EUROPEAN COMMISSION
EUROSTAT

Directorate C: Macro-economic statistics
Unit C-4: Price statistics. Purchasing power parities. Housing statistics

Index Compilation Techniques for Scanner Data

An overview

Paper prepared by Claude Lamboray
(claude.lamboray@ec.europa.eu)
for the Group of Experts on Consumer Price Indices
UNECE , Online meeting, June 2021

Index Compilation Techniques for Scanner Data

An Overview

Abstract

When calculating a price index with scanner data, decisions need first to be made on the definition of the individual product that is priced over time. In a second step, the prices of the individual products are aggregated using a specific compilation method. We propose to classify compilation methods according to the following criteria: (i) weights and product universe; (ii) the properties of the price index; (iii) the strategy that specifies how reference periods are updated and series are linked; and (iv) the sampling method that may be applied to restrict the individual products that enter the calculations. We also discuss how elementary aggregation is regulated in the HICP legal framework and provide an overview of the methods used by European countries. Empirically, we test the main bilateral and multilateral methods by applying them to six product categories of the publicly available Dominick's Finer Food data set.

Keywords: Scanner Data, Bilateral Methods, Multilateral Methods, Harmonised Index of Consumer Prices

JEL classification: C43, E31

1. INTRODUCTION

Scanner data sets are increasingly available to National Statistical Institutes (NSIs). However, the compilation of price indices from such new data sources is not straightforward. Scanner datasets are large datasets often characterised by dynamic product assortments. In addition, the datasets contain information not only on prices but also on how often the products were sold. The aim of this paper is to review index compilation techniques that are applied in the Harmonised Index of Consumer Prices (HICP) and to describe the state of the art in Europe in terms of implementation of various bilateral and multilateral methods. This paper builds extensively on work conducted in an HICP Task Force on Multilateral Methods. This Task Force was established in order to exchange information on multilateral methods and to draft a guide that explains how these methods should be implemented in the context of the HICP([1]).

Before any index compilation method can be applied, a decision must be made on the definition of the individual product. The individual product is the unit for which a price is tracked over time. It is the result of a unit value aggregation over time, outlets, and item codes. The average prices and aggregated quantities of the resulting individual products will be the input of any index compilation method.

In practice, there are many index compilation methods that can be applied to scanner data. With a bilateral price index, prices of products in two periods are compared. A bilateral price index may, or may not, use weights. Bilateral indices can be applied with respect to a fixed price reference

([1]) See Eurostat (2020). Further information on the work of the Task Force can be accessed here.

period that is updated, for example, once per year. While moving away from that period, the overlap of products declines. A way of increasing the overlap of products is to update the base period every month and to chain 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 anymore a reasonable price change over longer periods.

Multilateral methods have been found to be a solution to these problems. They take into account all the products that are available in the different periods. They allow to explicitly weigh each individual product according to its importance in each period. Finally, they aim to avoid the chain-drift problems encountered with chained bilateral indices. When implementing a multilateral approach, choices must be made concerning the multilateral index formula. The length of the time window over which the multilateral index formula is applied must be determined. A decision must also be made on how to splice the indices compiled over different time windows. Different splicing techniques have been proposed so that the published indices do not need to be revised.

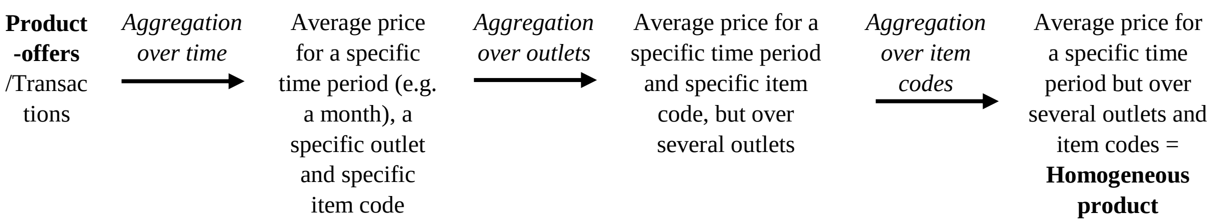
This paper is organised as follows. In Section 2, we discuss conceptual and practical arguments to specify the individual product. Section 3 presents a framework to classify the various compilations methods. Section 4 introduces some compilation methods and explains how they fit into the framework. Section 5 discusses how elementary aggregation is regulated in the HICP. An overview of country practices is given in Section 6. The results from simulations can be found in Section 7. Finally, we conclude by highlighting some challenges and open questions that arise from this analysis.

2. WHAT IS THE INDIVIDUAL PRODUCT TO BE PRICED?

In field price collection, prices for product-offers ([2]) are usually 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 unit values (sales divided by number of units sold) of some item codes.

When specifying individual products, one needs to consider the following three dimensions: (1) the time dimension; (2) the outlet dimension; and (3) the product dimension. At each level, an average price (unit value) is calculated, as Figure 1 highlights. The unit value price is calculated by dividing total expenditure by the related total quantity. 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 1: From product-offers to homogeneous products



([2]) 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.*

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 depend very much on the data characteristics (data may only be supplied at a more aggregated level), and on the product category (for example homogenous products may be especially relevant for clothing).

Conceptually, the main principle is that transactions/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 amongst which there are no significant quality differences ([3]). Quality differences must be evaluated with respect to the already mentioned time, outlet and product dimensions. If transactions are combined which are of different quality, unit value bias may occur.

In practice, the specification of individual products can be made on the grounds ([4]) discussed below. Some examples of specifications for the individual product can be found in Table 1.

Table 1: Examples of specifications for the individual product

Time dimension	Outlet dimension	Product dimension
First 3 weeks of the month	Unit value price by retail chain	SKU
First 14 days of the month	Unit value price by chain and by region	Article code
First 2 weeks of the month	Unit value price by chain and type of shop	GTIN
First 3 weeks of the month	Unit value price by outlet type and by province	GTIN
Full month	Unit value price by outlet	GTIN
Full month	Unit value price by outlet	'Article' consisting of several GTINs of the same quality

Time dimensions

In principle, it is appropriate to calculate a unit value when an item code 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 as homogeneous for the purpose of price aggregation.

Depending on the data supply arrangements and the production and publication calendar, the individual product typically covers the two, three (or sometimes the four) first weeks of the month. It is important to cover as much as possible of the reference month. Diewert, Fox and de

([3]) 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.*

([4]) This material builds upon the initial discussion that can be found in Dalèn (2017).

Haan (2016) showed that aggregation over only one week of the month can be upward biased compared to aggregation over the full month.

Outlet dimension

One 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 homogenous as long as the price of the product is found to be systematically the same across outlets. Outlets may also be grouped on the basis of some geographical criteria if for instance regional price indices are compiled. One reason for aggregating across outlets is that it significantly reduces the number of individual products which 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 nowadays be usually overcome.

The impact of aggregating (or not) across outlets is an empirical matter that can be assessed. One could 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. An extensive discussion on whether aggregation across outlets is warranted or not can be found in Ivancic and Fox (2013).

Product dimension

The Global Trade Identification Number (GTIN) is often found to be the most granular product level available in the dataset. Besides GTINs, some retailers may use Stock Keeping Units (SKUs) which can be slightly more stable than GTINs. There may be other product identifiers in the data. For most supermarket products, such item codes (e.g. GTIN, or preferably SKU) available in the data are used as a product identifier.

Relaunches can be an issue when tracking tight product codes. Therefore it is preferable to develop some procedure that identifies relaunches and links together two item codes referring to the 'same' product. Relaunches may come with changes in the package size which should also be adjusted for in the calculation of the average price. For example, a chocolate bar previously consisted of 100 g and is now sold in a package of only 80 g.

Item codes could be grouped together to form homogeneous products. Chessa (2019) argues that, when constructing homogeneous products, trade-offs must be made between homogeneity and stability over time. If homogenous 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 MARS method can be used as a tool for finding a compromise between these two objectives. Instead of using product characteristics to form homogenous products and thereby increasing the matching over time, one could also work with narrowly defined individual products and directly integrate the product characteristics in the indices.

3. A FRAMEWORK FOR 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 the previously defined individual products. The practical implementation of a compilation method is not only a question of a specific price index formula. There are different strategies to implement a price index. We distinguish the compilation methods ([5]) 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 2. Also some specific examples of compilation methods are included in Figure 2. These methods ([6]) are further defined in Section 4.

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. Different product universes can result from this. A key distinction must be made between, on the one hand, approaches that are based on a static product universe with fixed weights, and on the other hand, approaches that are based on a dynamic product universe and/or on individual products that are weighted according to their importance in each period.

Index property

There are many properties (or tests) that a price index may or may not satisfy. In order to discriminate between the different compilations 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.

- *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 (re-sampling). This may happen every month, or less frequently, for example only once per year. There are trade-offs in designing an update strategy. A more frequent updating better takes into account a dynamic product universe. At the same time frequent chaining can lead to chain-drift.

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 one month. There are trade-offs. For instance, splicing of indices compiled over shorter time windows could lead to unstable result 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.

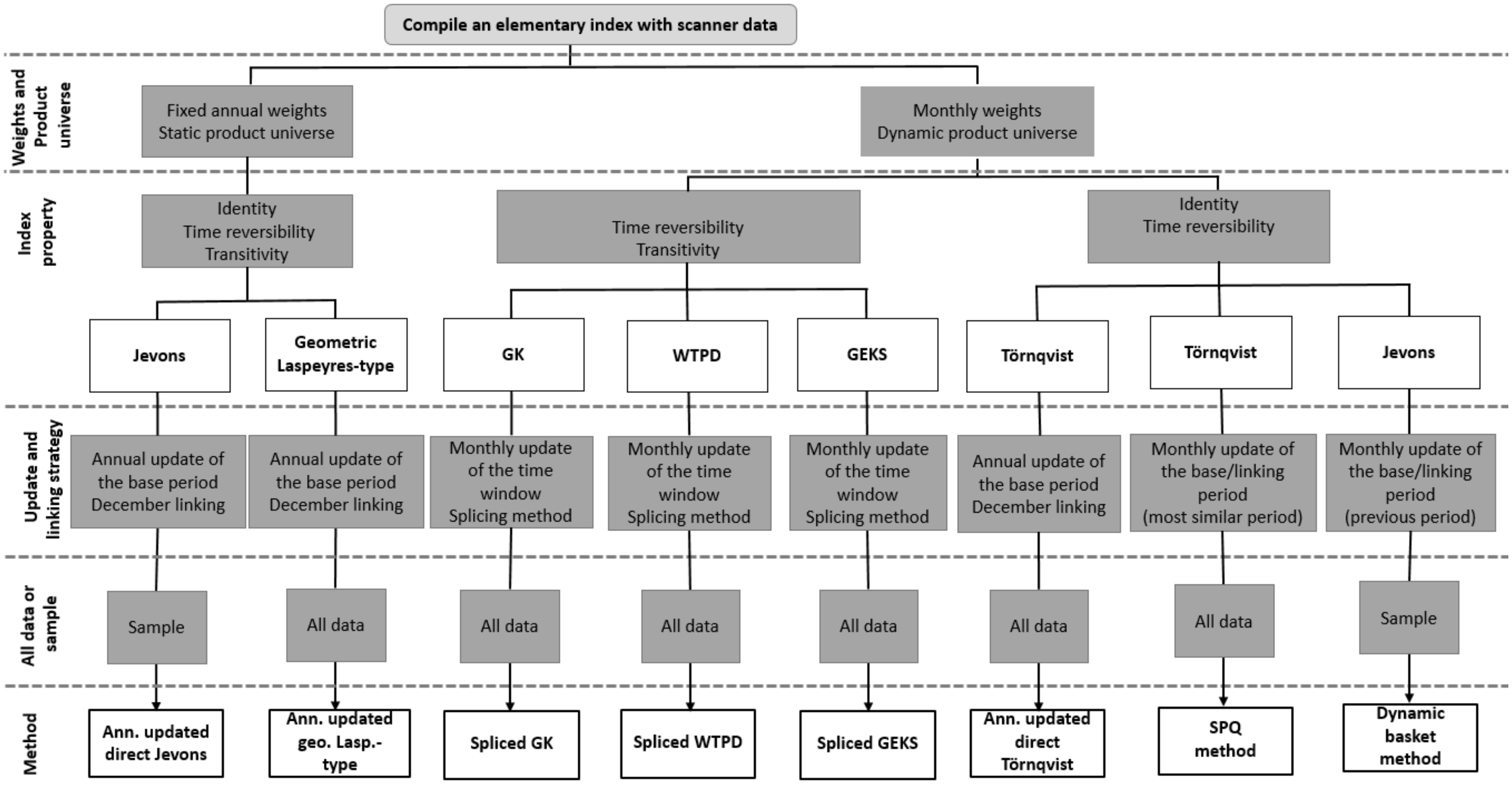
All data or sampling

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

([5]) A similar framework based on index formula, index base, reference universe and homogeneous product definition was formalised in Zhang, L. et al. (2019).

([6]) Linking Based on Relative Price and Quantity Similarity (SPQ method) will not be further examined in this paper. See Chapter 7 of the CPI Theory Manual (draft) for more information on this method.

Figure 2: A framework for price index compilation methods



4. EXAMPLES OF COMPILATION METHODS

4.1 Fixed weights and static product universe

In line with the fixed basket principle, one approach consists in pricing the same set of individual products (a 'basket') over time. In a geometric Laspeyres-type index, the weights attached to the products refer to some past period (e.g. the previous calendar year) and are kept fixed (e.g. for a year). Let b denote a period prior to the price reference period 0. We then have the following definition for a geometric Laspeyres-type ([7]) index:

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

New products can continuously be incorporated into the index compilation as one-to-one replacements of disappearing products. If a product initially available becomes missing, a replacement product is selected and brought into the index and a quality adjustment is performed. This ensures that the basket initially selected is kept fixed and remains representative. However, new individual products that are not used as replacements of disappearing products will be ignored. In practice, manual or automatic procedures must be designed that support the selection of the replacements ([8]).

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

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

The sample S is selected for the base period and priced over time. The sample S should be 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 (for example 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).

Both the geometric Laspeyres-type index and the Jevons index satisfy transitivity, under the condition that the underlying basket of individual products, and their weights, are kept constant over time. They also pass the identity test, and, as a consequence, also satisfy time reversibility.

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.

([7]) In the CPI Manual, this index is referred to as a Geometric Young.

([8]) An alternative would be to compile maximum overlap geometric Laspeyres-type indices. These indices are based on the matched products available in both the price reference and current periods. Maximum overlap indices would not be transitive anymore, but they would still satisfy identify and time reversibility.

4.2 Variable weights and/or dynamic product universe

4.2.1. Multilateral indices that satisfy transitivity

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 (see for example Ivancic, Diewert, and Fox (2011)). Multilateral methods are also presented in Chapter 10 of the CPI manual (ILO et al. (2020)). We focus here on the three most studied methods.

- The *Gini-Eltetö-Köves-Szulc (GEKS)* index is based on a bilateral index which is used to compare any two periods belonging to a time window (see paragraphs 10.80-10.81 of the CPI Manual). The GEKS index is obtained by averaging the bilateral indices. The initial GEKS method relies on a Fisher index. Very often, the GEKS approach is applied with an underlying bilateral Törnqvist index. This multilateral index is also known as CCDI (Caves, Christensen and Diewert, 1982).

- The *Weighted Time Product Dummy (WTPD)* index consists in running a regression which includes dummy variables for the products and time periods that belong to the time window (see paragraphs 10.85-10.88 of the CPI Manual). Each observation is weighted according to its share in a given period t. The regression is then estimated using the Weighted Least Squares (WLS) estimator.

- The *Geary-Khamis (GK)* index is obtained by solving a system of equations (see paragraphs 10.82-10.84 of the CPI Manual). This index can be seen as an implicit price index defined as a value index divided by a quantity index. This quantity index has the form of a basket index where the quantities in the two periods are compared using a fixed set of reference prices. These reference prices correspond to an average deflated price and are implicitly determined by the data.

All three multilateral indices are transitive (and they also satisfy time reversibility), but they do not satisfy the identity test. This means that the index may change between two periods although all prices in these two periods are identical.

A decision must be made regarding the length of the time window over which the multilateral index is applied. The multilateral index is transitive with respect to a given time window. Each time a new time window is used, the previously calculated indices may change. Therefore, splicing techniques must be used which link the latest multilateral index onto the previous results in order to avoid revisions of already published results. Transitivity is not satisfied anymore for the indices that are eventually published. Therefore, some degree of chain drift cannot be fully excluded for the published (spliced) indices.

4.2.2. Bilateral indices that satisfy identity

Fixed base Törnqvist

There are several options for symmetrically weighted price indices. The Törnqvist index is often used in practice. A maximum overlap Törnqvist index is defined as follows:

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

Such an index satisfies identity and time reversibility. However, it is not transitive. Chaining of a Törnqvist index can lead to the 'chain drift' problem. Therefore, a chained Törnqvist index should best be avoided. By directly comparing the current period to some fixed base period, this problem is avoided because no chaining is involved. Moreover, variable weights are attached to the individual products which allows capturing the economic importance of each individual product in each period.

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 only available in certain months of the year. With one-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 to 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.

Dynamic basket method

An approach referred to as the 'dynamic basket method' in the practical guide for processing supermarket scanner data (Eurostat, 2017) consists in updating the base period each month. Every month, the set of individual products that enter the index compilation is re-sampled. In practice, cut-off sampling is applied which selects the most sold products in two consecutive periods. A Jevons price index is compiled between the two periods taking into account only the selected individual products. The final price index is obtained by chaining the month-on-month Jevons indices.

Note that the Jevons index calculated between two successive month does satisfy identity and time reversibility. These properties are however not satisfied anymore for the chained indices. The chained Jevons indices calculated over different baskets are not transitive anymore.

The dynamic basket method was initially proposed in Van der Grient , H. and de Haan, J. (2010). This approach can be seen as a pragmatic compromise: on the one hand, it allows updating the basket 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 only implicitly used 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 to a chained Törnqvist index, the likelihood of chain drift is significantly reduced, it cannot be fully excluded. Some downward bias can appear in particular when products exit the sample at reduced prices. In order to mitigate this problem, a dumping filter can be implemented which excludes from the calculations those individual products that exhibit large decreases in both prices and quantities sold.

5. HICP LEGAL FRAMEWORK

Elementary aggregation was initially defined in Annex 2 of Regulation (EC) No 1749/96. This regulation focused on the Jevons and Dutot to compile an index for an elementary aggregate. This regulation was drafted well before new data sources (for example transaction data, web scraped data) became widely available NSIs.

The HICP legal framework has since been updated and modernised. A new framework regulation was finalised in 2016 ([9]), followed by a new implementing regulation in 2020 ([10]). These two regulations replace all previous legal acts. One of the objectives of the new implementing regulation was to adapt the provisions to the use of new data sources. It was decided that the regulation should define the conditions that the compilation method must satisfy.

Conceptually, the elementary aggregate is understood as the lowest level of a Laspeyres-type index ([11]). The aggregation from an elementary aggregate upwards is thus based on the standard higher-level aggregation procedures used in the HICP. The objective here is to cover compilation of the elementary price index, that is the index for an elementary aggregate ([12]). Article 12(1) of Regulation (EU) 2020/1148 frames how elementary price indices can be obtained:

> *The prices of individual products shall be aggregated to obtain elementary price indices using either of the following options:*
>
> *(a) an index formula that ensures transitivity. The price index of prior periods shall not be revised when using transitive index formulae; or*
>
> *(b) an index formula that ensures time reversibility and compares the prices of individual products in the current period with the prices of those products in the base period. The base period shall not be changed frequently if such change leads to significant violation of the transitivity principle.*

The indices described in Figure 2 comply with the principles set out in this Article.

In relation to paragraph (a), as discussed in section 4.1, indices, such as Jevons or Dutot, or their weighted counterparts, satisfy transitivity, under the condition that the underlying basket of individual products is kept constant over time. As discussed in section 4.2.1, multilateral indices, such as GEKS, Time Product Dummy or Geary-Khamis all satisfy transitivity. When applying a multilateral index in practice, one needs to decide how to combine indices compiled over different (rolling or expanding) time windows. It is important that the already published indices are not revised because of the use of a transitive index formula.

In relation to paragraph (b), as discussed in section 4.2.2, the Törnqvist index satisfies time reversibility. Such an index can be implemented as a fixed base index. Where these indices are used, the update of the base period must be implemented in a way not to significantly violate transitivity. As discussed in section 4.2.2, the Jevons index also satisfy time reversibility. These indices can be applied by updating the base period every month. This is only allowed to the

([9]) Regulation (EU) 2016/792 of the European Parliament and of the Council of 11 May 2016 on harmonised indices of consumer prices and the house price index, and repealing Council Regulation (EC) No 2494/95

([10]) Commission Implementing Regulation (EU) 2020/1148 of 31 July 2020 laying down the methodological and technical specifications in accordance with Regulation (EU) 2016/792 of the European Parliament and of the Council as regards harmonised indices of consumer prices and the house price index

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

([12]) 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* extent that the monthly update of the base period does not lead to a significantly violation of the transitivity criterion.

6. COUNTRY OVERVIEW

Tables 2 and 3 give an overview of methods used in selected EU and EFTA countries according to the criteria set out in this paper. For each country, only the most used method is reported. This method is typically applied to supermarket scanner data. Countries may use other methods for specific product categories for which scanner data is available. There are two groups of countries. One the one hand, there are countries that follow a fixed basket approach based on fixed weights from the past. On the other hand there are countries that have implemented a variable weight approach. Four countries have already implemented a multilateral method.

Table 2: Countries with fixed weights

Country	**Price index**	**Update and linking**	**All the data or sample**	***Reference***
Switzerland	Jevons	Direct index, resampling once per year	Sample of products	*Müller, R. (2010)*
Denmark	Weighted Jevons	Chained index, resampling once per year	Sample of products	*Mikkelsen, J. (2012)*
France	Geometric Laspeyres	Direct index, resampling once per year	All the data	*INSEE (2020)*
Iceland	Weighted Jevons	Direct index, resampling once per year	Sample of products	*Guðmundsdóttir, H. E. (2017)*
Sweden	Weighted Jevons	Direct index, resampling once per year	Sample of products	*Norberg, A. (2011).*

Table 3: Countries with variable weights

Country	**Price index**	**Update and linking strategy**	**All the data or sample**	***Reference***
Belgium	GEKS-Törnqvist (CCDI)	25 Month half splice on published index	All the data	*Van Loon, K. (2020)*
Finland	Törnqvist	Direct index, resampling once per year	All the data	*Vartia, Y. et al. (2018).*
Italy	Jevons	Chained index, resampling every month	Cut-off sampling	*ISTAT (2020)*
Luxembourg	GEKS-Törnqvist (CCDI)	25 Month half splice on published index	All the data	*Radjabov B., and Ferring M. (2021)*
Norway	GEKS-Törnqvist (CCDI)	25 Month half splice on published index	All the data	*Johansen, I. and Nygaard, R. (2021)*
Netherlands	Geary-Khamis (QU-method)	Fixed Base Expanding Window	All the data	*Chessa, A.G. (2016)*
Slovenia	Jevons	Chained index, resampling every month	Cut-off sampling	*Republic of Slovenia Statistical Office (2018)*
Spain	Jevons	Chained index, resampling every month	Cut-off sampling	*INE (2020)*

7. SIMULATIONS

We have implemented different methods ([13]) using the publicly available Dominick's scanner dataset. This dataset is presented in detail in Mehrhoff (2019). The weekly data has been transformed into monthly data by assigning each week to the month in which the week starts. We use data that covers 88 months (January 1990-April 1997). The data has been aggregated across all stores up to the chain level. The item code, and not the Universal Product Code (UPC), is used to define the individual product. We focus here on six categories:

- Bottled Juices (BJC),
- Front-end-candies (FEC),
- Fabric Softeners (FSF),
- Laundry Detergents (LND),
- Refrigerated Juices (RFJ),
- Canned Tuna (TNA).

The following direct indices are calculated: Törnqvist, Geometric Laspeyres, Laspeyres and Jevons. The base period is set to the previous year. This base period is updated every year. Two consecutive series are chain-linked via the December month. In addition, a chained Jevons index

([13]) All the detailed indices presented in this section can be accessed at CIRCABC.

combined with cut-off sampling (dynamic basket method) is calculated. Finally, a chained Törnqvist index is also calculated.

Table 4 includes the average index number over the period December 1990 to April 1997 for different methods. Figure 3 shows indices for the product FSF. Note that the multilateral index in Table 4 and in Figure 3 corresponds to a GEKS price index calculated on a 25 month rolling time window with a half splice on published indices (GEKS 25-HASP). The starting period for the multilateral index is January 1990. For the direct indices, the first comparison is between the year 1990 and December 1990. To make them comparable, the indices have been rescaled to December 1990 equals 100.

The following lessons can be learnt from these simulations:

- The multilateral index is in general rather close to an annually chained direct Törnqvist index. This is in line with the theory exposed in section 8 in Diewert (2018).

- The chained Törnqvist index has a downward bias with respect to an annually chained direct Törnqvist index. A chained Törnqvist index is usually not recommended because of possible downward chain drift.

- The annually chained direct Geometric Laspeyres index lies above an annually chained direct Törnqvist index, although in a few cases it can also be the other way around. This is in line with the theory exposed in section 6 in in Diewert (2018).

- The annually chained direct Laspeyres index is always higher than the annually chained direct Geometric Laspeyres.

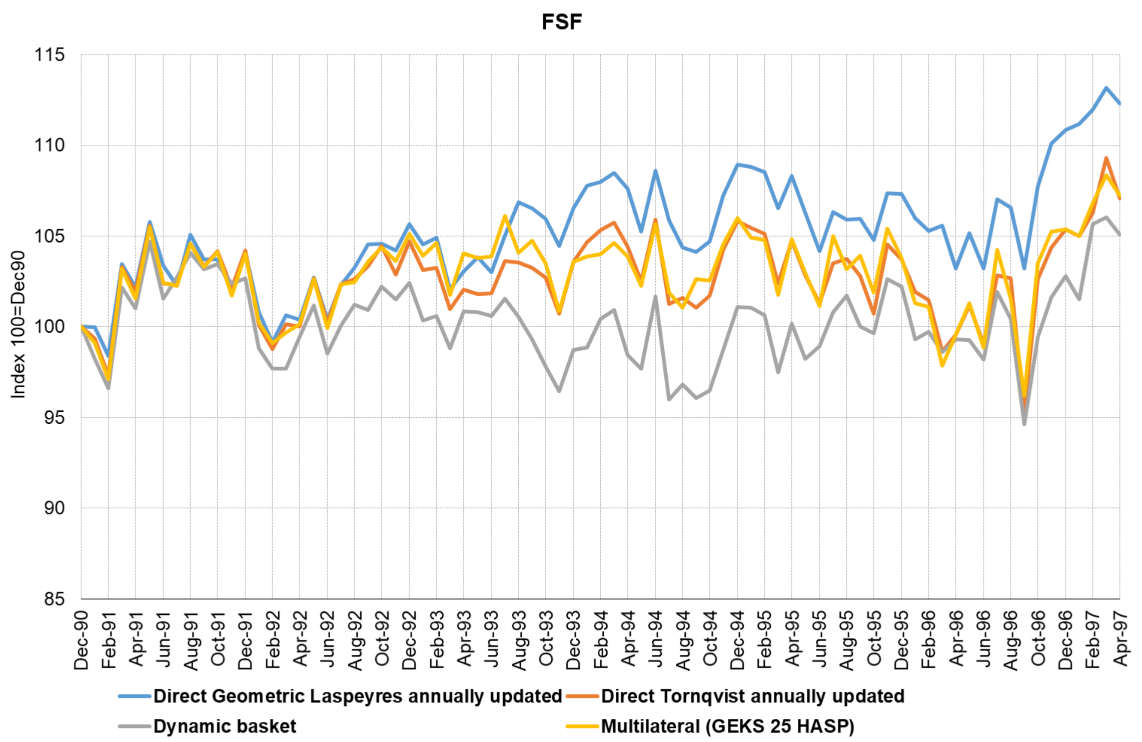
- Overall, the dynamic basket method is not that far from a multilateral, but there can still be some differences in some months. This is in line with findings in other studies. For instance, Chessa, A.G. (2018a) and Van Loon, K. and Roels, D. (2018) have also found that, at least for supermarket data, the 'dynamic approach', is often close to a multilateral index when compared at the more aggregated levels. However, differences may appear at the more detailed levels. We also noticed that the results of the dynamic approach can be very sensitive to the thresholds used for the low-sales filter the choice of the threshold. Białek and Beręsewicz (2020) have highlighted the same issue in their study. They also found that the dynamic basket method can lead to an index that sits slightly below an index obtained with a multilateral method.

- The annually chained direct Jevons looks less reliable and can sometimes be different from the other methods (for example for FEC, TNA or RFJ). This shows the importance of using weights in the index formula.

Table 4: Average index (Dec 1990 - April 1997), 100=Dec 1990

Product	Chained Törnqvist	Dynamic basket	Direct Törnqvist	GEKS 25 -HASP	Direct Geometric Laspeyres	Direct Laspeyres	Direct Jevons
BJC	103.1	106.7	103.5	104.6	105.4	106.2	105.2
FEC	113.4	114.7	114.8	114.8	115.3	116.0	111.7
FSF	99.3	100.3	102.7	102.9	105.2	106.9	100.8
LND	89.2	94.1	94.9	96.8	92.3	94.9	96.8
RFJ	74.3	85.7	86.6	86.8	88.3	89.5	94.3
TNA	92.6	97.6	97.5	97.5	97.7	99.8	105.5

Figure 3: Index (Dec 1990 - April 1997) for FSF, 100=Dec 1990



A multilateral method can be seen as a combination of an index formula, a window length and a splicing method. In total, 75 multilateral methods have been tested (3 index formulas, 3 rolling time windows with 8 splicing methods or 1 full time window):

- Index formula: GEKS-Törnqvist, Geary-Khamis (GK), Weighted Time Product Dummy (WTPD)

- Window length: rolling time windows of 13 months, 19 months, 25 months, 88 months (full time window)

- Splicing method: movement splice, window splice, window splice on published indices, half splice, half splice on published indices, mean splice, mean splice on published indices, December splice.

In order to summarise the results the various multilateral methods, the average index number over the period January 1990 to April 1997 for each of the 75 multilateral methods is plotted in Figure 4. As a benchmark, we highlight a GEKS price index calculated on a 25 month rolling time window with a half splice on published indices (25-HASP).

Results of the different methods are relatively consistent. This is largely in line with other empirical studies that compared different multilateral methods (for example ABS (2017), Van Loon, K. and Roels, D. (2018), Białek and Beręsewicz (2020)). The ONS (2020) also applied some multilateral methods to the Dominick's dataset. The following can be learnt from these simulations:

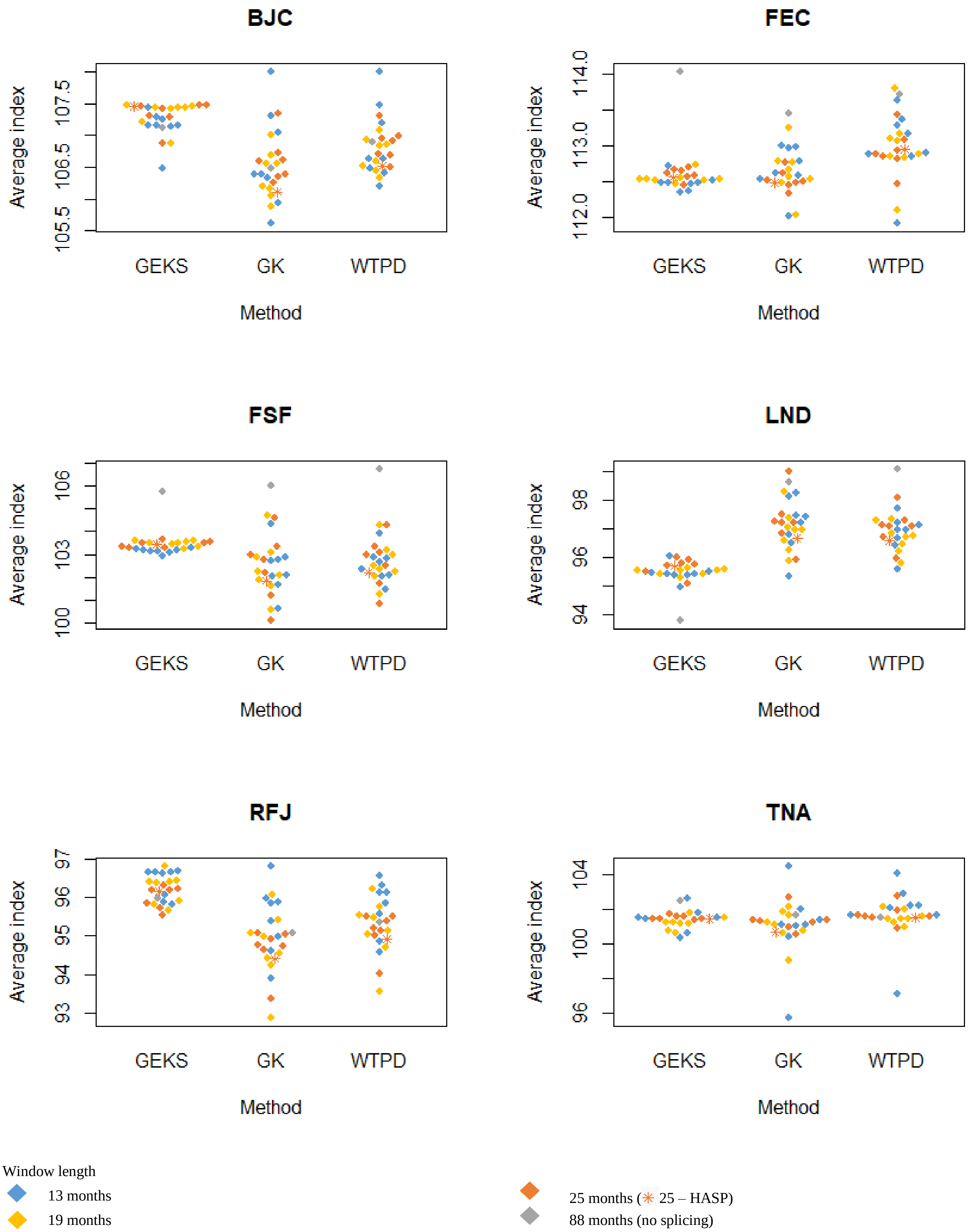
- The amplitude for the GEKS is less than for the GK and the WTPD. This shows that the GEKS is in general less sensitive to the choice of window length and splicing method than the GK or WTPD. The increased sensitivity of the GK and WTPD has also been noticed in other studies (for example Van Loon, K. and Roels, D. (2018)).

- In the case of the GEKS, the results of the different splicing method may not differ that much, conditional on a given window length. For the GK and WTPD, the results of different splicing method may differ more substantially, even for a given window length.

- The patterns for the GK and the WTPD look similar, but they are different from the pattern for the GEKS. It is not obvious to tell from these simulations if the GEKS is systematically lower or higher than the GK and WTPD.

- The results for extremely long time windows may look as outliers. This applies to all three index formulas (see for example FSF). This raises the question if multilateral indices compiled over long (full) time windows are suitable benchmarks, as it is commonly done in many studies. More work is needed to better understand the impact of window length.

Figure 4: Multilateral methods, average index (Jan 1990 - April 1997), 100=Jan 1990



8. Conclusions and open issues

There is no one method that can be applied to scanner data. In order to better understand and assess the different approaches that are implemented in practice, we clarified the specification of the individual product and classified the compilation methods according to different criteria.

In Europe, some countries use compilation techniques that are closer to the methods also used with prices collected in the field. Variants of direct indices based on fixed baskets and fixed weights from the past are implemented. In contrast, other countries adopt an approach with larger samples and with timely weights. The dynamic basket method is widely used, and countries are starting to switch to multilateral methods.

In order to reduce any representativity bias in the elementary indices, it is likely that using timely weights is better than using no weights, and that updating the samples more often in order to capture the product diversity over time is beneficial. From that perspective, multilateral methods seem to be a good compromise, although these methods also have their own technical challenges. Our simulations tend to indicate that direct Törnqvist indices and dynamic basket methods can approximate multilateral methods to some extent.

There is an increased complexity in the methods, from traditional small sample fixed basket methods to more advanced techniques such as multilateral methods. In practice, NSIs may apply a step-by-step approach. Over time, as the expertise is growing, the NSI may opt to implement methods that are methodologically more advanced. The increased complexity also creates challenges in the communication with users. The principles and methods underlying the inflation measurement should be communicated in a transparent manner.

The HICP implementing regulation covers many approaches to compile elementary indices. It also includes new methods that take advantage of the potential of scanner data. At the same time, it is acknowledged that there is a need to further develop recommendations and guidelines on the compilation techniques for scanner data and to enhance cross-country comparability. A dedicated HICP Task Force is currently developing guidelines on the use of multilateral methods in the HICP.

More generally, this raises the question of data integration. With the HICP becoming increasingly a multi-source statistics, more work is needed to understand how to best combine these different data sources and compilation methods.

References

ABS (2017). *Making Greater Use of Transactions Data to Compile the Consumer Price Index*. Information paper ABS.

Balk, B.M. (2005). *Price Indexes for Elementary Aggregates: The Sampling Approach*. Journal of Official Statistics, Vol. 21, No. 4, pp. 675–699.

Białek, J. and Beresewicz, M. (2020). *Scanner data in inflation measurement: from raw data to price indices*. arXiv:2005.11233.

Caves, D.W., Christensen, L.R. and Diewert, E.W. (1982). *Multilateral Comparisons of Output, Input, and Productivity using Superlative Index Numbers*. Economic Journal 92, 73-86.

Chicago Booth Kilts Center for Marketing (2013). *Dominick's Data Manual*. Version updated October 2018.

Chessa, A.G. (2016). *A new methodology for processing scanner data in the Dutch CPI*. EURONA, 1/2016, Eurostat, pp. 49-70.

Chessa, A.G. (2018). *A comparative study between the Geary-Khamis and chained Jevons method for Dutch supermarkets*. Paper presented at the UNECE Meeting of the Group of Experts on Consumer Price Indices, Geneva, 2018.

Chessa, A. G. (2019). *MARS: A method for defining products and linking barcodes of item relaunches*, paper presented at the 16th meeting of the Ottawa Group, Rio de Janeiro, Brazil.

Diewert, W.E., Fox, K.J. and de Haan, J. (2016). *A newly identified source of potential CPI bias: Weekly versus monthly unit value price indexes*, Economics Letters.

Diewert, W.E. (2018). *Scanner Data, Elementary Price Indexes and the Chain Drift Problem*. Discussion Paper 18-06, Vancouver School of Economics, University of British Columbia.

Dalèn, J. (2017). *Unit values in scanner data – some operational issues*. Paper presented at the 15th meeting of the Ottawa Group, Eltville, Germany.

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

Eurostat (2020). *Guide on the use of multilateral methods in the HICP* (Draft version November 2020). Annex of Item 4.5 of the Eurostat Price Statistics Working Group. Luxembourg, 25-26 November 2020.

Guðmundsdóttir, H. E. (2017). *How did Statistics Iceland start using scanner data?* Paper presented at the 15th Meeting of the Ottawa Group, Eltville, Germany.

ILO, IMF, OECD, Eurostat, UNECE, The World Bank (2020). *Consumer Price Index Manual: Concepts and Methods*.

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

INSEE (2020). *Using Scanner Data to Calculate the Consumer Price Index: The Impact on the CPI*. Available at: https://www.insee.fr/en/statistiques/documentation/note%20donn%C3%A9es%20de%20caisses%202020_en.pdf

ISTAT (2020). *Consumer prices indices. Basket, weight structure and survey updates*. Available at: https://www.istat.it/it/files/2020/02/Basket_2020.pdf

Ivancic, L., Diewert, E.W. and Fox, K.J. (2011). *Scanner data, time aggregation and the construction of price indexes*. Journal of Econometrics. Vol. 161(1), pages 24-35, March 2011.

Ivancic, L. and Fox, K.J. (2013). *Understanding price variation across stores and supermarket chains: some implications for CPI aggregation methods*. Review of Income and Wealth. Series 59, Number 4, December 2013.

Johansen, I. and Nygaard, R. (2021). *Ny beregningsmetode for prisindeks for matvarer og alkoholfrie drikkevarer i KPI*. Statistics Norway.

Norberg, A, Sammar, M. and Tongur, C. (2010). *A study on scanner data in the Swedish Consumer Price Index*. Paper presented at the 12th Meeting of the Ottawa Group, Wellington, New Zealand.

Mehrhoff, J., (2019*). Promoting the use of a publically available scanner dataset in price index research and for capacity building*. Paper presented at the 16th Meeting of the Ottawa Group, Rio de Janeiro, Brazil.

Mikkelsen, J. (2012). *Extending the Danish CPI with scanner data - A stepwise analysis*. Paper presented at the UNECE/ILO meeting of experts on Consumer Price Indices, Geneva, 2012.

Müller, R. (2010). *Scanner data in the Swiss CPI: An alternative to price collection in the fiel*d. Paper presented at the UNECE/ILO meeting of experts on Consumer Price Indices, Geneva, 2010.

ONS (2020). *New index number methods in consumer price statistics*. Research into the use of new index number methods to calculate price indices using web-scraped and scanner data. Available at: https://www.ons.gov.uk/economy/inflationandpriceindices/articles/newindexnumbermethodsinconsumerpricestatistics/2020-09-01

Radjabov B., Ferring, M. (2021).*The Implementation of a Multilateral Price Index Method for Scanner Data in the Luxembourg CPI*. STATEC Working Paper N° 121.

Republic of Slovenia Statistical Office (2018). *Changes in computing inflation in 2018*. Available at: https://www.stat.si/StatWeb/File/DocSysFile/9842/changes_%20in_computing_inflation_2018.pdf

Van Loon, K. and Roels, D. (2018). *Integrating big data in the Belgium CPI*, Belgium STATBEL. Paper presented at the UNECE/ILO Meeting of the Group of Experts on Consumer Price Indices, Geneva, 2018.

Van Loon, K. (2020). Van Loon, *Scanner data and web scraping in the Belgian CPI*. Presentation available here: https://www.nationalacademies.org/event/10-07-2020/docs/D9B1BADE81B7AD61F557A33D673AA79C7809506872D8

Van der Grient, H. and de Haan, J. (2010). *The use of supermarket scanner data in the Dutch CPI*. CBS.

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.

Zhang, L., Johansen, I. and Nygaard R. (2019). *Evaluating multilateral price indices in a dynamic item universe*. Discussion Papers No. 914, September 2019 Statistics Norway, Research Department