ such that $\hat{q}_i^k = \gamma q_i^k$ for all $i = 1, \ldots, n$. Then:

$$I^{0,t}\left(p^0,\ldots,p^k,\ldots,p^T,q^0,\ldots,q^k,\ldots,q^T\right) = I^{0,t}\left(p^0,\ldots,p^k,\ldots,p^T,q^0,\ldots,\hat{q}^k,\ldots,q^T\right) \forall \ t = 0,\ldots,T$$

***Test 7: Homogeneity in prices***

Rescaling the prices in a period changes the price index by the same proportion. For any period $k \neq 0$, there exists $\gamma$ such that $\hat{p}_i^k = \gamma p_i^k$ for all $i = 1, \ldots, n$. Then:

$$\gamma I^{0,k}\left(p^0,\ldots,p^k,\ldots,p^T,q^0,\ldots,q^k,\ldots,q^T\right) = I^{0,k}\left(p^0,\ldots,\hat{p}^k,\ldots,p^T,q^0,\ldots,q^k,\ldots,q^T\right)$$

***Test 8: Commensurability***

Changing the units in which prices and quantities are expressed does not change the price index. Let $\delta_i > 0 \ (i = 1,\ldots,n)$. Let $\hat{p}_i^t = \delta_i p_i^t$ and $\hat{q}_i^t = \frac{q_i^t}{\delta_i}$ $(i = 1, \ldots, n$ and $t = 0, \ldots, T)$. Then:

$$I^{0,t}\left(p^0,\ldots,p^T,q^0,\ldots,q^T\right) = I^{0,t}\left(\hat{p}^0,\ldots,\hat{p}^T,\hat{q}^0,\ldots,\hat{q}^T\right) \forall \ t = 0,\ldots,T$$

***Test 9: Symmetry in the treatment of time periods***

Reordering the time periods does not change the price change between the periods.

***Test 10: Symmetry in the treatment of products***

Reordering the products does not change the price index.

***Test 11: The basket test for prices***

If the quantities are the same in the base period and in the current period, the index in the current period corresponds to a basket index. If $q_i^0 = q_i^k = q_i$ for all $i = 1, \ldots, n$, then:

$$I^{0,k}(P,Q) = \frac{\sum_i p_i^k q_i}{\sum_i p_i^0 q_i}$$

***Test 12: Responsiveness test to imputed prices***

Let us consider a dynamic universe. If there are missing products in one or more periods, then there will be imputed prices for these missing products, while the quantities of such products will be set to zero. According to the responsiveness test, there must exist a different imputation of the missing prices that would give a different price index.

Table 13 summarises the performance of the three multilateral price indices with respect to these tests. This is only a selection of tests; other tests can be defined in order to clarify which properties are satisfied by which multilateral index formulas.

**Table 13: Test performance of the multilateral index formulas**

| | | GEKS–Tq | WTPD | GK |
|---|---|---|---|---|
| Test 1 | Transitivity | Yes | Yes | Yes |
| Test 2 | Identity | No | No | No |
| Test 3 | Multiperiod identity test | Yes | Yes | Yes |
| Test 4 | Continuity, positivity and normalisation | Yes | Yes | Yes |
| Test 5 | Proportional prices test | Yes | Yes | Yes |
| Test 6 | Homogeneity in quantities | Yes | Yes | No |
| Test 7 | Homogeneity in prices | Yes | Yes | Yes |
| Test 8 | Commensurability | Yes | Yes | Yes |
| Test 9 | Symmetry in the treatment of time periods | Yes | Yes | Yes |
| Test 10 | Symmetry in the treatment of products | Yes | Yes | Yes |
| Test 11 | The basket test for prices | No | No | Yes |
| Test 12 | Responsiveness test to imputed prices | Yes | No | No |

Note that all three multilateral indices satisfy tests 1, 3, 4, 5, 7, 8, 9 and 10 and that none of them satisfies test 2. The failure of test 2 can be seen as a limitation of the multilateral methods compared with bilateral methods, which typically satisfy the identity test. The SPQ method (see Section 2) is, however, an example of a multilateral method that passes the identity test.

A main feature of the GK is that it does not satisfy test 6. In fact, the GK depends on the quantities sold in each period. The months during which more products are sold have a larger impact on the result. This is not the case for the GEKS–Tq or the WTPD, which treat each month in the same way. At the same time, the GK is the only index that passes test 11.

Finally, a feature of the GEKS–Tq is that it can incorporate the imputed prices of missing products (test 12). Note that, in the standard application of the GEKS–Tq, prices are not imputed. Only matched products between any two periods of the time window are taken into account. This test says only that estimated prices can be taken into account even if the product has not been sold in that period. The estimation of prices can be done using, for example, a hedonic regression (see Section 4.4.3).

The test performance for the GEKS depends on the underlying bilateral index. The analysis presented here is based on a GEKS–Tq. The performance of the GEKS may be different with another bilateral index. For example, the GEKS–Walsh would not pass test 12.

### 4.4.3. Hedonic indices and multilateral methods

The GEKS–Tq, WTPD and GK indices take as inputs only the prices and quantities of the individual products. It is possible to extend these methods and explicitly incorporate characteristics of the individual products. From this perspective, all three methods are sufficiently flexible to include product characteristics. Let $p_i^t$ be the price, $q_i^t$ the quantity and $x_{i,k}$ the characteristics ($k = 1, \ldots, K$) of an individual product $i$ in period $t$.

### 4.4.3.1. GINI–ELTETÖ–KÖVES–SZULC

The GEKS–Tq method could be applied by calculating bilateral imputation Törnqvist indices instead of standard matched Törnqvist indices (de Haan and Krsinich, 2014; de Haan and Daalmans, 2019). For any two periods **0** and ***t***, we denote the matched products by $M_{0t}$ (products available in both periods), the new products by $N_{0t}$ (products available only in ***t***) and the disappearing products by $D_{0t}$ (products available only in **0**). The imputation Törnqvist index is then defined as follows:

$$I^{0,t} = \prod_{i \in M_{0t}} \left( \frac{p_i^t}{p_i^0} \right)^{0.5 \times \left( \frac{p_i^0 q_i^0}{\sum_{j \in M_{0t} \cup D_{0t}} p_j^0 q_j^0} + \frac{p_i^t q_i^t}{\sum_{j \in M_{0t} \cup N_{0t}} p_j^t q_j^t} \right)} \prod_{i \in D_{0t}} \left( \frac{\widehat{p_i^t}}{p_i^0} \right)^{\left( \frac{p_i^0 q_i^0}{\sum_{j \in M_{0t} \cup D_{0t}} p_j^0 q_j^0} \right)} \prod_{i \in N_{0t}} \left( \frac{p_i^t}{\widehat{p_i^0}} \right)^{\left( \frac{p_i^t q_i^t}{\sum_{j \in M_{0t} \cup N_{0t}} p_j^t q_j^t} \right)}$$

In an imputation Törnqvist index, the prices of the missing products can be estimated using a hedonic model. In practice, the following regression could be estimated for data pertaining to each period ***t***:

$$\ln\left(p_i^t\right) = \alpha^t + \sum_{k=1}^{K} \beta_k^t x_{i,k} + \varepsilon_i^t \;\; \forall \; i \in N_t$$

There are variants of this approach. Instead of applying single imputation, one could also apply double imputation and impute both observed and imputed prices of the sets $D_{0t}$ and $N_{0t}$. Another variant is to estimate a time dummy hedonic method between any two periods of the time window (see the imputation Törnqvist rolling yearGEKS method discussed in de Haan and Krsinich, 2014).

### 4.4.3.2. WEIGHTED TIME–PRODUCT DUMMY

The WTPD is based on a regression that includes dummy variables for each individual product. Instead, product characteristics could be used as explanatory variables of the model. In practice, the following hedonic model can be estimated over the pooled data pertaining to the full time window, using weights:

$$\ln p_i^t = \alpha + \sum_{\substack{r \in T \\ r \neq 0}} \delta^t D_i^r + \sum_{k=1}^{K} \beta_k x_{i,k} + \varepsilon_i^t \;\; \forall \; t \in T, \forall \; i \in N_t$$

The resulting index is then obtained by taking the exponential of the estimate of the time dummy variables. This method is sometimes referred to as the time dummy hedonic index. This method is regularly applied in the context of residential property price indices (see the International Monetary Fund's *Residential Property Price Index Practical Compilation Guide*, 2019).

### 4.4.3.3. GEARY–KHAMIS

In the GK, adjustment coefficients are estimated implicitly by the method. One could also derive explicit adjustment coefficients based on a hedonic regression. This is sometimes referred to as a quality-adjusted unit value index. In practice, the quality adjustment factors $v_i$ for a product ***i*** are obtained by comparing the estimated prices of that product (according to the time dummy hedonic model; see previous paragraph) with the estimated prices of a reference product (product 1).

$$v_i = \frac{\widehat{p_i^t}}{\widehat{p_1^t}} = \exp\left(\sum_{k=1}^{K} \widehat{\beta_k}(x_{i,k} - x_{1,k})\right)$$

This explicitly quality-adjusted unit value index can be quite different from the GK index. It can be shown (de Haan and Krsinich, 2018) that such an index is, however, very similar to a time dummy hedonic index.

### 4.4.4. Bilateral counterparts

A multilateral method can be applied to a time window that consists of only two periods. In that case, a multilateral index formula collapses to a bilateral index formula and the multilateral index formula can be seen as an extension of this bilateral index formula.

| Multilateral index formula | Bilateral index formula if time window consists of two periods only |
|---|---|
| GEKS–Tq | $I_{Tq}^{0,1} = \prod_{i \in N_0 \cap N_1} \left(\frac{p_i^1}{p_i^0}\right)^{0.5 \times \left(\frac{p_i^0 q_i^0}{\sum_{j \in N_0 \cap N_1} p_j^0 q_j^0} + \frac{p_i^1 q_i^1}{\sum_{j \in N_0 \cap N_1} p_j^1 q_j^1}\right)}$ |
| GK | $I_{BGK}^{0,1} = \frac{\sum_{i \in N_0 \cap N_1} \left(0.5 \times \left(\frac{1}{q_i^0} + \frac{1}{q_i^1}\right)\right)^{-1} p_i^1}{\sum_{i \in N_0 \cap N_1} \left(0.5 \times \left(\frac{1}{q_i^0} + \frac{1}{q_i^1}\right)\right)^{-1} p_i^0}$ |
| WTPD | $I_{GH}^{0,1} = \prod_{i \in N_0 \cap N_1} \left(\frac{p_i^1}{p_i^0}\right)^{\frac{\left(0.5 \times \left(\frac{1}{s_i^0} + \frac{1}{s_i^1}\right)\right)^{-1}}{\sum_{j \in N_0 \cap N_1} \left(0.5 \times \left(\frac{1}{s_j^0} + \frac{1}{s_j^1}\right)\right)^{-1}}}$<br>where $s_i^0$ and $s_i^1$ are the shares of the product ***i*** in the two comparison periods. These shares are calculated with respect to all products available in a given period (not only matched products). See also Section 4.2. |

With a time window consisting of only two periods, the GEKS–Tq naturally corresponds to the bilateral Törnqvist index. The GK corresponds to the bilateral GK, which is a basket price index ([19]) in which each product is weighted according to the harmonic average of the quantities observed in the two comparison periods. Finally, the WTPD corresponds to a weighted geometric average of the price ratios, whereby each price ratio is weighted according to the (normalised) harmonic average of the expenditure shares observed in the two comparison periods. In all cases, the bilateral index formulas are always matched indices: they rely only on those products that are available in the two comparison periods. From this perspective, the GEKS–Tq index can be preferred, as it is related to the Törnqvist index. The Törnqvist index has good properties and it can be easily explained to users.

### 4.4.5. Robustness

For a given index formula, decisions must be made on the time window length and on the splicing technique (see Section 5) in order to obtain non-revisable published indices. From an empirical perspective, some studies seem to suggest that the WTPD and the GK are more sensitive to these choices than the GEKS–Tq (e.g. Lamboray, 2021). In other words, the resulting indices after splicing (the indices eventually published) will be more different depending on the window length and splicing method adopted. These differences tend to be larger for the WTPD and the GK than for the GEKS–Tq.

Dumping occurs when a product is sold at reduced prices before being removed from the assortment. The number of units sold in the last month during which this product is available is usually low. The GEKS–Tq index gives more weight to the price decrease of the dumped product because this price decrease is implicitly affected by the (larger) expenditure shares of that product in past periods (Chessa, 2017). Therefore, dumping filters are best used together with the GEKS–Tq.

([19]) A basket price index is a price index that measures the proportionate change between periods **0** and ***t*** in the total value of a specified basket of goods and services, that is $\sum p^t q / \sum p^0 q$, where the terms $q$ are the specified quantities.

### 4.4.6. Explainability

Multilateral methods are more complex than bilateral methods. One aspect that could be considered when assessing a method is the need to explain it to users. There is an element of subjectivity in assessing if a method is more or less easy to explain. It also depends on the type of users. Advanced users might be interested in a more rigorous description of the methods, while other users would prefer receiving explanations of the methods in simpler terms.

The following can be said about the three multilateral methods discussed in this guide.

- The GEKS index can be explained by referring to the underlying bilateral index (e.g. the Törnqvist index).
- The WTPD index can be explained by referring to the underlying regression. This could be achievable because the estimation of regressions is a well-known statistical technique.
- The GK index can be explained as a quality-adjusted unit value index.

In addition to explainability of the methods, the interpretability can be facilitated by calculating contributions (see Section 8).

## Summary of key points

- Different multilateral index formulas can be regarded as appropriate for implementation. Theoretical, practical and empirical considerations should be taken into account when choosing a specific index formula.
- The GEKS–Tq is an index that is closest, on average, to the bilateral Törnqvist indices calculated between any periods of the time window. The WTPD is an index that is based on a model. The GK index is an example of an additive method.

# 5 Time windows and extension methods

This section examines the question of the length of the time window over which the multilateral index is applied. There are trade-offs to consider in that context. For instance, shorter time windows could lead to unstable results and may not solve the chain-drift problem. The longer the time window, the more data from the past will impact the current-month compilations. It also depends on the product type. For seasonal products, the window should be sufficiently long to cover two successive in-season periods. In practice, the time window should cover at least 13–14 months, if not longer (e.g. 25 months). For products with short life cycles (e.g. consumer electronics), a shorter time window may be warranted.

In a statistical production environment, each month, the time window has to be adjusted to include the data from the latest month. There are various strategies for extending time windows.

- **Rolling time windows**. Each month, the time window is shifted forward by 1 month. The length of the time window is kept constant. For instance, assuming a 13-month rolling time window, the first time window would cover January 2019 to January 2020 (compilation of the January 2020 index). The next time window would cover February 2019 to February 2020 (compilation of the February 2020 index), and so on. Rolling time windows are an intuitive choice for keeping window length constant. The latest month is included, and the oldest month is removed.
- **Expanding time windows**. Each month, the time window is extended by 1 month. The length of the time window increases each month by 1 month. After 1 year, the length of the time window could be reset to its initial length. For instance, the first time window would cover January 2019 to January 2020 (compilation of the January 2020 index). The next time window would cover January 2019 to February 2020 (compilation of the February 2020 index), and so on, until the compilation of the January 2021 index, which will be, again, based on a 2-month period (December 2020 to January 2021). With such a strategy, there is an imbalance in window length and the length at the beginning of the year is very short. The advantage of this strategy is that it can be implemented without the need for data that go back in time.

Each time a new time window is used, the previously calculated indices may change slightly. Therefore, splicing techniques must be used that link the latest multilateral index to previous results in order to avoid revisions of already published results.

Technically, the splicing of two series operates via a link month. There are different options for the selection of the link month when compiling the results in month $t$. The window length is denoted by $T$:

- **movement splice** – the period $t - 1$ is used as the link period;
- **window splice** – the period $t - T + 1$ is used as the link period;
- **half splice** – the period $t - ((T + 1)/ 2) + 1$ is used as the link period;
- **mean splice** – all the overlap periods are used as link periods;
- **fixed base** – the previous December is used as the link period.

Chessa (2019b) pointed out that there are, in fact, two main variants for splicing methods. Successive window shifts generate a sequence of recalculated or 'revised' indices alongside the initial published index in the same period. Both the recalculated and published indices are candidates for the index on which a new index series can be linked. Therefore, we will distinguish two variants of splicing (except for movement splicing, which has the index published in the previous period as the only link option):

- link the multilateral index compiled in period ***t*** with the multilateral index compiled in period ***t* – 1**;
- link the multilateral index compiled in period ***t*** with the published index.

The multilateral index is transitive with respect to a given time window. Transitivity is not satisfied any more for the indices that are eventually published. Therefore, some degree of chain drift cannot be fully excluded for the published (spliced) indices. However, splicing is the only option available if revision of already published figures is not possible.

The formulas of the different techniques are provided below. The methods are also illustrated on a numerical example. We describe the different splicing options using a rolling time window of 13 months, but the methods can be adapted to other time windows. In this example, there are three compilation rounds:

- the multilateral index compiled in period 13 covers periods 1–13;
- the multilateral index compiled in period 14 covers periods 2–14;
- the multilateral index compiled in period 15 covers periods 3–15.

***Movement splice***

Link to the published series:

$$I_{pub}^{0,t} = I_{pub}^{0,t-1} \times I_{[t-T+1,t]}^{t-1,t}$$

The movement splice method is illustrated in Table 14.

**Table 14: Illustration of the movement splice**

| Period | 1 | 2 | 3 | 4 | 5 | … | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| First compilation round in period 13 | 100.0 | 100.7 | 100.6 | 101.6 | 102.7 | … | 104.3 | 106.0 | 103.8 | | |
| Second compilation round in period 14 | | 100.0 | 100.2 | 101.1 | 102.2 | … | 103.8 | 105.5 | 103.3 | 104.6 | |
| Third compilation round in period 15 | | | 100.0 | 101.0 | 102.0 | … | 103.5 | 105.3 | 103.2 | 104.4 | 104.1 |
| **Published index (movement splice – link to published)** | **100.0** | **100.7** | **100.6** | **101.6** | **102.7** | **…** | **104.3** | **106.0** | **103.8** | **105.1** | **104.8** |

The splicing starts in period 14. The published indices for periods 1–13 are obtained at the first compilation round. The published index in period 14 is obtained by applying the change between the period-13 and period-14 indices of the second compilation round to the published index of period 13 (103.8 × (104.6 / 103.3) = 105.1). The published index in period 15 is obtained by applying the change between period-14 and period-15 indices of the third compilation round to the published index of period 14 (105.1 × (104.1 / 104.4) = 104.8).

***Window splice***

Link to the previously calculated series:

$$I_{pub}^{0,t} = I_{pub}^{0,t-1} \times I_{[t-T,t-1]}^{t-1,t-T+1} \times I_{[t-T+1,t]}^{t-T+1,t}$$

Link to the published series:

$$I_{pub}^{0,t} = I_{pub}^{0,t-1} \times I_{pub}^{t-1,t-T+1} \times I_{[t-T+1,t]}^{t-T+1,t} = I_{pub}^{0,t-T+1} \times I_{[t-T+1,t]}^{t-T+1,t}$$

The window splice method is illustrated in Table 15.

**Table 15:** **Illustration of the window splice**

| Period | 1 | 2 | 3 | 4 | 5 | … | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| First compilation round in period 13 | 100.0 | 100.7 | 100.6 | 101.6 | 102.7 | … | 104.3 | 106.0 | 103.8 | | |
| Second compilation round in period 14 | | 100.0 | 100.2 | 101.1 | 102.2 | … | 103.8 | 105.5 | 103.3 | 104.6 | |
| Third compilation round in period 15 | | | 100.0 | 101.0 | 102.0 | … | 103.5 | 105.3 | 103.2 | 104.4 | 104.1 |
| **Published index (window splice – link to previous)** | **100.0** | **100.7** | **100.6** | **101.6** | **102.7** | **…** | **104.3** | **106.0** | **103.8** | **105.3** | **105.0** |
| The splicing starts in period 14. The published indices for periods 1–13 are obtained at the first compilation round. The published index in period 14 is obtained by applying the change between the period-13 and period-2 indices of the first compilation round and the change between the period-2 and period-14 indices of the second compilation round to the published index of period 13 (103.8 × (100.7 / 103.8) × (104.6 / 100.0) = 105.3). The published index in period 15 is obtained by applying the change between the period-14 and period-3 indices of the second compilation round and the change between the period-3 and period-15 indices of the third compilation round to the published index of period 14 (105.3 × (100.2 / 104.6) × (104.1 / 100) = 105.0). | | | | | | | | | | | |
| **Published index (window splice – link to published)** | **100.0** | **100.7** | **100.6** | **101.6** | **102.7** | **…** | **104.3** | **106.0** | **103.8** | **105.3** | **104.7** |
| The splicing starts in period 14. The published indices for periods 1–13 are obtained at the first compilation round. The published index in period 14 is obtained by applying the change between the period-2 and period-14 indices of the second compilation round to the published index of period 2 (100.7 × (104.6 / 100.0) = 105.3). The published index in period 15 is obtained by applying the change between the period-3 and period-15 indices of the third compilation round to the published index of period 3 (100.6 × (104.1 / 100) = 104.7). | | | | | | | | | | | |

***Half splice***

Link to the previously calculated series:

$$I_{pub}^{0,t} = I_{pub}^{0,t-1} \times I_{[t-T,t-1]}^{t-1,t-\left(\frac{T+1}{2}\right)+1} \times I_{[t-T+1,t]}^{t-\left(\frac{T+1}{2}\right)+1,t}$$

Link to the published series:

$$I_{pub}^{0,t} = I_{pub}^{0,t-1} \times I_{pub}^{t-1,t-\left(\frac{T+1}{2}\right)+1} \times I_{[t-T+1,t]}^{t-\left(\frac{T+1}{2}\right)+1,t} = I_{pub}^{0,t-\left(\frac{T+1}{2}\right)+1} \times I_{[t-T+1,t]}^{t-\left(\frac{T+1}{2}\right)+1,t}$$

The half splice method is illustrated in Table 16.

**Table 16: Illustration of the half splice**

| Period | 1 | 2 | 3 | … | 7 | 8 | 9 | … | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| First compilation round in period 13 | 100.0 | 100.7 | 100.6 | … | 104.3 | 102.9 | 104.2 | … | 103.8 | | |
| Second compilation round in period 14 | | 100.0 | 100.2 | … | 103.8 | 102.5 | 103.7 | … | 103.3 | 104.6 | |
| Third compilation round in period 15 | | | 100.0 | … | 103.7 | 102.1 | 103.6 | … | 103.2 | 104.4 | 104.1 |
| **Published index (half splice – link to previous)** | **100.0** | **100.7** | **100.6** | **…** | **104.3** | **102.9** | **104.2** | **…** | **103.8** | **105.0** | **104.6** |
| The splicing starts in period 14. The published indices for periods 1–13 are obtained at the first compilation round. The published index in period 14 is obtained by applying the change between the period-13 and period-8 indices of the first compilation round and the change between the period-8 and period-14 indices of the second compilation round to the published index of period 13 (103.8 × (102.9 / 103.8) × (104.6 / 102.5) = 105.0). The published index in period 15 is obtained by applying the change between the period-14 and period-9 indices of the second compilation round and the change between the period-9 and period-15 indices of the third compilation round to the published index of period 14 (105.0 × (103.7 / 104.6) × (104.1 / 103.6) = 104.6). | | | | | | | | | | | |
| **Published index (half splice – link to published)** | **100.0** | **100.7** | **100.6** | **…** | **104.3** | **102.9** | **104.2** | **…** | **103.8** | **105.0** | **104.7** |
| The splicing starts in period 14. The published indices for periods 1–13 are obtained at the first compilation round. The published index in period 14 is obtained by applying the change between the period-8 and period-14 indices of the second compilation round to the published index of period 8 (102.9 × (104.6 / 102.5) = 105.0). The published index in period 15 is obtained by applying the change between the period-9 and period-15 indices of the third compilation round to the published index of period 9 (104.2 × (104.1 / 103.6) = 104.7). | | | | | | | | | | | |

***Mean splice***

Link to the previously calculated series:

$$I_{pub}^{0,t} = I_{pub}^{0,t-1} \times \prod_{k=t-T+1}^{t-1} \left( I_{[t-T,t-1]}^{t-1,k} \times I_{[t-T+1,t]}^{k,t} \right)^{\frac{1}{T-1}}$$

Link to the published series:

$$I_{pub}^{0,t} = I_{pub}^{0,t-1} \times \prod_{k=t-T+1}^{t-1} \left( I_{pub}^{t-1,k} \times I_{[t-T+1,t]}^{k,t} \right)^{\frac{1}{T-1}} = \prod_{k=t-T+1}^{t-1} \left( I_{pub}^{0,k} \times I_{[t-T+1,t]}^{k,t} \right)^{\frac{1}{T-1}}$$

The mean splice method is illustrated in Table 17.

**Table 17: Illustration of the mean splice**

| Period | 1 | 2 | 3 | 4 | 5 | … | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| First compilation round in period 13 | 100.0 | 100.7 | 100.6 | 101.6 | 102.7 | … | 104.3 | 106.0 | 103.8 | | |
| Second compilation round in period 14 | | 100.0 | 100.2 | 101.1 | 102.2 | … | 103.8 | 105.5 | 103.3 | 104.6 | |
| Third compilation round in period 15 | | | 100.0 | 101.0 | 102.0 | … | 103.5 | 105.3 | 103.2 | 104.4 | 104.1 |
| **Published index (mean splice – link to previous)** | **100.0** | **100.7** | **100.6** | **101.6** | **102.7** | **…** | **104.3** | **106.0** | **103.8** | **105.1** | **104.8** |
| The splicing starts in period 14. The published indices for periods 1–13 are obtained at the first compilation round. The published index in period 14 is obtained by applying the geometric average of the changes (*k* = 2, …, 13) between the period-13 and period-*k* indices of the first compilation round and the changes between the period-*k* and period-14 indices of the second compilation round to the published index of period 13 (103.8 × [(100.7 / 103.8)(104.6 / 100.0) × … × (103.8/103.8) × (104.6 / 103.3)]$^{1/12}$ = 103.8 × 1.012702 = 105.1).<br>The published index in period 15 is obtained by applying the geometric average of the changes (*k* = 3, …, 14) between the period-14 and period-*k* indices of the second compilation round and the changes between the period-*k* and period-15 indices of the third compilation round to the published index of period 14 (105.1 × [(100.2 / 104.6)(104.1 / 100.0) × … × (104.6 / 104.6) × (104.1 / 104.4)]$^{1/12}$ = 105.1 × 0.996914 = 104.8). | | | | | | | | | | | |
| **Published index (mean splice – link to published)** | **100.0** | **100.7** | **100.6** | **101.6** | **102.7** | **…** | **104.3** | **106.0** | **103.8** | **105.1** | **104.8** |
| The splicing starts in period 14. The published indices for periods 1–13 are obtained at the first compilation round. The published index in period 14 is obtained as a geometric average of the changes between the period-*k* and period-14 indices of the second compilation round to the published index of period *k* (*k* = 2, …, 13): [(100.7) × (104.6 / 100.0) × … × (103.8) × (104.6 / 103.3)]$^{1/12}$ = 105.1.<br>The published index in period 15 is obtained as a geometric average of the changes (*k* = 3, …, 14) between the period-*k* and period-15 indices of the third compilation round to the published index of period *k* (*k* = 3, …, 14): [(100.6) × (104.1/100.0) × … × (105.1) × (104.1 / 104.4)]$^{1/12}$ = 104.8. | | | | | | | | | | | |

***Fixed base (December of the previous year)***

Link to the published series:

$$I_{pub}^{0,t} = I_{pub}^{0,t-1} \times I_{pub}^{t-1,Dec(t)} \times I_{[t-T+1,t]}^{Dec(t),t} = I_{pub}^{0,Dec(t)} \times I_{[t-T+1,t]}^{Dec(t),t}$$

The fixed-base method is illustrated in Table 18.

**Table 18: Illustration of the fixed-base splicing**

| Period | 1 | 2 | 3 | 4 | 5 | … | 11 | 12 (base period) | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| First compilation round in period 13 | 100.0 | 100.7 | 100.6 | 101.6 | 102.7 | … | 104.3 | 106.0 | 103.8 | | |
| Second compilation round in period 14 | | 100.0 | 100.2 | 101.1 | 102.2 | … | 103.8 | 105.5 | 103.3 | 104.6 | |
| Third compilation round in period 15 | | | 100.0 | 101.0 | 102.0 | … | 103.5 | 105.3 | 103.2 | 104.4 | 104.1 |
| **Published index (fixed base – link to published)** | **100.0** | **100.7** | **100.6** | **101.6** | **102.7** | **…** | **104.3** | **106.0** | **103.8** | **105.1** | **104.8** |

The splicing starts in period 14. The published indices for periods 1–13 are obtained in the first compilation round. The published index in period 14 is obtained by applying the change between the period-12 and period-14 indices of the second compilation round to the published index of period 12 (106.0 × (104.6 / 105.5) = 105.1). The published index in period 15 is obtained by applying the change between the period-12 and period-15 indices of the third compilation round to the published index of period 12 (106.0 × (104.1 / 105.3) = 104.8).

The performance of a splicing method can be tested empirically by comparing the resulting index with the index compiled over a full time window. For example, consider data that covers several years. A multilateral index can be calculated over this entire time period. This index serves as a fully transitive benchmark. With the same data, a real-time index can be calculated using any of the splicing techniques discussed above. The difference between these two indices can help to evaluate the performance of the splicing method. Note, however, that the 'benchmark' index may also be slightly affected by a very long time window, as the result in a given month can be influenced by very distant data points.

The different splicing methods all have their advantages and disadvantages. Fixed-base splicing is more consistent with a bilateral index in which the December month of the previous year acts as a price reference period. However, such a strategy is heavily dependent on a specific month (December month of the previous year).

Movement splice is the method that can be most easily understood and explained. It is consistent with the month-on-month movements obtained with the multilateral index. However, some extent of chain drift cannot be excluded with this method. The same holds for window and half splice, which may all be subject to chain drift when linking to previously calculated series is performed. Mean splice may be a more robust alternative as it is based on all possible links. However, it is still possible that chain drift of movement or window splice propagates to the mean splice.

An alternative strategy would be to link to previously published series. Consider, for example, a 25-month rolling time window. The splicing is conducted using the half splice on published series method. Under this method, the annual rate of changes of the published indices corresponds to the annual rate of change of the latest calculated multilateral index ([20]). This is a useful property, as annual rates are a key indicator in the context of the HICP. We will refer to the half splice on published series splicing method with a 25-month rolling time window as 25-HASP (Chessa, 2021).

# Summary of key points

- Decisions must be made regarding the window length over which the multilateral method is compiled and the way that successive multilateral indices are linked without revising already published indices.

([20]) This can also be achieved by using window splice on published series over a 13-month rolling time window.

- With splicing, one cannot avoid chain drift entirely; hence the drift must be reduced as much as possible. The validity of a splicing method should be empirically validated.
- Consideration should be given to the half splice on published series method over a 25-month rolling time window (25-HASP).

# 6 Aggregation structures

It is important to identify the 'elementary aggregate' in the aggregation structure. Conceptually, according to the HICP legal framework, the elementary aggregate ([21]) corresponds to the lowest level of a Laspeyres-type index. Such an index uses the December month of the previous year as the price reference period and weights referring to the previous year. The weights attached to the elementary aggregate should be updated at the beginning of the year and they should be kept constant for the rest of the year.

The level at which the elementary aggregate is defined is crucial because it corresponds to the level at which the use of variable weights is changed to the use of fixed Laspeyres-type weights. In all cases, fixed weights must be applied at the ECOICOP subclass level at least.

It is possible that pushing the elementary aggregate level downwards in the index hierarchy will have an upward impact on results. The lower the elementary aggregate is defined, the less flexible the index compilation becomes to incorporate new products or outlets during the year. At the same time, tightly defined elementary aggregates introduce some stability in the index, which makes analysis and validation easier. It also ensures a closer consistency with a Laspeyres-type index. The impact of a multilateral method (e.g. compared with a fixed-weight method) is reduced if the multilateral method is applied only at a very detailed level.

Within an elementary aggregate, there can be one, or sometimes several, elementary price indices ([22]). An elementary price index is a price index that could be obtained with a multilateral method (or any other suitable bilateral method). The multilateral method is applied at the level at which the prices of the individual products are first combined.

In practice, the elementary price indices must be expressed in the price reference period (December of the previous year). This is achieved by dividing the multilateral index by its December $y - 1$ value. Because the index obtained from the multilateral method is (approximately) transitive, it can be expressed in any price reference period by simple re-referencing.

A typical aggregation structure can be represented as follows (Table 19). The ECOICOP subclass (five digits) is the lowest harmonised level of the HICP. The subclass may be further split up according to some classification used at the national level. For each (national) COICOP category, there are usually several retail chains that are explicitly weighted. Within each retail chain, some chain-specific classification could be used to define even finer product categories. Finally, it may be possible to further stratify by outlet.

([21]) Article 2(13) of Regulation (EU) 2020/1148: ' "elementary aggregate" means the smallest aggregate used in a Laspeyres-type index'.

([22]) Article 2(14) of Regulation (EU) 2020/1148: ' "elementary price index" means an index for an elementary aggregate or an index for a stratum within an elementary aggregate'.

**Table 19: Level of aggregation and aggregation methods**

| Level | Option 1 | Option 2 | Option 3 | Option 4 |
|---|---|---|---|---|
| ECOICOP (or some national refinement of ECOICOP) | Laspeyres-type | Laspeyres-type | Laspeyres-type | Laspeyres-type |
| Retail chain | Laspeyres-type | Laspeyres-type | Laspeyres-type (= EA) | Laspeyres-type (= EA) |
| Product category | Laspeyres-type | Laspeyres-type (= EA) | | Fixed-base Törnqvist |
| Outlet | Laspeyres-type (= EA) | | | |
| Individual product | Multilateral | Multilateral | Multilateral | Multilateral |
| Transacted product | Unit value | Unit value | Unit value | Unit value |

EA, elementary aggregate.

As discussed in Section 3, there are different choices for the individual product. The individual products themselves are the result of a unit value aggregation. If the individual product is already the result of an aggregation over outlets of the same retail chain (average price over all outlets of a retail chain), then option 1 is not possible any more. If the individual product is defined by outlet, then all options are possible.

In option 1, the elementary aggregate is defined rather tightly. It covers the transactions of a detailed product category in a specific outlet. In option 3, the elementary aggregate is defined rather broadly. It covers the transactions of a given COICOP category (e.g. ECOICOP subclass, or sub-subclass of a nationally defined COICOP) in a given retail chain. The disadvantage of option 1, compared with option 3, is that openings or closures of outlets and new or obsolete product categories are more difficult to handle. In practice, under option 1, such changes can be incorporated in the index through standard sample updates and chaining conducted, for example, once per year. In option 3, new outlets are incorporated implicitly in the unit value calculations (if the individual product is defined by retail chain), or they are incorporated after two matched price quotes are available (if the individual product is defined by outlet). Moreover, there can be changes every month in market shares of product categories or outlets. Such changes will not be captured in option 1, but they will be reflected in option 3. If there are no such shifts and the market shares of product categories or outlets are relatively stable over time, then the distinction between options 1 and 3 is less important. Option 2 is an example of an intermediate solution that skips the outlet level and directly aggregates the individual products up to the level of a retail-specific product category.

Option 4 is a more advanced solution. The elementary aggregate level is defined in the same way as in option 3. However, each elementary aggregate consists of several elementary price indices. In fact, a multilateral method is used to come up with price indices by product category. These elementary indices are then aggregated using, for instance, a Törnqvist index. Hence the aggregation at the product category level is still based on variable weights. It is only at the next level that fixed weights are used and the Laspeyres principle is applied. The result of option 4 is likely to be close to the result obtained with option 3. It has, however, the advantage that explicit subindices are compiled for each product category, which can facilitate the analysis and validation of the data.

One also needs to decide at which level the data from different sources (scanner data, web-scraped data, field price collection) could be combined. For instance, the ECOICOP subclass level (or any other level below the subclass) could be stratified by data source. Each data source corresponds to a stratum of the subclass to which a weight must be assigned and for which a price index is compiled. Fixed weights must be used at the level at which different data sources are combined because only fixed weights are available for the stratum that represents the traditionally collected prices.

# Summary of key points

- The advantage of applying a multilateral method up to a more aggregated level is the ability to cope with a dynamic universe of individual products. Keeping the variable weights as high as possible in the index hierarchy allows changes in the importance of products and/or outlets over time to be captured.
- The use of tightly defined elementary aggregates introduces more stability in the index. Such a strategy facilitates the analysis of results and ensures a closer consistency with a Laspeyres-type index, but captures a dynamic universe less well.
- The index structure below the subclass level and the elementary aggregate weights should be reviewed and updated at least once per year.

# 7 Seasonal products

In principle, multilateral methods can be applied to seasonal products. A seasonal product (23) is an individual product that is available for purchase, or is purchased in significant quantities, for only part of a year in a recurring pattern. In any given month, the product is considered to be either in season or out of season. The in-season period may vary from one year to another. As such, no specific treatment is required. However, because of a high variability in prices and quantities and a large and recurrent share in missing prices during the out-of-season periods, the different multilateral methods are likely to give more diverse results. Note that specifying homogeneous products sufficiently broadly (see Section 3) means that the homogeneous products are likely to be available in each month, and therefore the seasonal product problem vanishes.

Conceptually, multilateral methods can be seen as an example of a seasonal weight method (24). When an individual product is out of season, its weight is zero. When a product is in season, its weight corresponds to the actually observed quantities. The class-confined seasonal weight (CCSW) method (see Chapter 7.1 of the HICP manual) is another example of a seasonal weight method. In the CCSW method, a product that is out of season also has a zero weight. However, there are some differences. In the CCSW method, the weights for the in-season period are estimated a priori and derived from some past expenditure survey that does not necessarily reflect consumption expenditure in the current periods. Moreover, the weights in the CCSW method are adjusted so that the weights are fixed if the composition of the seasonal basket does not change from one month to the next.

Window length should be set so that it can include the last month of the previous in-season period and the first month of the new in-season period. This is important so that any price change between the previous and the current in-season periods is captured. If that is not the case, the first time that a seasonal product reappears, there is no match with the previous season and the price change is lost. That is why fixed-base expanding time window strategies (see Section 5) tend to perform less well with seasonal products. In order to capture products that are on the market only for 1 month per year, it is important to have windows of at least 13 months.

The choice of the splicing method is critical when dealing with seasonal products. Although largely an empirical matter, the spliced indices often tend to deviate (drift away) from a benchmark index calculated without splicing. There is an issue with the fixed-base splicing. With December being the 'link' month, the long-term trend of the index is determined, to a large extent, by those seasonal products that are available in December. A 25-HASP may be a good compromise because it ensures consistent annual rates. As a result, this method has some resemblance with over-the-year indices. At the same time, the multilateral index ties together all months of the time period, as opposed to over-the-year indices that only compare this month with the same month 1 year ago.

(23) Article 2(22) of Regulation (EU) 2020/1148: ' "seasonal product" means an individual product that is available for purchase or purchased in significant amounts only part of a year in a recurring pattern. In any given month, the product is considered to be either in-season or out-of-season. The in-season period may vary from one year to another'.

(24) Article 2(27) of Regulation (EU) 2020/1148: ' "seasonal weights method" means a treatment of seasonal products in which weights for out-of-season seasonal products are zero or set to zero'.

A data set with seasonal products and the corresponding multilateral indices can be found online [25] (example provided by STATEC).

# Summary of key points

- Individual seasonal products do not need to be explicitly identified in the data, but the compilation methodology for categories that include seasonal products should be carefully examined.
- In order to have a good coverage of two successive in-season periods, it is important to have time windows longer than 13 months.
- The impact of splicing should be closely examined with seasonal products, as some splicing methods may produce indices that are biased compared with (i.e. deviate systematically from) benchmark indices compiled without splicing.

[25] https://circabc.europa.eu/ui/group/7b031f10-ac19-4da3-a36f-58708a70133d/library/e52b1b37-38f3-4f2d-82be-14b005da1ce3/details

# 8 Contributions

The result of a multilateral index depends not only on the data of the two comparison periods, but on all the data of the time window. It is technically possible to calculate the impact that each individual product has on the result. The calculation of such contributions is an important tool for compilers to validate and to interpret the results. This section describes the methods presented in Webster and Tarnow-Mordi (2019) that can be used to compile such contributions.

The contributions refer to a multilateral price index compiled over a given time window, before any splicing is conducted. The contributions of the published results would have to be further adjusted in order to take into account the effect of the splicing. Note that, when using the 25-month half splice on published indices, the contributions of the annual rate of change from the time window equal the contributions for the published annual rate of change. Similarly, when using movement splicing, the contributions of the monthly rate of change from the time window equal the contributions for the published monthly rate of change.

The contribution calculations on a data set with missing products can be found online [26].

## 8.1. Contributions for the GEKS-Törnqvist

Bilateral shares for each individual product *i* with respect to any two periods *t*1 and *t*2 are defined as follows:

$$w_i^{t1,t2} = 0.5 \times \left( \frac{p_i^{t1} q_i^{t1}}{\sum_{j \in N_{t1} \cap N_{t2}} p_j^{t1} q_j^{t1}} + \frac{p_i^{t2} q_i^{t2}}{\sum_{j \in N_{t1} \cap N_{t2}} p_j^{t2} q_j^{t2}} \right)$$

Note that these shares are calculated with respect to the matched products in the two periods *t*1 and *t*2. If an individual product *i* is not available in both of the periods *t*1 and *t*2, then $w_i^{t1,t2}$ is set to zero.

An average bilateral share over the whole time window *W* is defined as follows for each individual product *i* with respect to a given period *t*:

$$w_i^{*,t} = \frac{1}{|W|} \sum_{r \in W} w_i^{r,t}$$

Note that, if an individual product is not available in period *t*, then $w_i^{*,t}$ must be zero.

Let us now consider a GEKS–Tq price index between periods *t*1 and *t*2. A multiplicative decomposition of the GEKS–Tq index can be obtained as follows:

([26]) https://circabc.europa.eu/ui/group/7b031f10-ac19-4da3-a36f-58708a70133d/library/4ffff85a-8485-42e8-b069-012313be0758/details

$$I_{W(GEKS-Tq)}^{t1,t2} = \prod_{i \in N} \frac{\left(p_i^{t2}\right)^{w_i^{*,t2}}}{\left(p_i^{t1}\right)^{w_i^{*,t1}}} \prod_{t \in W} \left(p_i^t\right)^{\frac{w_i^{t,t1} - w_i^{t,t2}}{|W|}}$$

All products available in at least one period of the time window must enter this formula. If a price of product ***i*** is not observed in period ***t*1** or ***t*2**, it is simply replaced in this formula with a '1'. The advantage of this decomposition is that the contributions of an individual product depend only on the prices and expenditure shares of that specific product, and not of the other products.

This method is illustrated in the example described in Section 4.1. We would like to decompose the aggregate price change in period 3, compared with period 0. According to the GEKS–Tq, this price change stands at 0.9822.

Below are the bilateral shares $w_1^{0,t}$ and $w_1^{3,t}$, as well as the average shares $w_1^{*,0}$ and $w_1^{*,3}$ for product 1.

| | Period 0 | Period 1 | Period 2 | Period 3 | Average |
|---|---|---|---|---|---|
| Period 0 | 6.8 % | 7.9 % | 10.7 % | 9.2 % | (6.8 % + 7.9 % + 10.7 % + 9.2 %) / 4 = 8.6 % |
| Period 3 | 9.2 % | 10.3 % | 13.0 % | 11.5 % | (9.2 % + 10.3 % + 13.0 % + 11.5 %) / 4 = 11.0 % |

The following contribution of product 1 in period 3, compared with period 0, can be calculated:

$$\frac{\left(p_1^3\right)^{w_1^{*,3}}}{\left(p_1^0\right)^{w_1^{*,0}}} \prod_{t \in W} \left(p_1^t\right)^{\frac{w_1^{t,0} - w_1^{t,3}}{4}} = \left(\frac{3.03^{11.0\,\%}}{2.97^{8.6\,\%}}\right)\left(2.97^{\left(\frac{6.8\,\% - 9.2\,\%}{4}\right)} \times 2.96^{\left(\frac{7.9\,\% - 10.3\,\%}{4}\right)} \times 2.93^{\left(\frac{10.7\,\% - 13\,\%}{4}\right)} \times 3.03^{\left(\frac{9.2\,\% - 11.5\,\%}{4}\right)}\right) = 1.00218$$

The contributions can be calculated in the same way for the other individual products. In the end, the following results are obtained:

- contribution of product 1 (period 3 compared with period 0): 1.00218;
- contribution of product 2 (period 3 compared with period 0): 0.98167;
- contribution of product 3 (period 3 compared with period 0): 0.99854;
- contribution of product 4 (period 3 compared with period 0): 0.99982.

The multiplication of these four contributions gives 0.9822.

## 8.2. Contributions for the weighted time–product dummy

A multiplicative decomposition of the WTPD index can be obtained as follows:

$$I_{W(WTPD)}^{t1,t2} = \prod_{i \in N} \frac{\left(p_i^{t2}\right)^{s_i^{t2}}}{\left(p_i^{t1}\right)^{s_i^{t1}}} \left(\exp\left(\hat{\alpha} + \widehat{\gamma}_i\right)\right)^{\left(s_i^{t2} - s_i^{t2}\right)}$$

When an individual product is not sold in a given period, its share is '0' and its price is '1' in the formula. Other decompositions of the WTPD can be developed (Webster and Tarnow-Mordi, 2019).

This method is illustrated in the example described in Section 4.2. We would like to decompose the aggregate price change in period 3, compared with period 0. According to the WTPD, this price change stands at 0.9831.

The following contribution of product 1 in period 3, compared with period 0, can be calculated:

$$\frac{\left(p_1^3\right)^{s_1^3}}{\left(p_1^0\right)^{s_1^3}}\left(\exp(\hat{\alpha})\right)^{\left(s_1^3 - s_1^0\right)} = \frac{3.03^{11.5\ \%}}{2.97^{6.8\ \%}}\left(\exp(1.10754)\right)^{(6.8\ \% - 11.5\ \%)} = 1.00141$$

The contributions can be calculated in the same way for the other individual products. In the end, the following results are obtained:

- contribution of product 1 (period 3 compared with period 0): 1.0014;
- contribution of product 2 (period 3 compared with period 0): 0.9815;
- contribution of product 3 (period 3 compared with period 0): 0.9997;
- contribution of product 4 (period 3 compared with period 0): 1.0005.

The multiplication of these four contributions gives 0.9831.

## 8.3. Contributions for the Geary–Khamis

In addition to the observed share $s_i^t$ for an individual product ***i*** in period ***t***, the following adjusted share can be calculated when the observed prices are replaced with the reference prices:

$$\sigma_i^t = \frac{q_i^t v_i}{\sum_{j \in N_t} q_j^t v_j}$$

Let us now consider a GK index between periods *t*1 and *t*2. An additive decomposition of the GK index can be obtained as follows:

$$I_{W(GK)}^{t1,t2} = \sum_{i:q_i^{t1} > 0} s_i^{t1} \frac{\sigma_i^{t2}}{\sigma_i^{t1}} \frac{p_i^{t2}}{p_i^{t1}} + \frac{1}{I_{W(GK)}^{0,t1}} \sum_{i:q_i^{t1} = 0} s_i^{t2} \frac{p_i^{t2}}{v_i}$$

The second term vanishes if the set of individual products is the same in the two comparison periods. Other decompositions of the GK can be developed (Webster and Tarnow-Mordi, 2019).

This method is illustrated in the example described in Section 4.3. We would like to decompose the aggregate price change in period 3, compared with period 0. According to the GK, this price change stands at 0.9833.

First, adjusted shares must be calculated.

| | Product 1 | Product 2 | Product 3 | Product 4 |
|---|---|---|---|---|
| Period 0 | 7.0 % | 23.7 % | 51.1 % | 18.2 % |
| Period 1 | 11.3 % | 35.9 % | 40.9 % | 11.9 % |

The following contribution of product 1 in period 3, compared with period 0, can be calculated:

$$s_1^0 \frac{\sigma_1^3}{\sigma_1^0} \frac{p_1^3}{p_1^0} = 6.8\ \% \frac{11.3\ \%}{7.0\ \%} \frac{3.03}{2.97} = 0.1131$$

The contributions can be calculated in the same way for the other individual products. In the end, the following results are obtained:

- contribution of product 1 (period 3 compared with period 0): 0.1131;
- contribution of product 2 (period 3 compared with period 0): 0.3483;
- contribution of product 3 (period 3 compared with period 0): 0.4036;
- contribution of product 4 (period 3 compared with period 0): 0.1182.

The sum of these four contributions gives 0.9833.

# 9 Harmonised Index of Consumer Prices at constant tax rates

The HICP at constant tax rates (HICP-CT) is an index that measures price changes in the same way as the HICP, except that rates of taxes on products are kept constant in the comparison period vis-à-vis the price reference period (December $y-1$). In practice, the prices observed in the HICP are replaced by constant-tax prices. The calculation of the constant-tax prices is explained in detail in Chapter 9 of the HICP methodological manual. The resulting elementary price indices are then aggregated with the same fixed weights both in the HICP and in the HICP-CT.

## 9.1. Treatment of the weights

The HICP-CT compilation becomes more complicated if variable weights are attached to the individual products within the elementary aggregate. The main principle is that the difference between the elementary price indices of the HICP and the HICP-CT should be driven only by differences in prices due to tax rate changes. The treatment of the weights depends on the index formula that is used to compile the elementary index.

Consider a bilateral index formula that is based on prices $\boldsymbol{p}_0$ and $\boldsymbol{p}_t$, and on the corresponding quantities $\boldsymbol{q}_0$ and $\boldsymbol{q}_t$. The turnover values are denoted by $\boldsymbol{v}_0$ and $\boldsymbol{v}_t$. A bilateral index is obtained either from the prices and quantities or from the prices and turnover values. Let us denote by $\boldsymbol{p}_t^{CT}$ the constant-tax price in period $t$ using the tax rates observed in the base period 0.

The same quantities should be used in the HICP and in the HICP-CT calculations if the index formula satisfies the following condition:

$$\frac{I\left(p_0, p_t^{CT}, q_0, q_t\right)}{I\left(p_0, p_t, q_0, q_t\right)} = I\left(p_t, p_t^{CT}, q_0, q_t\right)$$

For example, the Walsh and Fisher indices satisfy this condition. Therefore, when using a Walsh or Fisher index, the quantities should be kept unchanged.

Similarly, the same turnover values should be used in the HICP and in the HICP-CT calculations if the index formula satisfies the following condition:

$$\frac{I\left(p_0, p_t^{CT}, v_0, v_t\right)}{I\left(p_0, p_t, v_0, v_t\right)} = I\left(p_t, p_t^{CT}, v_0, v_t\right)$$

For example, the Törnqvist index satisfies this condition. Therefore, when using a Törnqvist index, the turnover should be kept unchanged.

In order to make a decision for a multilateral index formula, the time window can be restricted to two periods only. The corresponding bilateral index formulas are described in Section 4.4.4. It follows that the following strategy is to be applied when calculating an HICP-CT with a multilateral index formula.

| Index formula | Treatment of weights |
|---|---|
| GEKS–Tq | Turnover weights |
| WTPD | Turnover weights |
| GK | Quantity weights |

The computer program that is used to calculate indices may take as input either quantity or turnover weights ([27]). Note that changing the observed price to a constant-tax price while keeping the quantity (the turnover) unchanged implies that the turnover (the quantity) for that individual product in the constant-tax calculations is different from the turnover (the quantity) used in the HICP. In the following example, the observed price for an individual product in a given month is EUR 2.50. The corresponding constant-tax price is EUR 2.40. The relationship between price, quantity and turnover is always preserved (turnover equals price times quantity). Note that, if the tax rate change is proportional, then both quantity shares and turnover shares will not change within an elementary aggregate.

| | Initial data | Same turnover | Same quantity |
|---|---|---|---|
| ***p*** **(price)** | 2.50 | 2.40 (CT price) | 2.40 (CT price) |
| ***q*** **(quantities)** | 500 | 1 250 / 2.40 = 520.83 | 500 |
| ***v*** **(turnover)** | 1 250 | 1 250 | 2.40 × 500 = 1 200 |

CT, constant tax.

## 9.2. Compilation of the Harmonised Index of Consumer Prices at constant tax rates with a multilateral method

The calculation of the HICP-CT with a multilateral method should comply with the following two principles.

- It should be consistent with the framework of the HICP-CT. This means that the price change between a current month and the December month should be based on the tax rates of the December month.
- It should be consistent with the multilateral method used in the HICP. If there are no tax rate changes, then the HICP (calculated with a multilateral method) and the HICP-CT should be the same.

We suppose below that the multilateral index is calculated using a 25-month rolling time window with a half splice on published indices (25-HASP). The calculation of the HICP-CT then follows the steps outlined below.

***Step 1: Calculate constant-tax prices***

In a bilateral context, we would replace the prices in a given month with their corresponding constant-tax prices using the tax rates of the December of year ***y*** – 1. In a multilateral context, we will need to replace all prices of the entire time window that ends in the current month. The weights need to be adjusted, as discussed in Section 9.1.

In month ***m*** of year ***y***, the following adjusted data sets have to be constructed:

- a first data set in which all prices of the last 25 months are replaced using the tax rates of December of year ***y*** – 2;
- a second data set in which all prices of the last 25 months are replaced using the tax rates of December of year ***y*** – 1.

([27]) For example, a computer program could calculate a Törnqvist index taking as input the prices and quantities of the individual products. Expenditure shares would be calculated by multiplying the prices by the quantities.

In order to ensure consistency between the HICP and the HICP-CT, all individual products should be adjusted to the **$y-1$** and **$y-2$** December tax rates, independently, whether or not these individual products were actually sold, in the respective December month.

***Step 2: Calculate multilateral indices***

After replacing the observed prices, a multilateral price index is calculated from the adjusted data sets. For consistency reasons, the compilation of the HICP-CT should be based on the same index formula as the HICP.

- A first multilateral index is calculated with prices expressed in the tax rates of December **$y-2$** calculated over a period that spans from month **$m$** of year **$y-2$** to month **$m$** of year **$y$**.
- A second multilateral index is calculated with prices expressed in the tax rates of December **$y-1$** calculated over a period that spans from month **$m$** of year **$y-2$** to month **$m$** of year **$y$**.

In practice, the tax rates may not be known at the level of the individual product. Therefore, it may not be possible to derive constant-tax prices by individual product. It may, nevertheless, be possible to make some assumptions about the tax rates that are applied. One strategy consists of estimating a 'constant tax' multilateral index directly from the original multilateral index without going back to the level of individual products. Suppose that the ratio between the observed price and the constant-tax price is the same for all individual products in a given period. The same proportion will also be observed at the aggregate level when comparing the multilateral index (based on observed prices) with the multilateral index (based on constant-tax prices). This is because the multilateral methods discussed in this guide all pass the homogeneity-in-prices test (see Section 4.4.2). Further analysis is required to derive relationships between the original and constant-tax multilateral indices in settings that involve more complex tax schemes.

***Step 3: Splice the multilateral indices***

The HICP-CT is then obtained by augmenting the published HICP-CT 1 year ago with an annual price change. This annual price change is composed of two elements.

- The first element is a price change between month **$m$** of year **$y-1$** and December of year **$y-1$**. This price change uses the tax rates of December **$y-2$**.
- The second element is a price change between December of year **$y-1$** and month **$m$** of year **$y$**. This price change uses the tax rates of December **$y-1$**.

$$I_{HICP-CT}^{m,y} = I_{HICP-CT}^{m,y-1} \times \frac{I_{(m,y-2),(m,y)}^{Dec,y-1}\left(Taxes\ of\ Dec\ y-2\right)}{I_{(m,y-2),(m,y)}^{m,y-1}\left(Taxes\ of\ Dec\ y-2\right)} \times \frac{I_{(m,y-2),(m,y)}^{m,y}\left(Taxes\ of\ Dec\ y-1\right)}{I_{(m,y-2),(m,y)}^{Dec,y-1}\left(Taxes\ of\ Dec\ y-1\right)}$$

Note that, if there are no tax rate changes, which means that the current taxes, the taxes of December **$y-2$** and the taxes of December **$y-1$** are all the same, then the HICP-CT will be identical to the HICP with a 25-HASP splicing method.

# 9.3. Consistency in the monthly rates between the Harmonised Index of Consumer Prices and the Harmonised Index of Consumer Prices at constant tax rates

Differences in the monthly rates of change between the HICP and the HICP-CT appear in the month when there is change in the tax rate. In the other periods, there should be no significant differences as long as the tax is defined as a percentage of the price. However, this kind of consistency does not necessarily hold if taxes are defined as a monetary amount per physical unit of the product. In such a situation, the monthly rate of change of the HICP and of the HICP-CT can be different, even in those months when there are no tax rate changes. This is not specific to multilateral methods, but can also happen with weighted or unweighted bilateral methods.

Some experimental indices to illustrate the impact of changes in non-proportional taxes on multilateral and bilateral indices can be found online [28] (example provided by Statistics Norway).

# Summary of key points

- For the purpose of the calculation of the HICP-CT, observed prices should be replaced with prices based on the tax rates applicable in the respective December months. This treatment should be applied to all individual products in the entire time window. The weights to be used in the HICP-CT are dependent on the index formula.
- The index formula and window length in the HICP-CT should be consistent with the index formula and window length used in the HICP.
- If the overlap period for the splicing belongs to the previous year $y-1$, two multilateral series must be calculated involving two constant tax rates (referring to December $y-1$ and December $y-2$). The splicing method used in the HICP-CT must be adjusted and include the two series with the two constant tax rates.

[28] https://circabc.europa.eu/ui/group/7b031f10-ac19-4da3-a36f-58708a70133d/library/8e823072-4a89-4a4e-ba5b-686cccc578e7/details

# 10 Introducing multilateral methods in the regular production of the Harmonised Index of Consumer Prices

A number of preparatory steps are usually required before a multilateral method can eventually be used in the regular production of the HICP. This section explains the main phases, and highlights issues that should be considered for a successful introduction. Any project on multilateral methods must be adapted to national circumstances, such as the scope of the transaction data to be used (in terms of product and outlet coverage), the level of experience and expertise with scanner data and other alternative data sources, the resources and skills available within the NSI, the methods and sources applied in the HICP or the existence of specific user needs. The NSI should carefully plan the introduction of multilateral methods in order to make sure that the indices calculated with the new methodology are reliable. Transparency about the methods is an important element to safeguard the trust of users in the HICP.

## 10.1. Test phase

In order to get started, the NSI must first become familiar with the technical context of multilateral methods. This guide should help compilers to better understand the different aspects of implementing multilateral (or bilateral) methods in the HICP.

The methods should be applied and tested on real scanner data available to the NSI. One could first calculate indices for a few products, retailers or outlets before scaling up the analysis. Scanner data from supermarkets are usually well suited to price index compilation, but, in principle, multilateral methods can also be applied to other types of products or retailers. It may be that scanner data are already in use in the HICP albeit with a different methodology, so the multilateral methods can be applied to the same input data.

The objective of the explanatory analysis is to get a better understanding of the methods and to assess the impact of the methodological choices on results (see for example Białek and Beręsewicz, 2021). Different index formulas, window lengths and splicing techniques should be tried out and the resulting indices should be compared. It can also be interesting to compare multilateral with bilateral approaches. The experimental results should also be compared with the published indices, but it has to be kept in mind that there can be significant differences in terms of samples or index compilation methods between the multilateral and the published indices. The CPI manual (ILO et al., 2020, paragraph 10.89) notes the following:

*Several insights can be obtained from producing empirical results. Often these insights further reinforce the theoretical arguments for utilizing multilateral index methods to compile the CPI. This may include the impact of using contemporaneous information for weighting purpose that capture consumer behaviour, including substitution, over time. The empirical results should be communicated to CPI users and stakeholders.*

Some programming is required to calculate the indices. Each NSI can develop its own codes. Another possibility is to use available R packages ([29]). The *PriceIndices* package is explained in Białek (2021). Such packages can help an NSI to easily and quickly apply different methods ([30]) and can serve as a tool to benchmark or to further develop its own code.

During this initial test phase, any insights and results can be shared and discussed with users and other stakeholders. This gives visibility to the project and allows the NSI to collect further feedback.

It is recommended to have 1 year (at least) of test calculations before introducing a new method into the production.

## 10.2. Finalise the methodology

At the end of the test phase, a decision must be made concerning the specific method that will actually be implemented. The framework outlined in Section 2 describes the main price index compilation options. A key distinction has to be made between, on the one hand, methods that rely on fixed annual weights from the past and, on the other hand, methods that rely on weights observed in the price comparison periods. As outlined in Section 3, decisions must be made concerning the specification of the individual product. In Section 4, the main multilateral index formulas are presented. Decisions have to be made concerning window length and splicing method (see Section 5). In addition to the choice of a specific index compilation method, an appropriate aggregation structure must be designed (see Section 6). It must be decided how the multilateral method is integrated into the HICP, and combined with indices based on other data sources and compilation methods.

With multilateral methods, special care has to be taken when starting a new series. Suppose that a multilateral index is started in December of year $t$. Assuming a 25-month time window, this means that an index must be compiled over a period that spans from December of year $t - 2$ to December of year $t$. In January, the next index is compiled over the period from January of year $t - 1$ to January of year $t + 1$ and spliced according to the chosen method.

In principle, 25 months of back data are required in order to calculate a multilateral index over a 25-month time window. If, for example, only 13 months of back data are available, one workaround would be to estimate backwards prices using an appropriate proxy price index. Choices must also be made on how to estimate the corresponding quantities. Another possibility would be to introduce the multilateral method with an expanding window in its first year of introduction, and switch to a 25-month rolling time window from the second year of introduction onwards. Research on this topic, and on the impact of imputations, is ongoing.

When introducing a multilateral method into the HICP for the first time, the new series must be linked to the old series. Chain-linking is needed in order to make a transition from an old index to the (spliced) multilateral index.

## 10.3. Build a production system

A production system must be built in order to be able to regularly produce the indices according to the agreed methodology. Such a production system must be appropriately tested to ensure that the calculations are correct. The new system could also be tested in real conditions by running a shadow production in parallel with the actual production for some time.

([29]) See, for example, the packages *IndexNumR* and *PriceIndices*, available online (https://CRAN.R-project.org/package=IndexNumR and https://CRAN.R-project.org/package=PriceIndices).

([30]) Such an analysis, with many examples, can be found online (https://github.com/JacekBialek/important_documents/blob/main/SP.zip).

This guide focuses on the methodological aspects of multilateral indices. These steps must be embedded in a production pipeline and a workflow that deals with new data sources. Such a pipeline typically includes many processes and subprocesses, such as:

1. collecting the data;
2. cleaning the data;
3. classifying the data;
4. editing the data (imputation);
5. coding the units (specification of the individual product) (see Section 3);
6. calculating the indices (multilateral methods) (see Section 4);
7. integrating with the other data sources (see Section 6);
8. reviewing and validating the outputs (see Section 8);
9. calculating the higher-level aggregates.

The special features of multilateral methods must be taken into account when designing the processes and databases for the calculation of the indices. For example, it must be possible to work with data spanning more than 1 year and the resulting indices will also cover more than 1 year. A distinction must be made between the index calculated over a rolling window and the spliced index. The calculations also include some intermediate steps, such as a matrix of bilateral indices (in the case of GEKS) or quality-adjustment factors (in the case of WTPD and GK). Such intermediate results could be stored in order to facilitate validation and replicability.

The production system should set out tools to analyse and validate the results obtained with multilateral methods. Such an analysis can be conducted at the level of the price data, by using the decomposition methods presented in Section 8, or, at a more aggregated level, by examining the subindices. It is common practice to calculate and review rates of change of subindices, and the contributions of the subindices to the rates of change of higher-level aggregates. A time series analysis of the subindices may also help to identify unexpected behaviour. It may be efficient to adopt a combination of both bottom-up and top-down validation approaches.

The production system should also include a plan in case some of the data are missing (e.g. data from a retailer were not supplied). One strategy consists of imputing the data for that retailer based on the previous period data. For each individual product, the current period price (and quantity) is imputed by multiplying the price (and the quantity) of the previous period by a given factor. The calculation system can then be run as usual, including the imputed data.

## 10.4. Communicate about the methodological change

Before introducing a multilateral method in production, impact calculations should be conducted in order to assess how much the methodological change will impact the published indices. Article 9(3) of Regulation (EU) 2016/792 requires that, if a Member State intends to introduce a significant change in the production methods, it must inform Eurostat thereof at the latest 3 months in advance and provide Eurostat with a quantification of the impact of the change.

Every year in autumn, Eurostat conducts a survey in order to ask NSIs about planned methodological changes in the following year. Changes in relation to multilateral methods, along with other relevant changes, should be reported in this survey. Moreover, the HICP inventories are submitted to Eurostat by the end of March each year (Article 27(1) of Regulation (EU) 2020/1148). The relevant parts of the inventory (e.g. concepts related to scanner data) should be reviewed and updated if multilateral methods have been introduced in that year.

The introduction of a change related to multilateral methods should be communicated as clearly as possible to users and other parties interested in the index. This may involve the use of many additional different ways to communicate the introduction of multilateral methods at national and other levels. For example, major changes can be pre-announced, or announced in the national HICP/CPI press releases.

Methodological papers with more details can be published on the website of the NSI. Last but not least, the new methodology could also be presented in specialised technical workshops, in CPI advisory committees or to other specific stakeholders.

# Summary of key points

- The introduction of a multilateral method should be preceded by a test phase involving data covering at least 1 year.
- In addition to the choice of a specific multilateral method, decisions must be made on how to integrate the new series into the HICP.
- The calculation methodology must be embedded in a production system. Tools should be implemented to validate the results of the multilateral method.
- The introduction of a multilateral method in the HICP should be communicated to users in a transparent manner.

# Glossary

## BILATERAL INDICES

A category of index number formulas that measures the aggregate price change between two periods based on prices observed in these two periods only. Bilateral indices can be implemented in a direct form, by comparing the current period with a fixed base period, or in a chained form, by updating the base period in each period and chaining the month-on-month price developments

## CHAIN DRIFT

An undesired property associated with the application of chained indices, whereby the prices and quantities of products typically bounce up and down, thus causing indices to systematically drift away from their expected price trends. A way of measuring the degree of chain drift is to apply the multiperiod identity test

## CHAINED INDICES

An index number category in which indices are calculated through a chain system of price changes calculated using a bilateral index number formula for pairs of adjacent periods. Under this system, the basket of products could be updated every month, which reduces the lack of matching that occurs when fixed-base indices are used. The use of a chained index system often causes chain drift

## CHARACTERISTICITY

An index number property used in the multilateral context that assesses the degree to which a price change between two periods is not affected by prices from periods other than the ones being compared. There is a trade-off between this property and transitivity

## DUMPING FILTER

A procedure applied to data sets in which an observation is excluded from index calculations if both the price and the quantity go down considerably compared with the previous period. This filter is usually applied to identify and eliminate clearance sales, which may inappropriately influence the resulting price indices

## DYNAMIC-BASKET METHOD

A bilateral index number approach (also known as monthly resampling and chained Jevons index) in which the set of products that enters the index compilation is resampled every month. The resampling is carried out through the application of cut-off sampling to select the most-sold products in two consecutive periods, and the final price index is obtained by chaining month-on-month Jevons indices

## EXPANDING TIME WINDOWS

An approach in which the time window over which the multilateral index is calculated is extended by 1 month in each month. The length of the time window increases each month by 1 month. At some point (e.g. at the beginning of each year), the time window could be reset to its initial length

## FIXED-BASE SPLICING

A splicing technique used in the implementation of multilateral price indices in which the December month of the previous year acts as the link period

## GEARY–KHAMIS (GK)

A multilateral index number formula that is obtained by solving a system of equations. The index can be seen as an implicit price index defined as a value index divided by a quantity index, and therefore resembles the calculations of unit values used at the level of homogeneous products

## GINI–ELTETÖ–KÖVES–SZULC (GEKS)

A multilateral index number formula that uses, as its building block, a bilateral index formula. Although the initial GEKS method was built on a Fisher index, it can be based on any other bilateral index that satisfies the time reversal test, such as the Törnqvist index (GEKS–Tq, or CCDI (Caves, Christensen and Diewert)) or the Walsh price index formula

## GLOBAL TRADE IDENTIFICATION NUMBER (GTIN)

An international identifier for products, which has incorporated the former European Article Number (EAN) standard. The GTINs are available in 8-, 12-, 13- or 14-digit formats. GTINs are often found to be the most granular product level available in a data set

## HALF SPLICE

A splicing technique used in the implementation of multilateral price indices in which the period $t - ((|T| + 1) / 2) + 1)$ is used as the link period, where $t$ is the current period and $|T|$ is the chosen length of the time window

## HALF SPLICE ON PUBLISHED INDICES (HASP)

A specific splicing technique that is often combined with a 25-month rolling time window (25-HASP). With the 25-HASP, the annual rate of changes of the published index correspond to the annual rate of change of the latest calculated multilateral index

## IDENTITY TEST

A test under the axiomatic approach that requires that, if each and every price remains unchanged between the two periods, then the price index must equal unity. This is sometimes referred to as the 'strong' identity test. A weaker version of the identity test requires that, if all prices, and all quantities, are equal in the two comparison periods, then the price index must equal identity. The identity test is also related to the multiperiod identity test

## LOW-SALES FILTER

A procedure that identifies and excludes products with relatively small expenditures from a data set. It is not recommended to apply a low-sales filter with multilateral methods, as removing products with relatively small expenditures in a preprocessing step could bias results

## MATCH-ADJUSTED *R*-SQUARED (MARS)

A method to help price index compilers identify homogeneous products. If homogeneous products are specified too broadly, there may be the risk of a unit value bias. If they are defined too tightly, there may be the risk that relaunches are not captured. The MARS method provides a score that helps to strike a balance between these two competing requirements

## MEAN SPLICE

A splicing technique in which all the overlap periods are used as link periods

## MOVEMENT SPLICE

A splicing technique in which the previous period is used on a rolling basis as the link period

## MULTILATERAL INDICES

A category of index number formulas that measure the aggregate price change between two periods based on prices observed in multiple periods including the two comparison periods. For price comparisons over time, multilateral index formulas are mainly used with scanner data, and their main advantage is to avoid chain drift associated with the use of chained bilateral price index formulas in dealing with changing and dynamic consumption universes

## MULTIPERIOD IDENTITY TEST

A test based on the calculation of price changes over two consecutive periods, whereby an artificial final period is introduced in which the prices and quantities are the same as the prices and quantities of the first period. The test asks whether its result is equal to 1. Chain drift occurs when an index does not return to unity after this test is applied

## OUTLIER FILTER

A procedure applied to data in which an observation is excluded if a price or quantity change, when compared with the previous or any earlier month, is considered to be unusually high or low. Extreme changes may be an indication of data coding errors and other mistakes in the data. Thresholds for identifying outliers should be set very carefully, as they may have an impact on available information and on index results

## RELAUNCH PROBLEM

A situation in which products of the same quality are not matched because the level of product homogeneity is defined too tightly based on GTIN or other information. This situation, if not corrected, may lead to the price index missing out on important price changes associated with relaunches

## ROLLING TIME WINDOW

An approach in which the time window in which the multilateral method is calculated is shifted forward by 1 month. With this approach, the length of the time window is kept constant

## SPLICING

A technique (also known as linking) that consists of linking together two consecutive sequences of price indices that overlap in one or more periods. In the context of the calculation of multilateral index numbers, each time a new time window is used, the previously calculated indices may change. Splicing techniques are used to link the latest multilateral index to previous results in order to avoid revisions of published results

## STOCK-KEEPING UNIT (SKU)

A product identifier that, in contrast to the GTIN, is not regulated or standardised, as SKUs can be defined at the vendor or company level. When available, they may constitute an additional layer of granularity in the data

## TIME REVERSIBILITY

An index number property in which the index between periods *a* and *b* is equal to the inverse of the same index between periods *b* and *a*

## TIME WINDOW

The sequence of consecutive time periods over which the multilateral price index formula is calculated. The length of the window can be of an extensible nature (i.e. expanding time window) or of a fixed and rolling nature (i.e. rolling time window). Short lengths for the time window may lead to the compilation of unstable results and may not solve chain-drift problems. The longer the time window, the more data from the past will impact the current-month compilations

## TRANSITIVITY

An index number property in which an index that compares periods *a* and *b* indirectly through period *c* is identical to one that compares periods *a* and *b* directly. Transitivity, together with time reversibility, is one of the two index number properties considered in the implementing HICP regulation. While transitivity and identity imply time reversibility, the reverse is not true

## UNIT VALUE

The total value of the turnover divided by the sum of the quantities. Scanner data sets usually provide unit values, not individual prices. If unit values are not provided at a level at which products are homogeneous, their use in the compilation of price indices may give rise to unit value bias

## WEIGHTED TIME–PRODUCT DUMMY (WTPD)

A regression-based approach to provide multilateral price indices. It consists of running a regression using the WLS estimator, which includes dummy variables for the products and time periods that belong to the time window under consideration. The method can be alternatively formulated as a system of equations that is similar to the system of equations of the GK

## WINDOW SPLICE

A splicing technique in which the period $t - T + 1$ is used as the link period, where $t$ is the current period of the index and $|T|$ is the length of the time window

# Annex

## Research agenda for multilateral methods

### 1) METHODS

- While this guide focuses on three specific multilateral methods, other multilateral methods have been proposed that could be further investigated and that may have additional benefits. Examples include the SPQ method, the GEKS–Walsh, weighted variants of the GEKS and the Gerardi method.
- A methodology for calculating a HICP at constant tax rates with multilateral methods has been developed. Further empirical examples of this methodology and of the HICP-CT would be beneficial.

### 2) BENCHMARKING

- More research is needed to select the optimal window length and splicing method. There are trade-offs between transitivity and characteristicity, and these trade-offs, or other measurement targets, need to be formalised and assessed.
- There are different ways to calculate impacts and contributions in the context of multilateral methods. More work is needed to analyse and validate results obtained with multilateral methods.

### 3) IMPUTATIONS AND QUALITY ADJUSTMENT

- Prices and quantities can be imputed when data that go back far enough are not available to apply multilateral methods, or when data transmission fails. The role of such imputations needs to be further clarified and understood.
- If product characteristics are available, imputations can be based on hedonic functions. More work is needed on applying hedonic models to multilateral methods.
- The specification of the individual product still raises some questions because of its potentially significant impact on resulting indices. More guidance is needed to specify the individual product in the context of multilateral methods.
- Further clarifications are required on the treatment of seasonal products with multilateral methods. Both seasonal imputation and seasonal weights methods should be explored.

### 4) IMPLEMENTATION

- Against the background of a multitude of technical choices to be made when implementing a multilateral method, frameworks and decision aids should be developed for selecting a specific variant. Emphasis should be put on the harmonisation aspect and on the comparability of results.
- When implementing a multilateral method in the HICP, the indices obtained with multilateral methods must be combined with indices obtained with other index compilation methods. More work is needed on data integration and on consistency between methods.

- While multilateral methods have been examined in this guide in a generic way, further studies should be conducted on product-specific examples and implementation issues.

# References


Australian Bureau of Statistics (2016), *Making greater use of transactions data to compile the Consumer Price Index, Australia*, information paper, Australian Bureau of Statistics, Canberra (https://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/6401.0.60.003Explanatory%20Notes12016?OpenDocument).

Balk, B. M. (2001), 'Aggregation methods in international comparisons – what have we learned?', paper presented at the joint World Bank–OECD seminar on purchasing power parities, Washington DC, 30 January to 2 February.

Balk, B. M. (2005), 'Price indexes for elementary aggregates – the sampling approach', *Journal of Official Statistics*, Vol. 21, No 4, pp. 675–699.

Bertolotto, M. (2019), 'Online price index with product replacement – the closest-match approach', paper presented at the 16th meeting of the Ottawa Group, Rio de Janeiro, Brazil, 8–10 May.

Białek, J. (2021), 'PriceIndices – a new R package for bilateral and multilateral price index calculations', *Statistika – Statistics and Economy Journal*, Vol. 2/2021, pp. 122–141.

Białek, J. and Beręsewicz, M. (2021), 'Scanner data in inflation measurement: from raw data to price indices', *Statistical Journal of the IAOS*, Vol. 37, pp. 1315–1336.

Białek, J. and Bobel, A. (2019), 'Comparison of price index methods for CPI measurement using scanner data', paper presented at the 16th meeting of the Ottawa Group, Rio de Janeiro, Brazil, 8–10 May.

Caves, D. W., Christensen, L. R. and Diewert, W. E. (1982), 'Multilateral comparisons of output, input, and productivity using superlative index numbers', *Economic Journal*, Vol. 92, No 365, pp. 73–86.

Chessa, A. G. (2016), 'A new methodology for processing scanner data in the Dutch CPI', *EURONA – Eurostat Review on National Accounts and Macroeconomic Indicators*, No 1/2016, pp. 49–70.

Chessa, A. G. (2017), 'A comparison of price index methods for scanner data', paper presented at the 15th meeting of the Ottawa Group, Eltville am Rhein, Germany, 10–12 May.

Chessa, A. G. (2019a), 'MARS: a method for defining products and linking barcodes of item relaunches', paper presented at the New Techniques and Technologies for Statistics 2019 conference, Brussels, Belgium, 12–14 March.

Chessa, A. G. (2019b), 'A comparison of index extension methods for multilateral methods', paper presented at the 16th meeting of the Ottawa Group, Rio de Janeiro, Brazil, 8–10 May.

Chessa, A. G. (2021), 'Extension of multilateral index series over time: analysis and comparison of methods', paper presented at the UN Economic Commission for Europe meeting of the Group of Experts on Consumer Price Indices, 2–4, 9 and 10 June.

Chessa, A. G. and Griffioen, R. (2019), 'Comparing price indices of clothing and footwear for scanner data and web scraped data', *Economie et Statistique / Economics and Statistics*, No 509, Big Data and Statistics Part 2 – Big Data in the Consumer Price Index, pp. 49–68.

Dalèn, J. (2017), 'Unit values in scanner data – some operational issues', paper presented at the 15th meeting of the Ottawa Group, Eltville am Rhein, Germany, 10–12 May.

de Haan, J. and Daalmans, J. (2019), 'Scanner data in the CPI – the imputation CCDI index revisited', paper presented at the 16th meeting of the Ottawa Group, Rio de Janeiro, Brazil, 8–10 May.

de Haan, J. and Krsinich, F. (2014), 'Scanner data and the treatment of quality change in nonrevisable price indexes', *Journal of Business and Economic Statistics*, Vol. 32, No 3, pp. 341–358.

de Haan, J. and Krsinich, F. (2018), 'Time dummy hedonic and quality-adjusted unit value indexes – do they really differ?', *Review of Income and Wealth*, Vol. 64, No 4, pp. 757–776.

Diewert, W. E. (1999), 'Axiomatic and economic approaches to international comparisons', in Heston, A. and Lipsey, R. E. (eds), *International and interarea comparisons of income, output, and prices*, Studies in Income and Wealth, Vol. 61, University of Chicago Press, Chicago, pp. 13–87.

Diewert, W.E. (2020), 'The Chain Drift Problem and Multilateral Indexes', Discussion Paper 20-07, Vancouver School of Economics, The University of British Columbia.

Diewert, W. E. and Fox, K. (2020), 'Substitution bias in multilateral methods for CPI construction', *Journal of Business and Economic Statistics*, Vol. 40, No 1, pp. 355–369.

Diewert, W. E. and von der Lippe, P. (2010), 'Notes in unit value index bias', *Jahrbücher für Nationalökonomie und Statistik*, Vol. 230, pp. 690–708.

Diewert, W. E., Fox, K. and de Haan, J. (2016), 'A newly identified source of potential CPI bias – weekly versus monthly unit value price indexes', *Economics Letters*, Vol. 141, pp. 169–172.

Eurostat (2017), *Practical Guide for Processing Supermarket Scanner Data* (https://circabc.europa.eu/sd/a/8e1333df-ca16-40fc-bc6a-1ce1be37247c/Practical-Guide-Supermarket-Scanner-Data-September-2017.pdf).

Eurostat (2018), *Harmonised Index of Consumer Prices (HICP) Methodological Manual – 2018 edition* (https://ec.europa.eu/eurostat/en/web/products-manuals-and-guidelines/-/KS-GQ-17-015).

Henn, K., Islam, C. G., Schwind, P. and Wieland E. (2019), 'Measuring price dynamics of package holidays with transaction data', *EURONA – Eurostat Review on National Accounts and Macroeconomic Indicators*, No 2/2019, pp. 95–132.

ILO (International Labour Organization), IMF (International Monetary Fund), OECD (Organisation for Economic Co-operation and Development), Eurostat, UNECE (UN Economic Commission for Europe) and World Bank (2020), *Consumer Price Index Manual – Concepts and methods* (https://www.imf.org/-/media/Files/Data/CPI/cpi-manual-concepts-and-methods.ashx).

Institut National de la Statistique et des Études Économiques (INSEE) (2020), *Using scanner data to calculate the Consumer Price Index – The impact on the CPI* (https://www.insee.fr/en/statistiques/documentation/note%20donn%C3%A9es%20de%20caisses%202020_en.pdf).

International Monetary Fund (2019), *Residential Property Price Index Practical Compilation Guide*, draft version (https://www.imf.org/en/Data/Statistics/RPPI-guide).

Istituto Nazionale di Statistica (ISTAT) (2020), *Consumer Prices Indices – Basket, weight structure and survey updates* (https://www.istat.it/it/files/2020/02/Basket_2020.pdf).

Ivancic, L. and Fox, K. (2013), 'Understanding price variation across stores and supermarket chains – some implications for CPI aggregation methods', *Review of Income and Wealth*, Vol. 59, No 4, pp. 629–647.

Ivancic, L., Diewert, W. E. and Fox, K. J. (2011), 'Scanner data, time aggregation and the construction of price indexes', *Journal of Econometrics*, Vol. 161, No 1, pp. 24–35.

Lamboray, C. (2021), 'Index compilation techniques for scanner data: an overview', paper presented at the UNECE meeting of the Group of Experts on Consumer Price Indices, 2–4, 9 and 10 June.

Mikkelsen, J. (2012), 'Extending the Danish CPI with scanner data – a stepwise analysis', paper presented at the UNECE/ILO Meeting of the Group of Experts on Consumer Price Indices, Geneva, Switzerland, 30 May–1 June.

Müller, R. (2010), 'Scanner data in the Swiss CPI – an alternative to price collection in the field', paper presented at the UNECE/ILO meeting of the Group of Experts on Consumer Price Indices, Geneva, Switzerland, 10–12 May.

National Statistics Institute (INE) (2020), *Calculation methodology for the scanner data in the CPI and HICP* (https://www.ine.es/en/metodologia/t25/ipc_scanner_data_en.pdf).

Republic of Slovenia Statistical Office (2018), *Changes in Computing Inflation in 2018* (https://www.stat.si/StatWeb/File/DocSysFile/9842/changes_%20in_computing_inflation_2018.pdf).

Silver, M. (2009), 'An index number formula problem – the aggregation of broadly comparable items', IMF Working Paper 09/19, IMF, Washington, DC.

Ståhl, O. (2019), 'On the operational definition of homogeneous products in transaction data', paper presented at the New Techniques and Technologies for Statistics 2019 conference, Brussels, Belgium, 12–14 March.

Vartia, Y., Suoperä, A., Nieminen, K. and Markkanen, H. (2018), 'Chain error as a function of seasonal variation', paper presented at the 16th meeting of the Ottawa Group, Rio de Janeiro, Brazil, 8–10 May.

Webster, M. and Tarnow-Mordi, R. C. (2019), 'Decomposing multilateral price indexes into the contributions of individual commodities', *Journal of Official Statistics*, Vol. 35, No 2, pp. 461–486 (http://dx.doi.org/10.2478/JOS-2019-0020).

## Getting in touch with the EU

***In person***

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

***On the phone or by email***

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696 or
- by email via: https://europa.eu/european-union/contact_en

## Finding information about the EU

***Online***

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

***EU publications***

You can download or order free and priced EU publications at: https://op.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

***EU law and related documents***

For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex at: https://eur-lex.europa.eu

***Open data from the EU***

The EU Open Data Portal (https://data.europa.eu/en) provides access to datasets from the EU. Data can be downloaded and reused for free, for both commercial and non-commercial purposes.

KS-GQ-21-020-EN-N

# Guide on Multilateral Methods in the Harmonised Index of Consumer Prices

Scanner data are increasingly used in the calculation of the Harmonised Index of Consumer Prices (HICP). Multilateral methods are index compilation methods that have certain advantages over more traditional methods when scanner data are used.

Multilateral methods have been applied for many years for making price comparisons across space (e.g. between different countries or regions), and they have been adapted to make comparisons across time. The aim of this guide is to promote the use of multilateral methods for scanner data and to support countries in understanding and implementing multilateral methods in the context of the HICP. Guidance and recommendations are provided on the various steps needed to implement multilateral methods. This guide may be useful for compilers and users to become familiar with these methods, and for statistical offices to decide how to process scanner data.

**For more information**
**https://ec.europa.eu/eurostat/**

Publications Office
of the European Union

ISBN 978-92-76-44354-4